

Paulo Jorge Leitão Pessoa Guerreiro

**OS GENES POLIUBIQUITINA NO CILIADO *Tetrahymena*:
ESTRUTURA, EXPRESSÃO E ANÁLISE COMPARATIVA
ENTRE AS ESPÉCIES *T.pyriformis* E *T.thermophila***

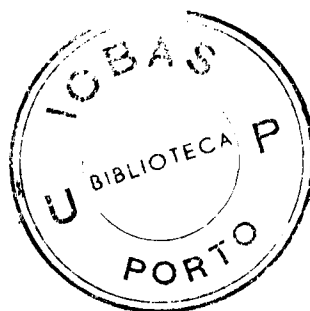
**PORTO
1995**

Paulo Jorge Leitão Pessoa Guerreiro

**OS GENES POLIUBIQUITINA NO CILIADO *Tetrahymena*:
ESTRUTURA, EXPRESSÃO E ANÁLISE COMPARATIVA
ENTRE AS ESPÉCIES *T.pyriformis* E *T.thermophila***

*Dissertação de candidatura ao
Grau de Doutor em Ciências
Biomédicas, Especialidade
Biologia Molecular (Genética
Molecular), apresentada ao
Instituto de Ciências Biomédicas
"Abel Salazar" da Universidade
do Porto.*

**PORTO
1995**



6664

À *Tetrahymena*, em particular
e aos ciliados, no geral

De acordo com o nº 2 do Artigo 8º do Decreto-Lei nº 388770, utilizaram-se neste trabalho resultados já publicados, ou em vias de publicação, que adiante se descrevem:

Guerreiro P., Neves A. & Rodrigues-Pousada C. (1995). The ubiquitin gene family in ciliates: *Tetrahymena pyriformis* has at least three different groups of polyubiquitin. Submetido a *Eur. J. Biochem.*

Guerreiro P., Neves A. & Rodrigues-Pousada C. (1993). Clusters of 5S rRNAs in the intergenic region of ubiquitin genes in *Tetrahymena pyriformis*. *Biochem. Biophys. Acta* **1216**, 137-139.

Neves, A., Guerreiro, P. & Rodrigues-Pousada, C. (1991). "The macronuclear polyubiquitin gene of the ciliate *Tetrahymena pyriformis*. *DNA Seq. - J. DNA Seq. Mapp.* **2**, 173-180.

Neves, A., Guerreiro, P. & Rodrigues-Pousada, C. (1990). Striking changes in polyubiquitin genes of *Tetrahymena pyriformis*. *Nucl. Ac. Res.* **18**, 656.

No cumprimento do disposto naquele Decreto-Lei, esclarecemos serem da nossa responsabilidade a execução das experiências que estiveram na base dos resultados aqui apresentados. São igualmente da nossa responsabilidade a interpretação dos resultados, sua discussão e redacção.

PRÓLOGO

A ubiquitina é, juntamente com a tubulina, actina e histonas, uma das proteínas intracelulares mais abundantes. Bastaria este facto para justificar a atenção que tem sido dada a esta proteína nas últimas duas décadas, mas os estudos efectuados têm demonstrado que a importância das funções da ubiquitina ultrapassa a sua abundância. Apesar de envolvida principalmente na degradação de proteínas de vida curta, a sua acção estende-se a diversos outros processos celulares, como sejam a transcrição e reparação do DNA, controlo do ciclo celular, "stress" celular e resposta hormonal, entre outros aspectos.

Esta multiplicidade funcional sugeriria, à partida, a existência de inúmeras variantes da proteína, mais ou menos especializadas nalgumas das suas funções. Na verdade, a descoberta dos genes codificantes para a ubiquitina parecia apoiar esta teoria. Constituindo uma família genética por vezes muito extensa, os genes codificantes para a ubiquitina são genes de fusão que se subdividem em duas classes: uma em que a unidade ubiquitina se encontra fundida a uma proteína ribossomal, e outra em que várias unidades ubiquitina se encontram fundidas entre si. Sendo assim, um organismo pode possuir um número elevado de unidades codificantes para a ubiquitina, que poderão originar de algumas unidades a várias dezenas de proteínas teoricamente distintas.

A realidade apresentar-se-ia, contudo, muito diferente. A ubiquitina é, a seguir à histona H4, a proteína mais conservada evolutivamente que se conhece. Em cada organismo, todas as suas unidades genéticas codificam para uma proteína idêntica; e em organismos tão distintos como animais, plantas, fungos e mesmo alguns protozoários, encontram-se apenas substituições em cerca de meia dúzia dos 76 resíduos de aminoácidos da ubiquitina. Assim, a especificidade das suas funções parece ser conferida não pela própria ubiquitina, mas pelas proteínas envolvidas no seu ciclo metabólico. De facto, as E2/UBC, as enzimas responsáveis pela transferência da ubiquitina para seus substratos, parecem ser as principais responsáveis por essa especificidade.

Apesar destes dados, a possibilidade de a ubiquitina apresentar variantes que lhe confirmam especificidade em determinadas funções não foi excluída, e tem sido apoiada por alguns dados recentes. Por um lado, têm sido encontrados em diversos organismos que vão desde fungos a animais vários genes que codificam para proteínas com óbvias semelhanças à ubiquitina. Estes genes apresentam por vezes estruturas semelhantes à dos genes ubiquitina, e para pelo menos um deles foi demonstrado que a proteína por ele codificada segue um ciclo metabólico semelhante ao ciclo da ubiquitina. A presença, em alguns vírus, de genes codificantes para proteínas semelhantes à ubiquitina tem sugerido igualmente que estas proteínas podem activar/bloquear no hospedeiro um conjunto de funções selecionado e favorável ao vírus. Finalmente, o último conjunto de dados tem sido conseguido junto dos organismos eucariotas que primeiro divergiram na escala evolutiva, principalmente unicelulares, como vários protozoários e algumas algas. Nestes organismos têm sido encontrados genes ubiquitina que permitem uma taxa de substituições muito superior à encontrada em organismos metazoários, mas não foi ainda determinado se as suas ubiquitinas possuem funções em tudo equivalentes às destes organismos ou se apresentam diferenças significativas.

O estudo da ubiquitina em organismos mais antigos apresenta assim diversos aliciantes. Sendo ubiquitinas "naturais", isto é, que terão um papel funcional equivalente nestes organismos e em organismos recentes, as diferenças observadas poderão então estar relacionadas ou com eventuais diferenças em algumas funções, ou com uma diferente adaptação da maquinaria do ciclo metabólico da ubiquitina a estas ubiquitinas diferentes. O

seu estudo permite assim determinar o modo como as substituições afectam a estrutura da proteína, de como esta estrutura interacciona com as outras enzimas do sistema, e de como essa diferente interacção afecta, ou não, as funções conhecidas da ubiquitina. Mais ainda, as diferenças observadas poderão ser correlacionadas com o desenvolvimento evolutivo das espécies, permitindo um melhor conhecimento da história evolutiva destas.

Dentro dos organismos unicelulares, o protozoário ciliado *Tetrahymena* é um organismo de eleição para o estudo da Biologia Molecular em organismos que mais cedo divergiram na árvore filogenética. Sob o ponto de vista laboratorial, trata-se de um organismo simples de cultivar, com um tempo de geração curto que permite uma manipulação fácil. Sendo um organismo de vida livre aquática, responde rapidamente a alterações do seu meio ambiente, como alterações de temperatura, minerais ou drogas, que são facilmente reprodutíveis em laboratório. Mas, principalmente, possui uma organização celular e um grau de diferenciação assinalável, como sejam a presença de um cortex complexo, que permite a formação de um aparelho oral e de um citoprocto, e a existência de dois núcleos diferenciados que permitem separar a fase vegetativa da fase gametogénica da célula. O grau de diferenciação deste organismo é tal que a *Tetrahymena* pode ser considerada com alguma justiça, a par de outros ciliados, como estando perto do limite evolutivo permitido a um organismo não metazoário.

O ciliado *Tetrahymena* possui assim diversas vantagens como modelo para o estudo da ubiquitina. O aumento dos níveis de ubiquitina é habitualmente uma das respostas celulares a alterações drásticas do ambiente, e, por isso, é particularmente marcante neste ciliado. Por outro lado, o grau de diferenciação celular que apresenta permite que a análise das funções da ubiquitina seja enquadrada na célula, facilitando a extrapolação dos resultados a organismos metazoários. Finalmente, os dados preliminares sobre a presença de ubiquitina em *Tetrahymena* mostram que não só ambas as classes de genes de ubiquitina se encontram presentes, mas igualmente que uma das classes - genes poliubiquitina - compreende uma família de genes em que alguns dos seus membros codificam ubiquitinas idênticas e semelhantes às encontradas em organismos metazoários, enquanto outros codificam para ubiquitinas distintas das primeiras e entre si. Neste sentido, o ciliado *Tetrahymena* parece possuir em simultâneo um conjunto de genes característico de organismos metazoários e outro de organismos unicelulares menos evoluídos, pelo que se tornava importante aprofundar o estudo sobre esta classe de genes, analisar a sua expressão e eventuais funções das proteínas por eles codificadas, objectivo final desta tese.

Esta dissertação foi então organizada em quatro capítulos. O primeiro, denominado Introdução, foi dividido em duas partes, sendo a primeira uma revisão dos conhecimentos actuais sobre a ubiquitina, e a segunda uma revisão dos conhecimentos actuais sobre os ciliados, focando-se em particular o género *Tetrahymena*. Na revisão sobre a ubiquitina, analisa-se a sua estrutura e o principal ciclo metabólico em que se encontra envolvida, bem como as outras enzimas nele envolvidas e os principais processos celulares em que se envolve. São igualmente focados os genes que codificam esta proteína, fazendo-se um estudo mais exaustivo sobre a sua organização nos diversos organismos onde já foram caracterizados, e dos principais fenómenos que regulam a sua expressão. Um parágrafo foi dedicado em especial aos genes que apresentam semelhanças com os genes que codificam para a ubiquitina. Na segunda parte da Introdução são apresentadas as principais características dos ciliados e as suas principais contribuições para a Biologia Molecular. O ciclo de vida da *Tetrahymena* é analisado com algum detalhe, sendo aprofundada a parte respeitante ao desenvolvimento do macronúcleo, que é comparada com o observado noutros ciliados.

Os métodos utilizados na realização deste trabalhos são descritos no segundo capítulo, tendo-se procurado que esta descrição permitisse seguir e reproduzir o trabalho apresentado.

Embora alguns genes poliubiquitina tivessem já sido clonados e caracterizados em *Tetrahymena pyriformis*, a sua organização genómica não era conhecida, e o conhecimento sobre os gene poliubiquitina divergentes era ainda deficiente. O terceiro capítulo contém os resultados respeitantes à clonagem e caracterização de cinco destes genes, dois dos quais apenas parcialmente descritos anteriormente. Estes genes foram comparados com os genes já conhecidos, analisando-se em particular as diferenças estruturais e evolutivas encontradas, e apresentando-se hipóteses que as justifiquem. A organização genómica da família poliubiquitina é examinada, sendo apresentado um primeiro modelo.

O trabalho efectuado com os genes poliubiquitina divergentes encontrados em *T.pyriformis* levantou diversas questões para cujas respostas se considerou ser necessário conhecer a organização destes genes numa espécie relacionada mas possuidora de ambos os núcleos e, por isso, capaz de reprodução sexuada. O quarto capítulo apresenta um trabalho preliminar efectuado numa espécie com essas características, *Tetrahymena thermophila*. Foi clonado e caracterizado um gene poliubiquitina, e avançam-se com alguns dados sobre a organização genómica desta classe de genes neste ciliado. Foram igualmente analisadas as diferenças de expressão destes genes em ambos os organismos, quer em condições de crescimento exponencial quer em condições de choque hipertérmico, bem como a sua expressão durante a fase de reprodução sexuada em *T.thermophila*.

AGRADECIMENTOS

Longe vão os tempos em que a Ciência se fazia durante os nossos tempos livres num quarto desocupado de casa, equipado com umas provetas e retortas mandadas fazer sob encomenda ao vidreiro da aldeia. Hoje, um trabalho científico necessita de um laboratório bem equipado, uma Instituição de apoio, e é fundamentalmente um trabalho de equipa. Não faz sentido, pois, apresentar esta dissertação sem uma palavra de agradecimento a todos, Instituições e pessoas, que em maior ou menor grau permitiram a sua realização.

À Fundação Calouste Gulbenkian, que através do seu Instituto Gulbenkian de Ciência me proporcionou as excelentes condições que permitiram a realização deste trabalho.

À Junta Nacional de Investigação Científica e Tecnológica, que através dos programas Ciência e Praxis XXI me concedeu a Bolsa de Doutoramento que me permitiu dedicar-me por completo à realização deste trabalho.

À Professora Doutora Claudina Rodrigues-Pousada agradeço a confiança que sempre depositou em mim, mesmo nos períodos áureos da minha teimosia. Agradeço igualmente a orientação científica e pedagógica que me proporcionou, traduzidas numa disponibilidade e apoios constantes, e cujo alcance vai muito mais além do que a simples realização desta dissertação.

À Ana Neves, companheira no mundo das ubiquitinas, agradeço a introdução às técnicas laboratoriais, e o apoio, espírito crítico e disponibilidade para a análise dos sucessos e insucessos deste trabalho. Mas também a amizade, o método de trabalho que ainda hoje me conduz, e a sua capacidade de conciliar pontos de vista diferentes.

À Helena Soares, muito mais do que a amizade, apoio e disponibilidade constantes, agradeço o partilhar de um sonho já um pouco realidade - o de se ser cientista - e de uma paixão - a dos ciliados - que levam a que a Ciência, afinal, até tenha algo de poesia.

À Professora Doutora Ana Margarida Damas, agradeço a ajuda indispensável na análise da estrutura terciária da ubiquitina.

À Alexandra, pela amizade, ajuda na subclonagem do fragmento "inclonável" e principalmente pela paciência que teve para me ler e corrigir os erros desta tese.

À Luisa Cyrne, pela amizade, apoio e ensinamentos vários que me transmitiu ao longo deste trabalho.

À Lisete Fernandes, pela amizade e apoio que sempre me proporcionou, mas também pelo espírito prático que possuí, que me levava a ver sempre o outro lado da questão.

À Cristina Vilela, agradeço a simpatia e confiança que sempre depositou em mim.

Ao Rui Córias, por toda a discussão e comentários sobre os resultados deste trabalho.

À Tatiana Vassilievskaya, pelas amenas conversas de fim-de-tarde, e ao Orfeu Flores, pela boa disposição e discussões clubísticas, porque nem só de Ciência vive o homem.

A todos os meus amigos e colegas que de uma forma ou outra me ajudaram na realização deste trabalho.

Aos meus professores, cujos ensinamentos me permitiram embarcar na vida científica, e em particular ao Professor Doutor Pedro Moradas Ferreira, que me deu o primeiro "bote" ao introduzir-me bem antes do meu estágio de licenciatura nos trabalhos laboratoriais e na manipulação da *Tetrahymena*.

Aos meus pais, irmãos e restante família, um agradecimento muito especial porque, se aqui cheguei, a eles o devo.

À Paula, o ter aceite partilhar a sua vida com um "homem dos bichinhos", o que não é propriamente fácil, e de ainda assim o apoiar e incentivar sempre.

E ao Bruno, que ainda não me viu, mas que já me está a aturar!

SÍMBOLOS E ABREVIATURAS

ψ gene	- pseudogene
A	- nucleótido de adenina
aa	- aminoácido
agrup.	- agrupados
AMP	- monofosfato de adenosina
ATP	- 5'-trifosfato de adenosina
bis-acrilamida	- N,N'-metileno-bis-acrilamida
BSA	- albumina sérica bovina
C	- nucleótido de citosina
c.	- cerca de
C-terminal	- extremidade da cadeia polipeptídica com o grupo α -carboxílico livre
Cbs	- sequência de clivagem do cromossoma
CCT	- chaperonina-contendo-TCP1
cDNA	- DNA complementar
cel.	- células
CTP	- 5'-trifosfato de citidina
dATP	- 5'-trifosfato de desoxiadenosina
dCTP	- 5'-trifosfato de desoxicitidina
DEPC	- dietilpirocarbonato
dGTP	- 5'-trifosfato de desoxiguanosina
DMSO	- dimetilsulfóxido
DNA	- ácido desoxirribonucleico
DNase	- desoxirribonuclease
dNTP	- 5'-trifosfato de desoxinucleósido
ds	- DNA de cadeia dupla
DTT	- ditioreitol
dTTP	- 5'-trifosfato de timidina
E1	- enzima activadora da ubiquitina (=UBA)
E2	- enzima de conjugação da ubiquitina (=UBC)
E3	- ligase da ubiquitina
EDTA	- etilenodiaminatetracetato
est	- (fase) estacionária
exp	- (crescimento) exponencial
fgm	- fragmento
G	- nucleótido de guanina
GTP	- 5'-trifosfato de guanosina
HS	- choque hipertérmico
HSE	- elemento de choque térmico
IES	- sequências internas eliminadas
inc.	- incompletos
IPTG	- isopropilo- β -D-galactósido
kb	- quilopares de bases
kDa	- quiloDalton
MDS	- sequências destinadas ao macronúcleo
MOPS	- (3-[N-morfolino]propanosulfonato
mRNA	- RNA mensageiro

N	- nucleótido
NMR	- ressonância magnética nuclear
N-terminal	- extremidade da cadeia polipeptídica com o grupo α -amino livre
NTP	- 5'-trifosfato de nucleósido
O.D.	- densidade óptica
pags	- páginas
pb	- pares de bases
PCR	- reacção de cadeias múltiplas (de "polichain reaction")
PEG	- polietilenoglicol
pfu	- unidade de placas fágicas
PPi	- pirofosfato inorgânico
PSA	- persulfato de amónio
pTU	- plasmídeo recombinante contendo sequências ubiquitina de <i>T.pyriformis</i>
pTTU	- plasmídeo recombinante contendo sequências ubiquitina de <i>T.thermophila</i>
R	- nucleótido de purina
rDNA	- DNA codificante para rRNA
RNA	- ácido ribonucleico
RNA poli(A) ⁺	- RNA com cauda poli(A) na extremidade 3'
RNA poli(A) ⁻	- RNA sem cauda poli(A) na extremidade 3'
RNase	- ribonuclease
rpm	- rotações por minuto
rRNA	- RNA ribossomal
SDS	- dodecilsulfato de sódio
spp	- specimen, espécie
ss	- DNA de cadeia simples
SSC	- tampão sal/citrato de sódio [150 mM NaCl; 15 mM citrato de sódio]
T	- nucleótido de timina
TCA	- ácido tricloroacético
TEMED	- <i>N,N,N',N'</i> -tetraetilenodiamina
Tris-HCl	- tris(hidroximetil)aminometano ajustado ao pH indicado com HCl
tRNA	- RNA de transferência
TU	- fragmento de DNA contendo sequências ubiquitna de <i>T.pyriformis</i>
TTU	- fragmento de DNA contendo sequências ubiquitna de <i>T.thermophila</i>
U	- unidades enzimáticas
u	- unidades ou repetições ubiquitina
ub	- ubiquitina
Ub52	- gene de fusão ubiquitina-pequena proteína ribossomal
Ub76	- gene de fusão ubiquitina-grande proteína ribossomal
UBA	- enzima activadora da ubiquitina (=E1)
UBC	- enzima de conjugação da ubiquitina (=E2)
uH	- histona ubiquitinada (ex: uH2A, uH2B)
UTP	- 5'-trifosfato de uracilo
UV	- ultravioleta
X-Gal	- 5-bromo-4-cloro-3-indoil- β -D-galactósido
Y	- nucleótido de pirimidina

RESUMO

A ubiquitina é uma pequena proteína eucariota muito abundante na célula, onde pode existir livre ou ligada covalentemente a vários substratos proteicos. A sua principal função desenrola-se no citoplasma, actuando na degradação de proteínas anormais ou de vida curta num processo dependente do ATP, mas encontra-se envolvida em diversos outros processos celulares como por exemplo a transcrição e reparação do DNA, resposta ao "stress" e resposta hormonal.

Os genes ubiquitina são genes de fusão que se dividem em duas classes: uma onde uma unidade ubiquitina se encontra fundida a uma proteína ribossomal, e outra onde várias unidades se encontram fundidas entre si. Os genes desta última classe, ou genes poliubiquitina, encontram-se presentes em todos os organismos estudados até agora, possuindo habitualmente de três a cerca de uma dezena de unidades. A presença destes genes em protozoários ciliados foi demonstrada pela primeira vez no ciliado *Tetrahymena pyriformis*, tendo-se clonado e caracterizado vários genes deste organismo. A análise das suas sequências nucleótídicas e de aminoácidos sugere que podem ser classificados em dois grupos principais contendo vários genes cada, e um terceiro contendo um único gene que poderá estar relacionado com o segundo grupo.

O primeiro grupo é constituído por genes poliubiquitina que codificam proteínas idênticas, muito semelhantes às encontradas noutros organismos. Este grupo é muito homogêneo nas suas sequências nucleotídicas, apresenta um a dois codões extra após a última unidade de cada gene. Existe uma preferência clara na utilização de codões, próxima da encontrada em genes muito expressos, como tubulinas e histonas. Os seus genes transcrevem pelo menos três mensageiros, e são regulados por agentes de "stress".

O segundo grupo apresenta genes poliubiquitina de características particulares. Todas as suas unidades são diferentes entre si e com as do primeiro grupo, não apresentando igualmente a Gly76 da última unidade de cada gene, essencial para a função da ubiquitina. Apesar da elevada taxa de substituições observada, vários dados sugerem que este grupo foi ou tem sido sujeito a uma pressão selectiva, e que não podem ser considerados pseudogenes do primeiro grupo. De facto, não se observam mudanças de grelha de leitura ou codões de terminação internos nos genes deste grupo, nem homologia entre as regiões flanqueantes não codificantes dos dois grupos. A taxa de substituições de nucleótidos é muito superior à de aminoácidos, o que se traduz numa utilização de codões distinta, mesmo em regiões onde a sequência de aminoácidos é absolutamente conservada entre os genes dos dois grupos. Finalmente, as substituições de aminoácidos encontram-se localizadas em regiões demarcadas da sequência primária da ubiquitina, e não só não afectam o seu perfil de hidropatia, como se localizam igualmente nas mesmas superfícies da estrutura terciária da molécula. Apesar destes dados, podem-se observar igualmente várias substituições não conservadas, que podem afectar o funcionamento da proteína. Por outro lado, não foram detectados sinais de actividade nestes genes nas diferentes condições estudadas, o que sugere que estes genes poderão ter perdido recentemente a sua funcionalidade.

Em *T. pyriformis*, os genes poliubiquitina constituem uma família de pelo menos onze genes, sendo cinco pertencente ao primeiro grupo e seis ou mais ao segundo. A análise da sua organização genómica revela que se encontram dispostos em série, mas que os genes do primeiro grupo se encontram localizados numa região do genoma distinta da dos do segundo grupo.

É possível que a presença de um grupo de genes evolucionariamente distinto mas não funcional em *T. pyriformis* possa estar relacionada com a perda e/ou ausência do

micronúcleo neste organismo. Neste sentido, analisou-se a sua presença numa espécie micronucleada, *T.thermophila*, capaz por isso de reprodução sexuada. A análise preliminar da sua organização genómica sugere que este organismo contém um número de genes poliubiquitina semelhante ao encontrado em *T.pyriformis* igualmente dispostos em série. Um destes genes foi clonado e caracterizado, revelando possuir todas as particularidades dos genes do primeiro grupo de *T.pyriformis*. Não foi caracterizado nenhum gene do segundo grupo, mas obtiveram-se resultados que sugerem que este pode estar presente em *T.thermophila*.

A expressão dos genes poliubiquitina foi analisada em *T.thermophila* de modo a aquilatar eventuais diferenças entre os dois organismos. Ambos os organismos possuem um padrão de expressão equivalente, apresentando um conjunto de mensageiros expressos durante o crescimento exponencial das células e induzido durante situações de "stress" hipertérmico, um segundo principalmente expresso durante situações de "stress" hipertémico, e um terceiro com uma expressão residual em ambas as situações. Contudo, em *T.thermophila* os genes principalmente transcritos durante o "stress" hipertérmico possuem aparentemente menos uma unidade ubiquitina que os genes equivalentes em *T.pyriformis*, enquanto os genes residualmente transcritos possuem aparentemente uma unidade a mais. Durante a reprodução sexuada das células, apenas o primeiro conjunto de genes é transcrito, não se observando diferenças significativas relativamente ao crescimento exponencial. Todavia, é possível observar-se um decréscimo dos níveis dos mensageiros codificantes para a ubiquitina de fusão com a proteína ribossomal.

ABSTRACT

Ubiquitin is a small eukaryotic protein of high intracellular abundance, where it exists in either free or covalently joined to various intracellular protein substrates. In the cytoplasm it is involved in the ATP-dependent proteolysis of abnormal and short-lived proteins, but ubiquitin has in addition other cellular roles, as for example in the DNA transcription and repair, stress and hormonal responses.

The ubiquitin genes are fusion genes that fall into two classes: one where one ubiquitin unit is fused to a ribosomal protein and other where several ubiquitin units are lined up head-to-tail. These genes, or polyubiquitin genes, are present in all organisms examined to date, with a number of repeats usually varying from three to about ten. The presence of polyubiquitin genes in ciliates was first demonstrated in *Tetrahymena pyriformis*, being several genes cloned and characterised. Their nucleotide and aminoacid sequence analysis suggest that they may be classified into two major groups with several genes each, and a third one that contains only one gene that may be related with the second group.

The first group is composed by polyubiquitin genes that code for identical proteins related to those found in other organisms. This group is highly similar at nucleotide level, and each gene presents one or two extra codons in the last ubiquitin unit. Their codon usage is strongly biased, with values corresponding to the ones of highly expressed genes, such the ones coding for tubulins and histones. They code indeed for at least three mRNAs and are under regulation by stress.

The second group is composed by polyubiquitin genes showing several striking features. All their units are different from each other or from units of the first group, and the last unit of each gene lacks the Gly76 codon, known to be essential for the ubiquitin function. Besides the high rate of mutations observed, the data suggest that this group is, or was until recently, subjected to an evolutionary pressure and cannot be accounted as pseudogenes related to the first group of polyubiquitins. Their genes do not present internal stop codons or loss of the reading frame, neither their non-coding regions are homologous to those of the first group. The rate of nucleotide substitutions is higher than that of aminoacid substitutions, leading to a different codon usage even in regions where the aminoacid sequence of the two groups is strictly conserved. Finally, the observed aminoacid mutations are clustered in discrete regions of the ubiquitin molecule and in the same regions of the tertiary structure of the molecule, and do not change the hydropathic profile obtained. However, it is also possible to observe some random and not conserved mutations that may affect the ubiquitin function. Moreover, no mRNA transcription was detected for these genes in the several conditions analysed, suggesting that they may have lost recently their functionality.

The polyubiquitin genes in *T.pyriformis* are a family of at least eleven genes, being five from the first group and six or more from the second one. The analysis of their genomic organisation shows that they are grouped in tandem, but with the genes of the first group localised in a different genomic region of that of the second group.

It is possible that the presence of a non-functional unrelated group of polyubiquitin genes in *T.pyriformis* could be related with the loss and/or lack of the micronuclei in this ciliate. A micronucleated species, *T.thermophila*, a specie with sexual reproduction, was then analysed for the presence of this group. The preliminary analysis of its genomic organisation suggests that this ciliate has a number of polyubiquitin genes similar to that found in *T.pyriformis*, which are also grouped in tandem. One gene was cloned and

characterised, showing all the first group features. None of the second group gene was characterised, but the results obtained suggest that they may be present in *T.thermophila*.

The ubiquitin expression was analysed in *T.thermophila* and compared with those of *T.pyriformis*. Both ciliates show a similar pattern of expression, presenting one band of mRNAs expressed in exponentially growing cells that are induced during heat stress, a second one of mRNAs mainly expressed during heat stress, and a third one of mRNAs barely expressed in both situations. However, the genes mainly transcribed during the heat stress in *T.thermophila* have apparently one unit less than those of *T.pyriformis*, whereas the genes barely transcribed have apparently one unit more. During sexual reproduction, only the first band is observed without significant differences with respect to the control. However, the mRNA levels of the ubiquitin-ribosomal protein fusion clearly decrease during this period.

RÉSUMÉ

L'ubiquitine est une petite protéine eucariotique très abondante dans la cellule, où elle existe libre ou liée d'une façon covalente à plusieurs substrats protéiques. Sa principale fonction a lieu dans le cytoplasme, où elle agit sur la dégradation de protéines anormales ou de courte demi-vie dans un processus dépendant de l'ATP, mais elle joue aussi un rôle sur d'autres processus cellulaires, comme la transcription et réparation de l'ADN, la réponse au stress et la réponse hormonale.

Les gènes ubiquitine sont des gènes de fusion qui se divisent en deux classes: l'une, où une unité ubiquitine est fusionnée à une protéine ribosomale, et l'autre où plusieurs unités sont fusionnées entre elles. Les gènes de cette dernière classe, aussi appelés gènes polyubiquitine, sont présents dans tous les organismes étudiés jusqu'à maintenant, et il y en a entre trois à environ une dizaine d'unités. La présence de ces gènes chez les protozoaires ciliés a été mise en évidence par la première fois chez le cilié *Tetrahymena pyriformis*, où ont été clonés et caractérisés plusieurs gènes. L'analyse de leurs séquences nucléotidiques et polypeptidiques suggère qu'ils peuvent être classés en deux groupes principaux qui contiennent plusieurs gènes chacun, et encore dans un troisième avec un seul gène qui pourra être en rapport avec le deuxième groupe.

Le premier groupe présente des gènes polyubiquitine qui codifient des protéines égales qui sont très semblables à celles trouvées chez les autres organismes. Ce groupe est très homogène dans ses séquences nucléotidiques, et il présente un ou deux codons extra après la dernière unité de chaque gène. Il y a une préférence nette dans l'utilisation des codons qui est très proche de celle trouvée pour les gènes très exprimés, comme les tubulines et les histones. Les gènes de ce groupe transcrivent au moins trois ARNm, et ils sont régulés par des agents de stress.

Le deuxième groupe présente des gènes polyubiquitine avec des caractéristiques très particulières. Toutes leurs unités sont différentes entre elles ou avec les unités du premier groupe, et il n'y existe pas la Gly76 de la dernière unité de chaque gène, qui est essentielle à la fonction de l'ubiquitine. Malgré le haut taux de substitutions observé, plusieurs données suggèrent que ce groupe est ou il a été soumis à une pression sélective, ne pouvant pas ainsi être considérés comme pseudogènes du premier groupe. En effet, on n'observe aucun changement de phase de lecture ou la présence de codons de terminaison internes dans les gènes de ce groupe, ni aucune homologie entre les régions flanquantes et non-codantes de ces deux groupes. Le taux de substitutions de nucléotides est bien supérieur à celui des acides aminés, ce qui se traduit dans une utilisation de codons distincte même dans les régions où la séquence d'acides aminés est absolument conservée entre les gènes de ces deux groupes. Finalement, les substitutions d'acides aminés sont localisées sur des régions définies dans la séquence primaire de l'ubiquitine, et non seulement elles n'affectent pas leur profil de hydrophobie, comme elles sont également localisées sur les mêmes surfaces de la structure tertiaire de la molécule. Malgré ces données, on peut observer aussi quelques substitutions qui ne sont pas conservées et qui peuvent affecter le fonctionnement de la protéine. Par contre, on n'a trouvé aucun signal d'activité à ces gènes dans les différentes conditions étudiées, ce qui suggère que ces gènes pourront avoir perdu récemment leur fonctionnalité.

Chez *T. pyriformis*, les gènes polyubiquitine font partie d'une famille de au moins onze gènes, cinq pour le premier groupe et six ou plus pour le deuxième. L'analyse de son organisation génomique montre qu'ils sont disposés en série, mais aussi que les gènes du

premier groupe se localisent dans une région du génome différente de celle du deuxième groupe.

C'est possible que la présence d'un groupe de gènes evolutionairement différente mais aussi non fonctionel chez *T.pyriformis* soit en relation avec l'elimination et/ou absence du micronoyau chez cet organisme. Dans ce but, on a analysé sa présence chez une espèce micronucléé, *T.thermophila*, qui est capable de reproduction sexuée. L'analyse preliminaire de son organization génomique suggère que cet organisme contient un nombre de gènes polyubiquitine semblables aux trouvés chez *T.pyriformis*, également disposés en série. L'un de ces gènes a été cloné et caracterisé, révélant avoir toutes les particularités des gènes du premier groupe de *T.pyriformis*. Aucun gène du deuxième groupe a été caracterisé, mais les resultats obtenues suggèrent que ce groupe peut être présent chez *T.thermophila*.

L'expression des gènes ubiquitine a été analysée chez *T.thermophila* ayant le but de chercher les différences possibles parmi les deux organismes. Tous les deux ont un profile d'expression semblable, avec une première population de messagers exprimés pendant la croissance exponentielle des cellules et qui est augmentée lors d'un choc thermique, une deuxième qui est exprimé surtout lors d'un choc thermique, et une troisième qui a une expression résiduelle dans les deux situations. Cependant, chez *T.thermophila* les gènes qui sont transcrits surtout lors d'un choc thermique ont apparemment une unité ubiquitine en moins que les gènes equivalents chez *T.pyriformis*, tandis que les gènes qui ont une expression résiduelle ont une unité en plus. Pendant la reproduction sexuelle des cellules, seulement la première population est transcrite, et on n'observe pas de différences significatives par rapport aux cellules en croissance exponentielle. Néanmoins, on peut même observer une décroissance des niveaux de messagers codants pour l'ubiquitine de fusion avec la protéine ribosomale.

ÍNDICE

PRÓLOGO	vii
AGRADECIMENTOS.....	xi
SÍMBOLOS E ABREVIATURAS.....	xiii
RESUMO	xv
ABSTRACT	xvii
RÉSUMÉ.....	xix
ÍNDICE.....	xxi

CAPÍTULO I - INTRODUÇÃO 1

O sistema ubiquitina	1
1. A ubiquitina: breve história.	3
2. O sistema ubiquitina.....	3
2.1. A ubiquitina: propriedades estruturais.	3
2.2. Conjugação da ubiquitina.	4
2.3. Degradação proteica mediada pelo sistema ubiquitina.	6
2.4. Enzimas do sistema ubiquitina.	8
2.4.1. Hidrolases C-terminais da ubiquitina.....	8
2.4.2. E1/UBA - Enzimas de activação da ubiquitina.	9
2.4.3. E2/UBC - Enzimas de conjugação da ubiquitina.	9
2.4.4. E3 - Ligases da ubiquitina.	10
2.4.5. O proteassoma.	11
3. Funções da ubiquitinação.	12
3.1. Reparação do DNA.	13
3.2. Controlo do ciclo celular.....	14
3.3. Transcrição.....	15
3.4. Resposta ao "stress" celular.....	16
3.5. Ubiquitinação de proteínas membranares.....	17
3.6. Infecção viral.	18
3.7. Outros aspectos da ubiquitinação.	19
4. Os genes ubiquitina.....	26
4.1 Estrutura dos genes ubiquitina.....	26
4.1.1. Classe I: genes poliubiquitina.	28
4.1.2. Classe II: genes ubiquitina-proteínas ribossomais.	31
4.2. Genes relacionados aos genes ubiquitina.....	34
4.3. Organização genómica e expressão dos genes ubiquitina.....	36

Capítulo I - Introdução 2

O ciliado <i>Tetrahymena</i>	41
1. Caracterização do ciliado <i>Tetrahymena</i>	43
1.1. Introdução ao phylum Ciliophora	43
1.2. Características e distribuição do género <i>Tetrahymena</i>	44
2. O ciclo de vida em <i>Tetrahymena</i>	46
2.1. O dimorfismo nuclear em ciliados	46
2.2. O ciclo de vida em <i>Tetrahymena</i>	47
2.2.1. O ciclo vegetativo	47
2.2.2. O ciclo sexual ou conjugação	48
2.2.2.1. A conjugação normal	48
2.2.2.2. Desvios importantes à conjugação normal	51
2.2.2.3. Os tipos de "mating" em <i>Tetrahymena</i>	53
2.3. O desenvolvimento macronuclear	55
2.3.1. Em <i>Tetrahymena</i>	55
2.3.2. Em ciliados Hipotricos	58
3. Outras contribuições nos conceitos da biologia molecular	60
3.1. O código genético	60
3.2. Os genes de rRNA e o conceito de autoexcisão ("self-splicing")	61
3.3. Os telómeros	62

CAPÍTULO II - MATERIAL E MÉTODOS 65

1. Células e condições de cultura	67
1.1. Células eucariotas: <i>Tetrahymena</i> spp.	67
1.1.1. Espécies e estirpes usadas	67
1.1.2. Condições de cultura	67
1.2. Células bacterianas: <i>Escherichia coli</i>	68
1.2.1. Estirpes usadas	68
1.2.2. Meios e condições de cultura	69
2. Vectores de clonagem	70
2.1. Vectores plasmídicos	70
2.2. Vectores fágicos: o fago M13 e seus derivados	70
2.3. Vectores fágicos: o fago λ DashII	71
3. Preparação e análise de DNA de <i>Tetrahymena</i>	71
3.1. Preparação do DNA nuclear	71
3.2. Análise do DNA nuclear (Southern)	72
3.2.1. Hidrólise e separação do DNA por electroforese em geis de agarose	72
3.2.2. Imobilização do DNA	72
3.2.3. Preparação de sondas radioactivas	73
3.2.4. Condições de hibridação dos filtros	74
3.2.4.1. Hibridação com sondas não oligoméricas	74
3.2.4.2. Hibridação com sondas oligoméricas	74

4. Preparação e análise do RNA de <i>Tetrahymena</i>.	75
4.1. Extracção de RNA.	75
4.1.1. Extracção de RNA citoplasmático.	75
4.1.2. Isolamento de mRNA poli(A)+.	75
4.2. Análise do RNA citoplasmático (Northern).	76
4.2.1. Separação do RNA por electroforese em geis de agarose.	76
4.2.2. Imobilização do RNA.	76
4.2.3. Condições de hibridação dos filtros.	77
4.3. Análise da transcrição aparente ("run on").	77
4.3.1. Preparação de núcleos de <i>Tetrahymena</i> .	77
4.3.2. Imobilização das sondas de DNA.	78
4.3.3. Transcrição nuclear e isolamento de RNA.	78
4.3.4. Condições de hibridação dos filtros.	78
5. Construção de bancos de DNA de <i>Tetrahymena</i>.	79
5.1. Construção de bancos de DNA de <i>Tetrahymena pyriformis</i> em plasmídeos.	79
5.1.1. Banco total de fragmentos <i>Hind</i> III em pBR322.	79
5.1.2. Banco parcial de fragmentos <i>Cla</i> I em pUC19.	79
5.2. Construção de bancos totais em λ DashII.	79
5.2.1. Bancos de DNA total de <i>Tetrahymena</i> parcialmente digerido com <i>Sau</i> 3AI.	79
5.2.2. Selecção dos clones recombinantes.	80
5.2.3. Preparação de DNA de fagos λ .	81
5.3. Análise dos clones recombinantes.	82
6. Subclonagem em plasmídeo.	82
6.1. Preparação de fragmento.	82
6.2. Ligação do fragmento ao plasmídeo.	83
6.3. Transformação de células <i>E. coli</i> .	83
6.3.1. Preparação de células competentes.	83
6.3.2. Transformação de células competentes.	84
6.4. Preparação de DNA plasmídico.	84
6.4.1. Extracção de DNA em pequena escala (miniprep).	84
6.4.2. Extracção de DNA em média escala (mediprep).	85
6.4.3. Extracção do DNA em grande escala.	86
6.5. Análise dos subclones obtidos.	86
7. Determinação da sequência nucleotídica do DNA.	86
7.1. Subclonagem em fagos M13.	86
7.1.1. Geração de subclones por clonagem directa ou aleatória.	86
7.1.2. Geração de subclones por deleção parcial do fragmento inserido.	86
7.2. Transformação com o fago M13.	87
7.3. Preparação e análise dos fagos recombinantes.	88
7.3.1. Preparação da cadeia simples (ss) dos fagos.	88
7.3.2. Preparação da cadeia dupla (ds) dos fagos.	88
7.3.3. Análise dos fagos recombinantes.	88
7.4. Sequenciação do DNA de cadeia simples.	89
7.5. Análise das sequências nucleotídicas.	89

CAPÍTULO III - RESULTADOS 1

Genes poliubiquitina em <i>T.pyriformis</i>	91
1. Isolamento de genes poliubiquitina em <i>Tetrahymena pyriformis</i>	93
2. Análise dos genes poliubiquitina	93
2.1. Análise das sequências codificantes	93
2.2. Análise das sequências não codificantes.	95
2.3. Utilização de códons pelos genes poliubiquitina.	99
3. Análise comparativa entre ubiquitinas de diferentes espécies	102
4. Análise da estrutura da molécula de ubiquitina de <i>Tetrahymena pyriformis</i>	104
4.1. Perfis de hidropatia.	104
4.2. Estrutura terciária.	106
5. Expressão dos genes poliubiquitina	107
6. Organização genômica dos genes poliubiquitina de <i>Tetrahymena pyriformis</i>	108
7. Discussão	111

CAPÍTULO IV - RESULTADOS 2

Genes poliubiquitina em <i>T.thermophila</i>	115
1. Análise de sequências ubiquitina no DNA genômico de <i>T.thermophila</i>	117
2. Isolamento e caracterização dos genes poliubiquitina em <i>T.thermophila</i>	119
2.1. Isolamento de genes ubiquitina	119
2.2. Análise do gene poliubiquitina TTU3.	120
3. Expressão dos genes ubiquitina em <i>T.thermophila</i>	122
4. Discussão	124

CONCLUSÕES E PERSPECTIVAS	127
BIBLIOGRAFIA	131

APÊNDICE 1 - Clusters of 5S rRNAs in the intergenic region of ubiquitin genes in <i>Tetrahymena pyriformis</i>	165
APÊNDICE 2 - Sequenciação parcial e análise da subunidade 2 da RNA polimerase II de <i>Tetrahymena thermophila</i>	171

Capítulo I

INTRODUÇÃO 1

O sistema ubiquitina

1. A ubiquitina: breve história.

Durante a sua evolução as células eucariotas desenvolveram diversos mecanismos reguladores dos níveis de proteína intracelulares. Estes mecanismos regulam quer a sua síntese, desde a transcrição de DNA em mRNA até à tradução deste, quer a sua degradação. De facto, dentro de uma célula, coexistem proteínas de vida curta, com semividas de poucos minutos a poucas horas, até proteínas de vida longa cujas semividas podem atingir vários dias. Esta diferença na estabilidade proteica requer um controlo preciso da selectividade do processo degradativo.

As células eucariotas desenvolveram diferentes vias para a degradação de proteínas. Em jejum, a degradação proteica é efectuada principalmente por via lisossomal de uma maneira não selectiva, embora tenham sido detectadas em algumas proteínas sequências que poderão regular a sua degradação lisossomal (Dice, 1990). Outra via degradativa, aparentemente responsável pela eliminação de proteínas membranares recém-sintetizadas mal formadas, encontra-se localizada no retículo endoplasmático (Klausner & Sitia, 1990). A principal via degradativa para proteínas de vida curta, anormais ou mal formadas é todavia citosólica, e dependente de energia metabólica. A elucidação deste sistema iniciou-se há quinze anos, quando foi posta em evidência a existência de um factor proteico necessário à actividade de um sistema proteolítico dependente de ATP em reticulócitos (Ciechanover *et al.*, 1978). Dois anos depois, esse factor foi isolado (Wilkinson *et al.*, 1980), e identificado com a proteína descrita em 1975 por Goldstein e colaboradores capaz de induzir a diferenciação de linfócitos B. Neste mesmo ano, já Schlesinger e colaboradores (1975) publicara a sequência desse factor, que, inicialmente denominado de "ubiquitous immunopoietic peptide" devido à sua detecção por anticorpos em células animais, de plantas e de levedura, passou então a ser designado pelo nome actual - ubiquitina.

O interesse pela ubiquitina aumentou desde então. Tal como a actina, a tubulina e as histonas, é uma das proteínas intracelulares mais abundantes. Inicialmente apenas implicada no processo de proteólise celular, a sua acção atinge os mais diversos níveis da actividade celular, como a reparação do DNA, controlo do ciclo celular, resposta ao "stress", modificação de histonas e receptores celulares, infecção viral, etc. Este interesse crescente por esta proteína é facilmente evidenciado pelo largo número de publicações surgidas nos últimos anos, incluindo diversas revisões (Rechsteiner, 1987; Jentsch *et al.*, 1991; Finley & Chau, 1991; Hershko & Ciechanover, 1992; Jentsch, 1992), mini-revisões (Jentsch *et al.*, 1990; Hershko, 1991; Hochstrasser, 1992; Sommer & Seufert, 1992) e monografias (Rechsteiner, 1988; Schlesinger & Hershko, 1988).

2. O sistema ubiquitina.

2.1. A ubiquitina: propriedades estruturais.

A ubiquitina é uma pequena proteína de 76 aminoácidos, de ponto isoeléctrico $pI=6.7$ e com carga praticamente neutra a valores de pH inferiores a 9 (Low *et al.*, 1979). A sua estrutura terciária foi já determinada por cristalografia de raios X a uma resolução de 1.8 Å (Vijay-Kumar *et al.*, 1985; 1987a) e por NMR bidimensional (Weber *et al.*, 1987). A ubiquitina possui uma estrutura globular extremamente compacta, da qual apenas se destaca a região C-terminal (Fig.11.1). A molécula contém três voltas e meia em α -hélice, uma pequena região de 3_{10} -hélice, cinco cadeias de folhas β e sete inversões. Curiosamente, foi encontrada uma estrutura equivalente para a imunoglobulina G, apesar da inexistência de

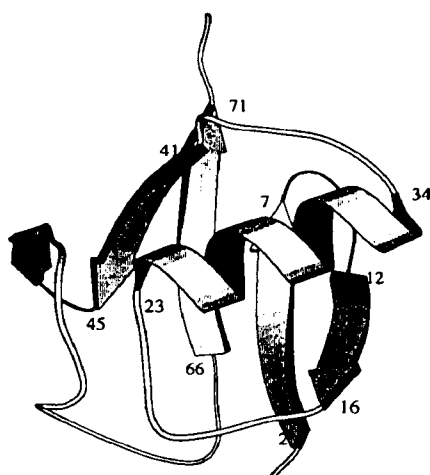


Fig.II.1. Diagrama esquemático da ubiquitina. As setas representam as cadeias de folha β e a serpentina a região de α -hélice. Retirado de Kraulis (1991).

homologia a nível das respectivas sequências primárias (Gronenborn & Clore, 1991; Kraulis, 1991). Cerca de 87% da molécula está envolvida em ligações de hidrogénio na estrutura secundária, o que explicará a sua extraordinária resistência à desnaturação quer pelo calor, sendo possível reverter a desnaturação até temperaturas de 85°C (Low *et al.*, 1979), quer por agentes químicos como a ureia, guanidina-HCl e álcool (Jenson *et al.*, 1980; Wilkinson & Mayer, 1986).

Na molécula de ubiquitina foram identificados dois locais funcionais. O primeiro, na região C-terminal, é o local de ligação a outras proteínas, num processo denominado conjugação ou ubiquitinação. Este processo é mediado pelo grupo carboxil do resíduo terminal Gly76, que se une covalentemente aos grupos ϵ -amino de resíduos lisina de diversas

proteínas. Esta reacção é extremamente dependente da estrutura da região C-terminal, já que quer a mutação da Gly76 a Ala, quer o encurtamento da cadeia através da deleção da Leu73 inibem a actividade da ubiquitina (Hodgins *et al.*, 1992; Ecker *et al.*, 1987). O segundo local funcional da ubiquitina está associado à Lys48: este resíduo serve como local aceitador para ubiquitinação de maneira equivalente a resíduos lisina em outras proteínas (Chau *et al.*, 1989). A mutação desta lisina em arginina inibe a actividade da ubiquitina, provocando lesões graves em plantas (Bachmair *et al.*, 1990). Outro resíduo lisina, Lys63, poderá eventualmente funcionar também como local para autoubiquitinação, permitindo assim a existência de conjugados proteicos com multiubiquitina ramificada (Jentsch, 1992). Mutações em outros resíduos afectam igualmente a actividade da ubiquitina, embora de uma forma menos acentuada: a mutação His68 em lisina inibe 70% da actividade proteolítica, enquanto a iodinação da Tyr59 ou a sua mutação em fenilalanina pode chegar a afectá-la em 30% (Ecker *et al.*, 1987; Pickart *et al.*, 1992). Contudo, estas mutações parecem actuar a nível de alterações na estrutura da proteína, não constituindo locais activos da mesma. A estrutura globular e compacta da molécula de ubiquitina não tem permitido determinar até ao momento outros possíveis locais activos, nomeadamente os de reconhecimento da ubiquitina pelas enzimas do sistema de conjugação.

2.2. Conjugação da ubiquitina.

A conjugação da ubiquitina com substratos proteicos é um processo constituído por várias etapas (Fig.II.2). A ubiquitina é codificada por uma família multigénica e traduzida como uma proteína precursora, que é convertida em monómeros ubiquitina por acção da enzima hidrolase C-terminal (a). O primeiro passo do ciclo de conjugação consiste na activação do resíduo C-terminal da ubiquitina, num processo catalisado pela enzima de activação da ubiquitina, E1 ou UBA (b). Este passo requer ATP, e consiste na formação de um intermediário adenilato de ubiquitina seguido da transferência da ubiquitina para um grupo tiol de uma cisteína interna da enzima E1 (Haas & Rose, 1982).

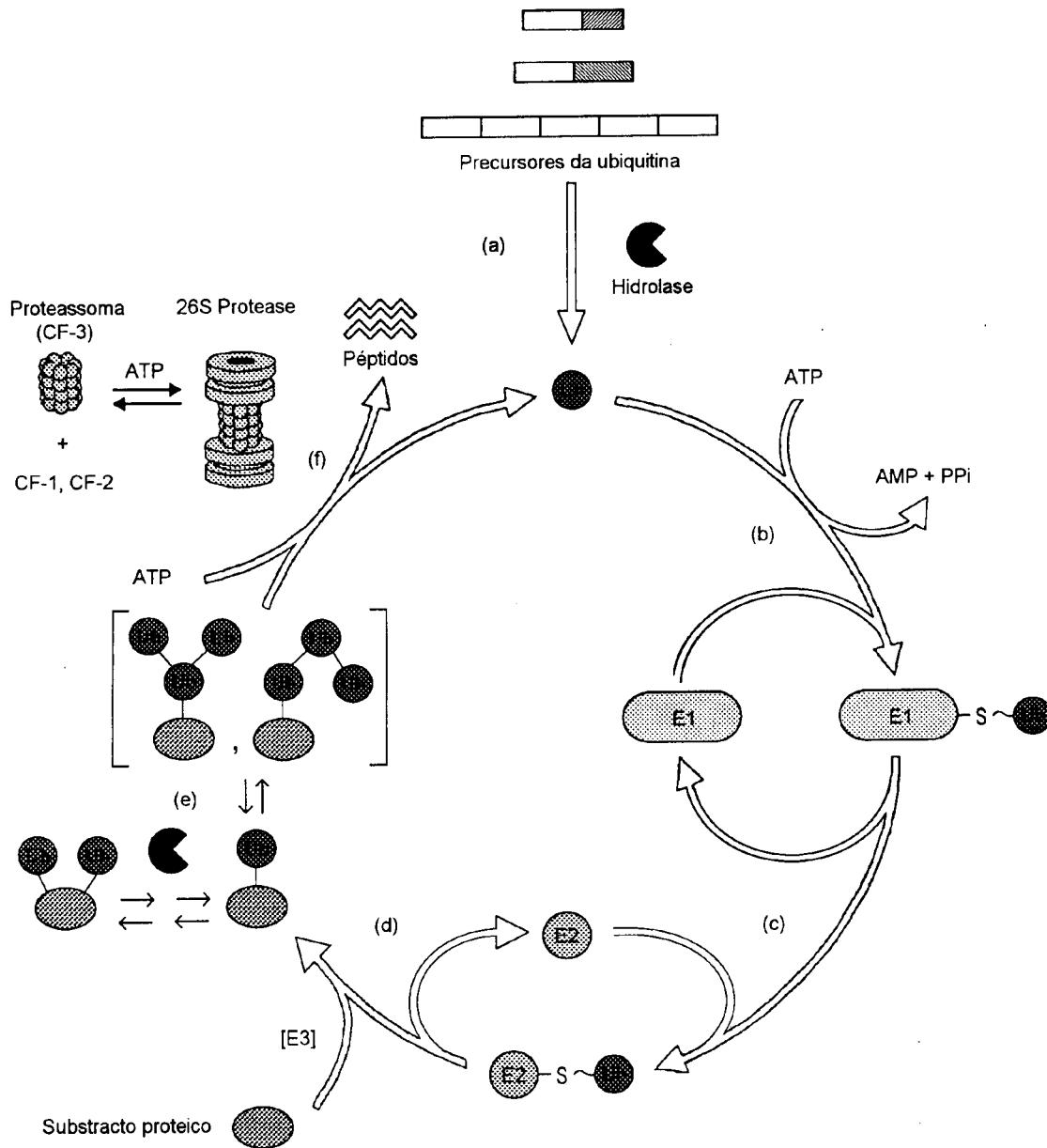


Fig.II.2. Ciclo da conjugação da ubiquitina. Os precursoros da ubiquitina (poliubiquitinas, proteínas de fusão) são hidrolisados pelas hidrolases da ubiquitina (a) originando ubiquitina livre. Esta é activada pela enzima de activação da ubiquitina, E1, num processo dependente do ATP (b), e transferida para as enzimas de conjugação da ubiquitina, E2 (c). As E2 catalisam a ubiquitinação dos substratos, na presença ou ausência das ligases da ubiquitina, E3 (d). Os substratos são mono ou multiubiquitinados, num processo que pode ser revertido pelas hidrolases (e) mas que habitualmente segue a via degradativa, catalisada pelo proteassoma (f) e regenerando a ubiquitina livre. Adaptado de Hochstrasser (1992).

A ubiquitina activada é seguidamente transferida para um resíduo cisteína específico de uma de várias enzimas de conjugação da ubiquitina, E2 ou UBC (c) (Pickart & Rose, 1985a). Estas enzimas são dadores de ubiquitina para a formação de uma ligação isopeptídica entre o resíduo glicina C-terminal da ubiquitina e grupos ϵ -amino de resíduos lisina de proteínas (d). Algumas destas ligações requerem enzimas acessórias, as ligases da ubiquitina (E3) (Reiss *et al.*, 1989). Os substratos proteicos podem ser monoubiquitinados, de que são exemplos típicos as histonas (Goldknopf & Busch, 1977), ou multiubiquitinados. A multiubiquitinação pode ser conseguida de três maneiras: por monoubiquitinação de vários resíduos lisina dos substratos, por autoubiquitinação da ubiquitina no seu resíduo

Lys48 formando longas cadeias lineares, ou por autoubiquitinação da ubiquitina nos seus resíduos Lys48 e Lys63, formando cadeias ramificadas (Chau *et al.*, 1989; Jentsch, 1992). Os substratos ubiquitinados podem ser desubiquitinados pela enzima hidrolase da ubiquitina (e), regenerando substrato e ubiquitina livre (Hershko & Ciechanover, 1992). É o que se passa com os substratos monoubiquitinados, que se crê não serem degradados, mas também pode ocorrer com substratos multiubiquitinados. Contudo, estes últimos normalmente são degradados por um complexo multiproteico em forma de framboesa, o proteassoma, com a regeneração da ubiquitina, que reiniciará o ciclo (f).

2.3. Degradação proteica mediada pelo sistema ubiquitina.

Embora a ubiquitinação de proteínas possa não conduzir à sua degradação e poder aparentemente constituir um mecanismo próprio de regulação celular, como com as histonas H2A e H2B (# 3.3), a maioria das proteínas ubiquitinadas é efectivamente degradada. Todavia, o sistema ubiquitina tanto actua como um meio de eliminar proteínas mal formadas ou anormais, como constitui um mecanismo fino de regulação de diversos sistemas celulares através do controlo da concentração de proteínas reguladoras, de que são exemplo as ciclinas e o repressor transcrricional MATα2 (# 3.2).

A compreensão deste problema motivou a pesquisa de determinantes estruturais que fossem reconhecidos pelo sistema ubiquitina. Os primeiros dados foram conseguidos por Bachmair *et al.* (1986), que mostraram que o tempo de semivida de uma proteína depende fortemente do seu resíduo N-terminal. Estes autores realizaram fusões de ubiquitina com a enzima β-galactosidase do tipo Ub-X-βgal, representando X um aminoácido qualquer. Independentemente do aminoácido X (com a única excepção da Pro), esta proteína de fusão é eficientemente desubiquitinada, à semelhança das proteínas de fusão Ub52 e Ub79 (# 4.1.2). Todavia, dependendo da natureza de X, a proteína X-βgal resultante é metabolicamente instável ou estável. Ao código resultante chamou-se Regra do N-terminal, ou "N-end rule" (Fig.II.3) (Varshavsky, 1992). Esta regra foi demonstrada não só em células eucariotas, como células de levedura ou reticulócitos de coelho (Bachmair & Varshavsky, 1989; Gonda *et al.*, 1989), mas também em *Escherichia coli* (Tobias *et al.*, 1991), embora cada organismo possua grupos de aminoácidos estabilizantes ou destabilizantes distintos (Fig.II.3).

Aminoácidos	F	L	W	Y	R	K	H	I	N	Q	D	E	C	A	S	T	G	V	P	M
<i>E.coli</i>	●	●	●	●	■	■	○	○	○	○	○	○	○	○	○	○	○	○	?	○
<i>S.cerevisiae</i>	●	●	●	●	●	●	●	●	◆	◆	■	■	○	○	○	○	○	○	○	○
Reticulócito	●	●	●	●	●	●	●	●	◆	◆	■	■	■	●	●	●	○	○	○	○

Fig.II.3. Comparação da Regra do N-terminal entre bactérias e eucariotas. Os círculos abertos representam resíduos estabilizantes; os círculos, quadrados e losangos fechados representam, respectivamente, resíduos estabilizantes primários, secundários e terciários nas Regras do N-terminal de *E.coli*, *S.cerevisiae* e reticulócitos de coelho. O ponto de interrogação indica um estatuto desconhecido para o resíduo. Retirado de Varshavsky (1992).

A actividade dos resíduos N-terminais pode ser explicada em eucariotas por um modelo no qual são directamente reconhecidos pelas ligases da ubiquitina (# 2.4.4). Estas enzimas reconhecem aparentemente resíduos básicos ou com cadeias hidrofóbicas ramificadas (Reiss *et al.*, 1988; Gonda *et al.*, 1989), que por isso são denominados resíduos destabilizantes primários. Os resíduos ácidos e amidados, igualmente destabilizantes, necessitam primeiro de ser modificados e são por isso denominados, respectivamente, secundários e terciários (Fig.11.3). Os resíduos amidados, Gln e Asn, são convertidos aos respectivos grupos ácidos, Glu e Asp, através da acção de uma desamidase específica (Gonda *et al.*, 1989). A estes, por sua vez, é adicionado um resíduo Arg por uma arginil-tRNA-transferase (R-transferase), permitindo o seu reconhecimento pelas ligases da ubiquitina (Ferber & Ciechanover, 1987; Ciechanover *et al.*, 1988). Existe todavia uma aparente anormalidade neste sistema, uma vez que o tempo de semivida da Gln-βgal (10 min) em levedura é cerca de um terço da Glu-βgal (30 min), embora a primeira tenha de ser convertida na segunda (Bachmair *et al.*, 1986; Finley & Chau, 1991). Em *E.coli*, existem também resíduos secundários (Arg e Lys; Fig.11.3) que necessitam de ser conjugados a resíduos Leu ou Phe pela enzima Leu,Phe-tRNA-transferase (L/F-transferase), mas não existem resíduos terciários (Tobias *et al.*, 1991).

A existência de um resíduo N-terminal destabilizante não é contudo suficiente para determinar a degradação de uma proteína. É necessário um segundo determinante, que consiste na presença de um resíduo interno específico Lys (Bachmair & Varshavsky, 1989). Este resíduo é o local de ligação de uma cadeia multiubiquitina, cuja formação é requerida para a degradação da maioria dos substratos sujeitos à regra do N-terminal (Chau *et al.*, 1989). O(s) resíduo(s) Lys seleccionado(s) para a ubiquitinação têm aparentemente de se encontrar espacialmente próximo(s) do resíduo N-terminal destabilizante, e quer a região contendo a lisina quer a região contendo o resíduo N-terminal necessitam de ser móveis (Bachmair & Varshavsky, 1989; Johnson *et al.*, 1990). Estes dois determinantes não necessitam necessariamente de se localizar na mesma subunidade em caso de proteínas multiméricas: neste caso, pode ocorrer um reconhecimento em *trans* (Johnson *et al.*, 1990). Todavia, apenas a subunidade contendo a cadeia multiubiquitina é efectivamente degradada, o que permite postular a existência de mecanismos de selecção de subunidades para degradação (Varshavsky, 1992).

A maioria das proteínas não compartimentalizadas não possuem resíduos N-terminais destabilizantes, o que se deve às propriedades da enzima Met-aminopeptidase, que remove o resíduo N-terminal Met apenas se o resíduo seguinte não for destabilizante (Arfin & Bradshaw, 1988; Hirel *et al.*, 1989; Sherman *et al.*, 1985). Assim, a população das proteínas celulares que possam estar sujeitas ao reconhecimento e degradação pela regra do N-terminal parece ser limitada, o que é demonstrado pelo facto de que um mutante de levedura sem o gene codificante para a ligase da ubiquitina apenas apresenta um fenótipo ligeiro (Bartel *et al.*, 1990). É possível que esta regra reconheça predominantemente proteínas compartimentalizadas cujo sinal peptídico fosse removido, mas que estejam mal localizadas no citosol (Hershko & Ciechanover, 1992), ou que tenha um papel importante em certo tipo de células, como p. ex. os reticulócitos (# 2.4.4). Por outro lado, existem várias evidências de que o sistema ubiquitina é reconhecido por diferentes sinais. Várias proteínas são acetiladas no resíduo N-terminal, e pelo menos uma proteína não relacionada com a ligase da ubiquitina, o factor H, é necessária para a degradação de uma destas proteínas acetiladas, a histona H2A (Gonen *et al.*, 1991). Por análise computacional, Rogers e colaboradores (1986) encontraram sequências ricas em resíduos de Pro, Glu, Ser e Thr (sequências PEST), tipicamente rodeadas por resíduos carregados, que existem em 95%

das proteínas de vida curta conhecidas. Contudo, o seu mecanismo de actuação é ainda desconhecido (Hershko & Ciechanover, 1992).

2.4. Enzimas do sistema ubiquitina.

2.4.1. Hidrolases C-terminais da ubiquitina.

As hidrolases C-terminais da ubiquitina são enzimas essenciais do sistema ubiquitina que catalisam a remoção da ubiquitina de diversos substratos proteicos (Fig.11.2). As suas funções não estão satisfatoriamente estabelecidas, mas crê-se que actuem a pelo menos quatro níveis no sistema ubiquitina (Hershko & Ciechanover, 1992). Na via proteolítica, as hidrolases são necessárias para a remoção da ubiquitina das ligações isopeptídicas dos resíduos Lys dos substratos proteicos e à despolimerização das cadeias multiubiquitina (Eytan *et al.*, 1993). É possível que as hidrolases possuam também um papel corrector da proteólise, nomeadamente por remoção de proteínas incorrectamente ubiquitinadas ou por reajuste de estruturas multiubiquitina anormais (Rose, 1988). As hidrolases são também necessárias ao processamento dos precursores da biosíntese da ubiquitina, como as poliubiquitinas e proteínas de fusão, e à remoção dos aminoácidos C-terminais extra que existem nas poliubiquitinas (§ 4.1). Duas outras possíveis funções das hidrolases consistirão na reciclagem da ubiquitina que reagiu com nucleófilos intracelulares, como a glutatona ou poliaminas (Rose, 1988), ou na desubiquitinação de proteínas reversivelmente ubiquitinadas que não seguem a via degradativa, como as histonas H2A e H2B (Mueller *et al.*, 1985).

As hidrolases da ubiquitina constituem uma grande família, constituída por várias classes de proteínas. A mais conhecida é composta por proteínas de cerca de 30 kDa, que aparentemente actuam sobre conjugados de baixa massa molecular (Pickart & Rose, 1985b; Mayer & Wilkinson, 1989). Assim, são capazes de hidrolisar ligações da ubiquitina a grupos ϵ -amino de pequenas proteínas, como histonas, ou processar proteínas de fusão como a Ub52 (Hershko & Ciechanover, 1992). Estas proteínas também hidrolisam ésteres tiol entre ubiquitina e DTT, pelo que poderão estar envolvidas na reciclagem da ubiquitina ligada a nucleófilos intracelulares (Rose & Warms, 1983; Pickart & Rose, 1985b). As hidrolases de 30 kDa têm sido clonadas de vários organismos, tendo um desses genes previamente identificado como a proteína PGP 9.5, uma proteína que constitui 1-5% do total de proteínas das células neuronais (Wilkinson *et al.*, 1989; Miller *et al.*, 1989). O significado desta abundância não é conhecido, mas é de notar que a ubiquitina tem sido associada a diversas doenças neuronais (para revisão, ver Mayer *et al.*, 1991).

A hidrolase de 30 kDa é incapaz de hidrolisar fusões da ubiquitina a proteínas de elevada massa molecular, como a β -galactosidase. Este tipo de fusões são processadas mais eficientemente por uma outra classe de hidrolases, de 90 kDa (Tobias & Varshavsky, 1991). Esta enzima actua também em fusões naturais, como nas proteínas de fusão Ub52 e Ub79 ou nas poliubiquitinas. Contudo, estas só são processadas quando o gene poliubiquitina é coexpresso com a enzima em *E.coli*, mas não em extractos ao qual se adiciona posteriormente a enzima, o que sugere que o processamento do precursor poliubiquitina pode ser cotraducional (Tobias & Varshavsky, 1991). Uma outra hidrolase, que actua preferencialmente multiubiquitina, foi recentemente caracterizada por Hadari *et al.* (1992). Esta enzima, de 100 kDa, converte conjugados de elevada massa molecular em formas de baixa massa molecular, que se tornam resistentes à acção da enzima.

Estes dados sugerem que as diversas classes de hidrolases possam ter funções complementares, actuando a diferentes níveis do sistema ubiquitina. Todavia, é provável que algumas funções sejam suportadas por mais de uma classe. A deleção dos quatro genes codificantes para hidrolases da ubiquitina conhecidos em levedura conduz a um mutante ainda viável, o que sugere a presença de outras hidrolases com funções equivalentes (Tobias & Varshavsky, 1991; Baker *et al.*, 1992).

2.4.2. E1/UBA - Enzimas de activação da ubiquitina.

As enzimas de activação da ubiquitina são enzimas que formam um éster tiol E1-ubiquitina num processo que envolve um intermediário ubiquitina-adenilato (Fig.11.2). A enzima foi purificada em vários organismos, e aparenta ser um homodímero com uma massa molecular de cerca de 210 kDa (Ciechanover *et al.* 1982). Os genes codificantes para as subunidades da enzima têm sido identificados em plantas, animais e levedura (Hatfield & Vierstra, 1989; 1992; Imai *et al.*, 1992; McGrath *et al.*, 1991). A sua análise revela que esta enzima, à semelhança de outras do sistema ubiquitina, é muito conservada evolutivamente. As sequências apresentam dois motivos Gly-X-Gly-X-X-Gly, característicos de domínios de ligação a nucleótidos, e 5 resíduos Cys conservados, um dos quais (Cys626 em trigo; Hatfield & Vierstra, 1992) é o resíduo aceitador da ubiquitina activada.

As E1 são enzimas abundantes no citosol e no núcleo, sendo também encontradas associadas aos diversos componentes do citoesqueleto (Cook & Chock, 1991a; Trausch *et al.*, 1993). A localização nuclear parece ser dependente do ciclo celular, já que as E1 desaparecem durante a fase S para reaparecerem na fase G2, associando-se durante a mitose ao fuso mitótico mas não aos cromossomas (Grenfell *et al.*, 1994). Curiosamente, vários mutantes termossensíveis das E1 são incapazes de prosseguir o ciclo celular para além das fases S/G2 ou S (Mita *et al.*, 1980; Finley *et al.*, 1984; Cook & Chock, 1991b). Apesar disso, o papel destas enzimas no núcleo não é conhecido, podendo estar envolvidas na ubiquitinação de proteínas nucleares específicas, como as histonas, ou na degradação de outras proteínas nucleares (Hershko & Ciechanover, 1992).

A enzima E1 existe sob diversas isoformas, tendo sido encontrados vários genes homólogos a esta enzima em diversos organismos (Cook & Chock, 1992; Hatfield & Vierstra, 1992; Mitchell *et al.*, 1991; 1992; Leyser *et al.*, 1993). Contudo, é possível que as proteínas codificadas por estes genes possuam papeis distintos das E1 canónicas: o gene UBA1 de levedura é essencial à viabilidade da célula, apesar da aparente presença de um segundo gene homólogo (McGrath *et al.*, 1991; Jentsch, 1992). As possíveis funções destes genes podem estar relacionadas com a resistência à auxina em plantas (Leyser *et al.*, 1993) e espermatogénese em mamíferos (Mitchell *et al.*, 1991; 1992).

2.4.3. E2/UBC - Enzimas de conjugação da ubiquitina.

As enzimas de conjugação da ubiquitina constituem uma família de enzimas capaz de mediar a transferência da ubiquitina das enzimas de activação da ubiquitina, E1, aos substratos proteicos na presença ou ausência das ligases da ubiquitina, E3 (Fig.11.2). Várias enzimas e os seus correspondentes genes têm sido identificados nos mais diversos organismos, incluindo o vírus da peste suína africana, revelando-se uma família muito

heterogênea na sua estrutura e função (Jentsch, 1992; Zhen *et al.*, 1993; Bartling *et al.*, 1993; Rodriguez *et al.*, 1992; Hingamp *et al.*, 1992). Algumas enzimas são extremamente raras, enquanto outras são insolúveis, com massas moleculares que variam de 14 kDa a cerca de 500 kDa (Jentsch, 1992). Todas as UBC possuem um domínio muito conservado com cerca de 150 resíduos de aminoácidos (domínio UBC), que inclui uma Cys central necessária à formação do tioéster ubiquitina-enzima. A montante da cisteína encontra-se uma região conservada rica em prolinas, e na região N-terminal do domínio UBC existe um motivo rico em resíduos básicos, também muito conservado. A substituição de uma das prolinas invariantes produz uma enzima parcialmente termoinstável (Ellison *et al.*, 1991), enquanto a deleção parcial da zona básica reduz a capacidade da enzima formar ligações tioéster com a ubiquitina, sugerindo que esta região se encontra envolvida no reconhecimento da, ou na ligação à, E1 (Sullivan & Vierstra, 1991).

De acordo com a sua estrutura, as enzimas de conjugação da ubiquitina podem ser divididas em três classes (Jentsch *et al.*, 1990). A classe I é constituída por enzimas que consistem quase exclusivamente no domínio UBC. Estas enzimas transferem deficientemente a ubiquitina, e necessitam aparentemente da ligase da ubiquitina, E3, para o seu funcionamento (Jentsch *et al.*, 1991; Parag *et al.*, 1993). As enzimas da classe II possuem extensões C-terminais não relacionadas entre si que contribuem para a especificidade destas enzimas. A remoção destas extensões reduz a capacidade da enzima transferir a ubiquitina para substratos, embora continue a interagir com a E1 (Sung *et al.*, 1988). Em pelo menos uma das enzimas, a região C-terminal contribui para a sua localização celular: a UBC6 de levedura possui um sinal hidrofóbico em âncora que posiciona a proteína no retículo endoplasmático (Sommer & Jentsch, 1993). A classe III contém enzimas que possuem extensões N-terminais mas não C-terminais, cuja função é desconhecida (Jentsch, 1992). Existem ainda duas UBC de elevada massa molecular que não são englobáveis nesta classificação: uma proteína solúvel de 230 kDa de reticulócitos de coelho (Klemperer *et al.*, 1989), e uma proteína membranaar de 500 kDa de rato que possui, para além de uma pequena extensão C-terminal, uma longa região N-terminal com várias possíveis zonas transmembranares (Jentsch, 1992).

2.4.4. E3 - Ligases da ubiquitina.

Algumas enzimas de conjugação da ubiquitina necessitam de factores acessórios, as ligases da ubiquitina, E3, para o reconhecimento dos substratos proteicos (Fig. I1.2). O reconhecimento parece dar-se fundamentalmente no resíduo N-terminal dos substratos, pelo que estas enzimas parecem estar envolvidas sobretudo na degradação proteica pela regra do N-terminal (# 2.3). As E3 têm sido encontradas em levedura e em reticulócitos de coelho (Bartel *et al.*, 1990; Reiss *et al.*, 1989; Reiss & Hershko, 1990), apresentando ser homodímeros com subunidades de cerca de 200 kDa.

Uma destas enzimas, a E3 α de reticulócito de coelho, tem sido profundamente estudada. Esta enzima possui dois locais distintos de interacção com os resíduos N-terminais dos substratos (Reiss *et al.*, 1988). Um dos locais reconhece resíduos básicos (Arg, His, Lys), designados como Tipo I, enquanto o segundo reconhece resíduos hidrofóbicos ramificados (Leu, Phe, Trp, Tyr), designados como Tipo II. Uma terceira classe de substratos, Tipo III, que não possui resíduos N-terminais de nenhum dos tipos anteriores, é também reconhecida pela enzima, mas o local responsável pelo reconhecimento permanece desconhecido (Reiss *et al.*, 1988; Reiss & Hershko, 1990). A enzima possui aparentemente

um local de ligação ao "corpo" do substrato, e que poderá ser o responsável pelo reconhecimento dos substratos do tipo III. Finalmente, a E3 α possui ainda um local de ligação à E2 e outro à ubiquitina, que estabiliza os conjugados ubiquitina-substrato e deverá ser necessário à adição progressiva de unidades múltiplas de ubiquitina (Reiss *et al.*, 1989). Os conjugados multiubiquitinados não são libertados da E3 α , parecendo ser directamente canalizados para o complexo proteolítico, o proteassoma.

Apesar da sua aparente importância, a E3 α não parece ser essencial à célula, já que a remoção do gene que codifica para o seu homólogo em levedura não produz alterações fenotípicas significativas (Bartel *et al.*, 1990). Aparentemente, a regra do N-terminal parece funcionar principalmente para proteínas compartimentalizadas que estejam mal localizadas (# 2.3). Em reticulócito de coelho, foi caracterizada uma segunda ligase, E3 β , que, apesar de possuir diversas propriedades comuns à E3 α , reconhece principalmente como substrato os resíduos N-terminais Thr, Ser e Ala (Gonda *et al.*, 1989; Heller & Hershko, 1990). Esta ligase poderá explicar a divergência observada para a regra do N-terminal entre a levedura e os reticulócitos de coelho (Fig.11.3). Os três resíduos Thr, Ser e Ala aparecem frequentemente na N-terminal de proteínas de vida longa não compartimentalizadas (Gonda *et al.*, 1989), pelo que é possível que a regra do N-terminal do reticulócito seja diferente da do seu precursor, já que a destruição selectiva de proteínas de vida longa é particularmente extensa durante a conversão do reticulócito em eritrócito (Varshavsky, 1992).

Outras proteínas que facilitem as reacções de conjugação da ubiquitina, mesmo sem interagir fisicamente com as E2, poderão também ser classificadas como E3 (Jentsch, 1992). É o caso da oncoproteína E6 codificada pelo papillomavírus, já que a sua ligação à proteína supressora do tumor p53 promove a ubiquitinação e degradação desta (Huibregtse *et al.*, 1991; Scheffner *et al.*, 1990). Várias outras proteínas que tenham um papel na reparação proteica, como as proteínas de "stress" hsp60 e hsp70 ou outras chaperones como a CCT, poderão também ser consideradas como E3 (Jentsch, 1992).

2.4.5. O proteassoma.

O proteassoma é uma partícula intracelular de elevada massa molecular (cerca de 650 kDa) que possui propriedades proteolíticas. Foi isolado em vários organismos e encontra-se principalmente no citoplasma e no núcleo das células (Rivett, 1993). É constituído por cerca de duas dezenas de subunidades diferentes, entre 20-35 kDa, que se organizam numa estrutura cilíndrica de quatro anéis com 6-7 subunidades em cada anel (Heinemeyer *et al.*, 1991; Rivett & Sweeney, 1991; Djaballah *et al.*, 1993). Na arqueobactéria *Thermoplasma acidophilum*, cujo proteassoma só possui duas subunidades diferentes - α e β - o primeiro e o quarto anel são constituídos por subunidades α e os dois anéis centrais por subunidades β (Grziwa *et al.*, 1991). Vários genes codificantes para diversas subunidades têm sido isolados em levedura e animais, e a análise das suas sequências permite classificá-los na sua maior parte (mas não na globalidade) em dois grupos - A e B - baseados na semelhança com as subunidades α e β de *T.acidophilum* (Rivett, 1993).

O proteassoma possui funções proteolíticas múltiplas, de que fazem parte as hidrólises específicas das ligações peptídicas seguintes aos resíduos básicos, ácidos ou hidrofóbicos. Todavia, o proteassoma arqueobacteriano possui apenas o último tipo de actividade catalítica (Dahlmann *et al.*, 1989; Rivett, 1993). Estas funções são aparentemente independentes, com pelo menos cinco locais activos distintos no proteassoma de mamíferos

(Rivett, 1993). Todavia, pouco é conhecido acerca dos diferentes centros catalíticos, ou das subunidades envolvidas. Em levedura, mutantes defectivos em alguns genes de subunidades do proteassoma (PRE1-4) apresentam actividades reduzidas de alguns tipos de catálise (Heinemeyer *et al.*, 1991; 1993; Hilt *et al.*, 1993). Vários genes são essenciais à viabilidade celular, mas outros não apresentam fenótipos visíveis (Emori *et al.*, 1991). A actividade proteolítica do proteassoma pode também ser afectada pela presença de proteínas inibidoras ou activadoras (Rivett, 1993).

A degradação de proteínas ubiquitinadas não é contudo feita directamente pelo proteassoma. Este possui uma actividade proteolítica própria, não-lisossomal e dependente do ATP que tem sido associada a diversos processos celulares, como a geração de pequenos péptidos que serão presentes ao sistema imunitário (Goldberg & Rock, 1992). Apesar disso, o proteassoma é essencial ao sistema ubiquitina, já que o mutante de levedura *pre1*, deficiente numa subunidade do proteassoma que afecta a actividade sobre resíduos hidrofóbicos, não é capaz de degradar proteínas habitualmente sujeitas à proteólise dependente da ubiquitina em células normais (Seufert & Jentsch, 1992). A degradação de proteínas ubiquitinadas é feita por um complexo proteolítico de 1500 kDa (proteínase 26S), que para além do padrão proteico característico do proteassoma, possui ainda vários polipéptidos de 35 a 110 kDa (Hough *et al.*, 1987). Aparentemente, a proteína 26S consiste numa associação dependente do ATP do proteassoma com vários outros factores (Eytan *et al.*, 1989; Driscoll & Goldberg, 1990), dois dos quais - CF1 e CF2 - poderão constituir um activador e um inibidor do proteassoma (Goldberg & Rock, 1992; Driscoll *et al.*, 1992). Um outro componente da proteína 26S é uma protease dependente do ATP, denominada multipaina, cuja relação com os factores CF1 ou CF2 é desconhecida. Esta protease é capaz só por si de degradar conjugados da ubiquitina, mas a sua inclusão no complexo aumenta a sua actividade e permite uma proteólise mais extensa dos substratos (Hershko & Ciechanover, 1992). Todavia, a própria presença do proteassoma no complexo 26S não é consensual: alguns autores identificam um factor de massa semelhante à do proteassoma mas totalmente distinto deste, que seria o verdadeiro constituinte do complexo. Estes autores argumentam que a actividade característica do proteassoma detectada na fracção contendo o complexo é causada por agregados do proteassoma formados durante o processo de purificação, já que, usando processos mais suaves, conseguem obter complexos 26S onde não detectam a presença de actividade característica do proteassoma nem a presença de subunidades deste por ensaios imunológicos (Sellig *et al.*, 1991; Kuehn *et al.*, 1992).

3. Funções da ubiquitinação.

Nos últimos anos tem-se encontrado um envolvimento do sistema ubiquitina nas mais diversas funções celulares, desde processos tão distintos como o ciclo celular e a reparação do DNA à translocação de proteínas. Na generalidade dos casos, tem-se verificado que o sistema actua através de um ajuste fino da sua acção degradativa a proteínas reguladoras e/ou envolvidas nas diversas funções estudadas, como a degradação das ciclinas durante o ciclo celular ou a degradação do repressor transcrricional MAT α 2 (Glotzer *et al.*, 1991; Chen *et al.*, 1993). Porém, têm sido encontradas várias proteínas ubiquitinadas que aparentemente não seguem a via degradativa. Exemplos típicos são as histonas H2A e H2B, e certas proteínas da superfície celular (Goldknopf & Busch, 1977; Siegelman *et al.*, 1986). Normalmente, estas proteínas são monoubiquitinadas, e a sua função não é conhecida; não é

contudo de excluir que constituam precursores com vida longa de produtos finais multiubiquitinados que sejam rapidamente degradados (Finley & Chau, 1991).

Para além das características próprias dos substratos, a especificidade do sistema ubiquitina para as diversas funções estudadas parece ser conferido sobretudo pelas diversas enzimas de conjugação existentes, as E2 ou UBC. Das cerca de dez enzimas descobertas em levedura, apenas um pequeno grupo actua em cada uma das funções, parecendo observar-se nesses casos uma acção coordenada entre elas. Todavia, é raro encontrar-se uma dependência total de uma função a uma determinada E2, e frequentemente o papel desempenhado por uma E2 pode ser também desempenhado com maior ou menor eficiência por várias outras (Jentsch, 1992).

3.1. Reparação do DNA.

A importância do sistema ubiquitina na reparação e mutagénese do DNA celular veio com a descoberta em levedura de que dois dos genes envolvidos neste processo, os genes de reparação do DNA *RAD6* e *RAD23*, codificam respectivamente para uma enzima de conjugação da ubiquitina, a UBC2, e uma proteína contendo uma região semelhante à ubiquitina (Jentsch *et al.*, 1987; Watkins *et al.*, 1993). Mutações no gene *RAD6* conferem hipersensibilidade à luz ultravioleta, raios X e agentes mutagénicos, bem como uma incapacidade de converter lesões do DNA em mutações (Prakash *et al.*, 1990). Para além disso, os mutantes apresentam vários outros fenótipos, como um aumentado da frequência de retrotransposições, crescimento lento, esporulação defeituosa e paragem do ciclo celular (Picologlou *et al.*, 1990; Ellison *et al.*, 1991). Esta multiplicidade funcional da *RAD6*, que reflecte provavelmente a existência de substratos fisiológicos múltiplos, é dependente da sua actividade como enzima de conjugação: a substituição do seu único resíduo Cys, no qual ocorre a ligação tioéster com a ubiquitina, por Ala ou Val origina mutantes fenotipicamente equivalentes a mutantes que não possuam o gene (Sung *et al.*, 1990). *In vitro*, a *RAD6* pode monoubiquitinar ou multiubiquitinar (neste caso, envolvendo cadeias ligadas a Lys63; Jentsch, 1992) histonas na ausência das ligases da ubiquitina E3 (Jentsch *et al.*, 1987; Sung *et al.*, 1988; Haas *et al.*, 1990). Todavia, a *RAD6* pode formar complexos com as ligases da ubiquitina E3 para outras funções, nomeadamente para a degradação de proteínas via regra do N-terminal (Dohmen *et al.*, 1991; Sung *et al.*, 1991).

Os mecanismos de actuação da *RAD6* na reparação do DNA ou noutros fenótipos que produz, bem como a maior parte dos substratos onde actua, permanece contudo desconhecida. A *RAD6* é uma enzima de conjugação da classe II, contendo um núcleo catalítico comum às UBC e uma região C-terminal rica em resíduos ácidos (Reynolds *et al.*, 1985). A ubiquitinação de histonas requer a presença desta cauda ácida, mas mutantes em que esta foi deletada mantêm a eficiência na reparação do DNA e na resistência aos UV (Morrison *et al.*, 1988; Sung *et al.*, 1988). É de notar ainda que os genes homólogos da *RAD6* que foram encontrados noutros organismos pertencem à classe I, sem extensões C-terminais (Jentsch, 1992). Todavia, apesar de os diversos núcleos catalíticos das diversas UBC serem muito homólogos, as funções da *RAD6* não são substituíveis por outro núcleo catalítico de outra E2, como o da UBC3, mesmo que a este se lhe adicione a região C-terminal da *RAD6* (Silver *et al.*, 1992). A presença da região C-terminal é necessária para restaurar a esporulação, pelo que pode estar envolvida no reconhecimento dos substratos envolvidos neste processo, que serão distintos dos envolvidos na reparação do DNA (Finley & Chau, 1991).

O gene *RAD23* de levedura codifica para uma proteína cuja região N-terminal é semelhante à ubiquitina, sendo por isso um gene relacionado com a ubiquitina da classe B (# 4.2). Mutações neste gene causam uma sensibilidade moderada aos raios UV e uma eficiência reduzida na reparação do DNA (Prakash *et al.*, 1990). A função da proteína não é conhecida, mas em humanos, onde também foram encontrados dois genes, forma um complexo com uma segunda proteína envolvida na reparação do DNA - XP-C -, uma proteína equivalente à *RAD4* de levedura (Masutani *et al.*, 1994). A região semelhante à ubiquitina é essencial à função, embora possa ser substituída pela ubiquitina canónica sem perda de eficiência. A *RAD23* parece ser degradada pelo sistema ubiquitina, mas não existem evidências de que a região ubiquitina esteja directamente envolvida, já que a mutação das lisinas desta região em argininas não afectam a degradação da proteína. Por outro lado, o sistema ubiquitina não parece actuar como um regulador fino dos níveis desta proteína, uma vez que a *RAD23* é uma proteína muito estável, com tempos de semivida superiores a 6 horas (Watkins *et al.*, 1993).

3.2. Controlo do ciclo celular.

A importância do sistema ubiquitina no ciclo celular foi sugerida pelos estudos realizados em algumas linhas celulares mutantes de mamíferos, onde se identificou o factor termolábil responsável pela paragem na fase G2 e/ou S à temperatura não permissiva como sendo a enzima de activação da ubiquitina, E1 (Finley *et al.*, 1984; Kulka *et al.*, 1988). A ubiquitinação parece ser então um processo necessário ao correcto desenvolvimento do ciclo celular. De todas as suas possíveis funções, a melhor caracterizada prende-se com a regulação da progressão da mitose. A mitose é iniciada com a activação do factor promotor da fase M (MPF - "Mitose Protein Factor"), uma proteína cinase cujas principais subunidades são o p34^{cdc2}, um polipéptido possuindo homologia com várias cinases conhecidas, e a ciclina, uma subunidade reguladora da cinase cuja abundância varia durante o ciclo celular. A transição da metáfase para a anáfase, que marca o fim da mitose, é induzida pela degradação da ciclina e consequente inactivação do MPF (Feilotter *et al.*, 1992; Murray *et al.*, 1989). Recentemente, foi demonstrado que a degradação da ciclina é acompanhada por uma multiubiquitinação desta proteína, sendo a cinética de degradação consistente com a degradação via sistema ubiquitina (Glutzer *et al.*, 1991; Hershko *et al.*, 1991). Dois motivos estruturais presentes nas ciclinas parecem mediar a sua degradação: uma Arg invariante numa região consensual perto da região N-terminal, denominada "caixa de destruição", e um grupo de lisinas que presumivelmente conterão o local de ubiquitinação. Estes dois motivos são suficientes para conferir a uma proteína heteróloga um ciclo degradativo dependente do ciclo celular (Glutzer *et al.*, 1991; Finley & Chau, 1991). Mutantes da ciclina cuja "caixa de destruição" esteja mutada ou ausente não são degradados, e consequentemente o MPF não é inactivado (Murray *et al.*, 1989). Estas células não terminam a mitose, mas são ainda capazes de iniciar a anáfase, permitindo a separação das cromátidas irmãs. Todavia, esta separação é atrasada quando às células é adicionado o fragmento N-terminal da ciclina, que actua como competidor da degradação das ciclinas, ou com a adição de ubiquitina metilada, que inibe a proteólise pelo sistema ubiquitina. Este atraso verifica-se mesmo na presença de mutantes contendo ciclinas não degradáveis, o que sugere a presença de outras proteínas que necessitam de ser degradadas para dissolver a ligação entre as cromátidas irmãs (Holloway *et al.*, 1993).

Estudos sobre a divisão celular em levedura permitiram evidenciar a importância do sistema ubiquitina no controlo do ciclo celular. Em levedura, as células que tenham iniciado a fase S do ciclo celular necessitam de executar as funções conferidas pelos genes *CDC34* e *CDC4* antes que a replicação do DNA cromossomal possa ocorrer (Goebel *et al.*, 1988). O gene *CDC4* codifica para uma proteína contendo motivos repetidos semelhantes aos encontrados nas transducinas ou Gs de mamíferos (Yochem & Byers, 1987), enquanto o gene *CDC34* se verificou codificar para uma enzima de conjugação de ubiquitina da classe II, UBC3 (Goebel *et al.*, 1988). A região C-terminal da UBC3 contém, sensivelmente a meio, uma zona muito rica em resíduos ácidos que, contudo, não é essencial para a função da proteína, suportada apenas pela região compreendida entre essa zona e o núcleo catalítico (Kolman *et al.*, 1992). Já o núcleo catalítico da UBC3 não parece conter nenhuma particularidade essencial à função, contrariamente ao observado com a RAD6 (# 3.1): uma enzima quimérica que contenha o núcleo catalítico da UBC2 (RAD6) e a região C-terminal da UBC3 é capaz, curiosamente, de compensar mutantes tanto da RAD6 como da CDC34 (Silver *et al.*, 1992). O gene *CDC34* é transcrito como um mRNA raro (uma cópia por célula) e, apesar de terem sido identificadas várias ciclinas para a fase G1 em levedura, os substratos fisiológicos para a UBC3 não são conhecidos (Goebel *et al.*, 1988; Richardson *et al.*, 1989). É possível que a sua função passe também por uma autoubiquitinação que controle a sua própria degradação (Banerjee *et al.*, 1993).

Um outro gene de levedura codificante para uma enzima de conjugação, *UBC9*, parece também estar envolvido no ciclo celular. A depleção da UBC9 leva a uma paragem na fase G₂/M, mas os substratos da enzima não foram identificados. Embora estes possam incluir ciclinas, estas não deverão ser os únicos substratos da enzima, já que a sua degradação só parece ser requerida numa fase posterior do ciclo celular (Jentsch, 1992).

3.3. Transcrição.

As histonas ubiquitinadas uH2A e uH2B encontram-se entre os conjugados da ubiquitina mais abundantes na célula, (Goldknopf & Busch, 1977; Mueller *et al.*, 1985; Shimogawara & Muto, 1992). Todavia, apesar dos estudos intensos realizados, a função celular destes conjugados não é clara. Algumas observações indirectas sugerem que as histonas ubiquitinadas possam estar relacionadas com a transcrição: as uH2 parecem estar preferencialmente enriquecidas em regiões transcricionalmente activas da cromatina, e os níveis de uH2B caem significativamente na presença de inibidores da RNA polimerase II (Levinger & Varshavsky, 1982; Nickel *et al.*, 1989; Davie & Murphy, 1990). Não existem contudo evidências conclusivas que suportem esta hipótese, e é certo que, caso as histonas ubiquitinadas possuam algum papel na transcrição, este não é universal para as células eucariotas. De facto, embora em alguns organismos as uH2 sejam muito abundantes, (em células de vertebrados a uH2A pode atingir 15% do total das H2A), as células de fungos ou da planta *Arabidopsis* não parecem conter, ou pouco contêm, histonas ubiquitinadas (Finley & Chau, 1991). Na levedura, não só não são detectadas histonas ubiquitinadas, como a substituição do resíduo Lys por Arg na sequência conservada nas H2A que corresponde ao local de ubiquitinação das células de mamíferos não produz nenhum fenótipo detectável (Swerdlow *et al.*, 1990). Curiosamente, as histonas de vertebrados e de *Tetrahymena* possuem, acima do local de ubiquitinação, uma região semelhante à "caixa de destruição" das ciclinas (Hunt, 1991; # 5.2). As ciclinas são degradadas após poliubiquitinação durante a mitose, mas a supressão da ubiquitinação das ciclinas durante a interfase não é completa,

podendo encontrar-se ciclinas mono ou diubiquitinadas (Finley & Chau, 1991). É possível então que as histonas monoubiquitinadas sejam um intermediário estável da degradação (que ocorre pelo menos *in vitro*; Haas *et al.*, 1990), ou que estejam também envolvidas no ciclo celular. Efectivamente, em células sincronizadas de *Physarum* os níveis de uH2 desaparecem durante a metáfase e reaparecem durante a anáfase (Mueller *et al.*, 1985).

Outro modo de actuação do sistema ubiquitina na transcrição ocorre a nível de proteínas reguladoras. Uma destas proteínas é o repressor MAT α 2, um regulador transcricional do tipo de "mating" na levedura necessário à repressão dos genes α -específicos em células α (Nasmyth, 1982; Herskowitz *et al.*, 1992). Este repressor possui um tempo de semivida de cerca de 5 min, sendo a sua degradação efectuada pelo sistema ubiquitina após ligação de uma cadeia poliubiquitina (Hochstrasser & Varshavsky, 1990; Hochstrasser *et al.*, 1991). Diversos estudos revelaram a existência de pelo menos dois sinais de degradação distintos na molécula MAT α 2, e o envolvimento de pelo menos quatro enzimas de conjugação da ubiquitina no processo degradativo (Hochstrasser & Varshavsky, 1990; Chen *et al.*, 1993). A degradação parece seguir duas vias distintas, uma envolvendo as enzimas UBC6 e UBC7, e a outra envolvendo as enzimas UBC4 e UBC5 (Chen *et al.*, 1993). A UBC6 é uma proteína membranar com um domínio catalítico citoplasmático, que se localiza no retículo endoplasmático e possivelmente no envelope nuclear (Sommer & Jentsch, 1993). Neste último caso, a UBC6 poderá controlar não só a semivida da MAT α 2, como a de outras proteínas nucleares que passem através dos poros nucleares (Jentsch, 1992).

3.4. Resposta ao "stress" celular.

A indução de proteínas de "stress" a temperaturas não permissivas em linhas celulares de mamífero com termossensibilidade para a enzima de activação da ubiquitina, E1, veio sugerir um relacionamento entre o sistema ubiquitina e o "stress" celular (Finley *et al.*, 1984; Ciechanover *et al.*, 1984). De facto, veio-se posteriormente a descobrir que a ubiquitina é ela própria uma proteína de "stress", sendo os genes poliubiquitina induzidos sob diversas condições de "stress" celular em vários organismos (Bond & Schlesinger, 1986; # 4.3.).

Em levedura, os genes codificantes para as enzimas de conjugação da ubiquitina UBC4 e UBC5 são também induzidos pelo "stress" (Seufert & Jentsch, 1990). As UBC4 e UBC5 são enzimas da classe I praticamente idênticas e aparentemente intercambiáveis, que são necessárias à degradação constitutiva das proteínas de vida curta de levedura. Os mutantes duplos, *ubc4ubc5*, apresentam um decréscimo acentuado na taxa de degradação das proteínas contendo análogos de aminoácidos e são hipersensíveis a incubações prolongadas a temperaturas elevadas, sugerindo igualmente um papel na degradação de proteínas anormais ou danificadas pelo "stress" (Seufert & Jentsch, 1990). Em condições de cultura normais, estes mutantes apresentam uma taxa de crescimento reduzida, sendo a sua função aparentemente compensada por uma enzima de conjugação codificada por um outro gene, UBC1: os mutantes triplos nestes três genes são letais (Seufert *et al.*, 1990). Todavia, a UBC1, uma enzima da classe II, possui especificidades não compartilhadas pelas UBC4/UBC5: assim, o gene UBC1 não é induzido por "stress", e os mutantes *ubc1* apresentam apenas uma ligeira quebra na taxa de crescimento, mas são manifestamente deficientes no crescimento após a germinação dos ascósporos (Seufert *et al.*, 1990).

Uma outra enzima de conjugação da ubiquitina de levedura, a UBC7, está provavelmente envolvida na degradação das proteínas anormais (Jungmann *et al.*, 1993). A transcrição do gene *UBC7* (e igualmente a dos genes *UBC4* e *UBC5*) é induzida por cádmio, um potente veneno celular e agente de "stress", mas não é induzido pelo "stress" hipertérmico. Os mutantes *ubc7* são hipersensíveis ao cádmio, embora não apresentem mais nenhum outro fenótipo detectável (Jungmann *et al.*, 1993; Vassal *et al.*, 1992).

3.5. Ubiquitinação de proteínas membranares.

A ubiquitina parece desempenhar um papel importante, que em alguns casos aparenta ser independente da via degradativa, em vários processos celulares associados à membrana, como sejam a recepção de hormonas, a translocação proteica e a inserção membranar de proteínas.

A ubiquitinação de receptores membranares tem sido extensamente descrita na literatura, parecendo existirem diversos mecanismos de actuação para cada caso. O receptor de linfócitos gp90^{MEL-14} é ubiquitinado no seu domínio extracelular, e os anticorpos antiubiquitina inibem a ligação do linfócito, sugerindo que a ubiquitinação pode ser funcionalmente relevante (Siegelman *et al.*, 1986; St.John *et al.*, 1986). Não sendo conhecido nenhum sistema de ubiquitinação extracelular, a ubiquitinação do receptor deverá ocorrer antes da sua inserção na membrana, o que sugere que os conjugados ramificados são competentemente translocados (Siegelman *et al.*, 1986). Os outros receptores ubiquitinados conhecidos são-no nos seus domínios citoplasmáticos (Yarden *et al.*, 1986; Mori *et al.*, 1992; Paolini & Kinet, 1993), mas a função destas ubiquitinações permanece desconhecida. O receptor do PDGF ("platelet-derived growth factor") é ubiquitinado na sua região C-terminal após a ligação do ligando, e mutantes deletados nesta região possuem uma actividade mitogénica amplificada. Estes dados sugerem um papel regulatório negativo no sinal mitogénico pela ubiquitinação, possivelmente através da promoção da degradação do receptor activado (Mori *et al.*, 1993). Já a ubiquitinação dos receptores da imunoglobulina E não parece promover a sua degradação. A activação destes receptores pelo ligando promove a sua rápida multiubiquitinação, mas o processo é imediatamente revertido com a remoção do ligando, sugerindo um papel "on/off" em alguma função do receptor até agora desconhecida (Paolini & Kinet, 1993).

As enzimas responsáveis pela ubiquitinação de proteínas membranares são ainda pouco conhecidas. Em levedura, foi descoberta uma enzima de conjugação da ubiquitina da classe II, UBC6, que é uma proteína integral da membrana do retículo endoplasmático e, possivelmente, do envelope nuclear (Sommer & Jentsch, 1993). Sabe-se que esta enzima actua na degradação do repressor MAT α 2 (Chen *et al.*, 1993; # 3.3), mas os mutantes *ubc6* não apresentam nenhum fenótipo visível. Todavia, a perda da função da UBC6 permite restaurar a viabilidade celular à temperatura não permissiva de mutantes termosensíveis de uma proteína requerida para a translocação de proteínas através do retículo endoplasmático, SEC1. Segundo o modelo proposto, a UBC6 promove a ubiquitinação e consequente degradação do aparato translocador distorcido, que contudo ainda é parcialmente funcional (Sommer & Jentsch, 1993). Um outro gene codificante para uma enzima de conjugação membranar foi recentemente descoberta em mamíferos. Este gene codifica para uma proteína de cerca de 500 kDa com vários domínios de α -hélice transmembranares e possíveis locais de glicosilação, aparentemente constituinte do aparelho de Golgi, que pode

estar envolvida nos processos degradativos do retículo endoplasmático (Jentsch, 1992; Bonifacino & Lippincott-Schwartz, 1991; Klausner & Sitia, 1990).

O sistema ubiquitina parece estar também envolvido na translocação para o mitocôndrio. De facto, anticorpos antiubiquitina inibem a inserção das monoamina oxidases A e B na membrana exterior do mitocôndrio, sugerindo um possível papel chaperone no transporte de proteínas (Zhuang & McCauley, 1989; Zhuang *et al.*, 1992).

3.6. Infecção viral.

A detecção, em partículas de vírus da leucose de aves, de ubiquitina encapsulada (Putterman *et al.*, 1990), e a presença de proteínas ubiquitinadas da cápsula do vírus do mosaico do tabaco (Dunigan *et al.*, 1988) vieram sugerir um possível envolvimento do sistema ubiquitina na infecção viral. Embora o significado biológico destas observações não tenha sido ainda determinado, vários outros dados sugerem que alguns vírus podem modular activamente o sistema ubiquitina do hospedeiro. O melhor exemplo é dado pela promoção da ubiquitinação e posterior degradação da proteína p53 pela oncoproteína E6 de alguns papillomavírus humanos, que assim funcionará como uma ligase da ubiquitina viral (Scheffner *et al.*, 1990). Outro exemplo é a presença de uma enzima de conjugação da ubiquitina da classe II codificada pelo vírus da peste suína africana, expressa continuamente durante a infecção viral (Hingamp *et al.*, 1992; Rodriguez *et al.*, 1992). A presença de uma extensão C-terminal sugere que a E2 viral pode não depender de ligases da ubiquitina do hospedeiro, E3, para o reconhecimento do substrato.

A acção viral sobre o sistema ubiquitina do hospedeiro pode ainda fazer-se por outros mecanismos. Pelo menos dois Baculovirus, um vírus de DNA que infecta células de insectos, contêm um gene essencial que codifica para uma proteína viral tardia relacionada com a ubiquitina (Guarino, 1990; Russell & Rohrmann, 1993; # 4.2). Estas proteínas têm uma região C-terminal idêntica à da ubiquitina e podem ser utilizadas pelo sistema ubiquitina (Finley & Chau, 1991), embora possam não assegurar todas as funções relacionadas com a ubiquitina. Inclusivamente, foi sugerido que podiam agir como inibidoras de algumas reacções de conjugação (Guarino, 1990; Jentsch, 1992). Um outro vírus que infecta células de mamíferos, FBR-MuSV, contém um gene praticamente idêntico a um outro encontrado em vários mamíferos, onde codifica para uma proteína contendo uma região homóloga à ubiquitina fundida à proteína ribossomal S30 (# 4.2). Contudo, no vírus este gene é transcrito como um mensageiro antiparalelo cuja presença duplica a eficiência de transformação do vírus (Michiels *et al.*, 1993).

Alguns genes do sistema ubiquitina do hospedeiro podem também ser induzidos com a infecção por vírus. O gene poliubiquitina *Ubb* humano (mas não os genes *Uba* e *Ubc*) é induzido após infecção de células humanas pelo vírus da herpes, um processo mediado pelo factor de transcrição viral ICP4 (Kemp & Latchman, 1988). Noutros casos, não é a ubiquitina mas os seus homólogos que são induzidos. A UCRP, uma proteína contendo dois domínios homólogos à ubiquitina e capaz de se conjugar com outras proteínas, é induzida especificamente na presença de interferão, e a sua expressão está intimamente correlacionada com a resposta antiviral (Haas *et al.*, 1987; # 4.2.).

Curiosamente, alguns Togavírus BDVD, um vírus de RNA que infecta células de bovinos, parecem usar o sistema ubiquitina de uma forma totalmente distinta. Dois destes vírus, CP1 e Osloss, possuem estirpes citopatogénicas que se distinguem por apresentarem uma inserção de uma unidade e meia de um gene poliubiquitina quase idêntico ao do

hospedeiro (Meyers *et al.*, 1989; 1991). A inserção ocorre dentro de um gene viral, p125, de forma que, após a transcrição e tradução do gene, a proteína resultante é processada de modo equivalente às poliubiquitinas ou proteínas de fusão (# 4.1), produzindo três péptidos: um contendo a N-terminal da proteína viral fundida a uma região C-terminal de ubiquitina, uma ubiquitina idêntica à do hospedeiro e um contendo a região C-terminal da proteína viral. Aparentemente, é a inactivação da proteína viral, clivada pela inserção da ubiquitina, que confere a citopatogenicidade às estirpes, embora não seja de excluir uma função para os péptidos resultantes (Tautz *et al.*, 1993).

3.7. Outros aspectos da ubiquitinação.

Nos últimos anos, várias têm sido as proteínas ou funções celulares encontradas envolvidas directa ou indirectamente com o sistema ubiquitina. Todavia, para além dos exemplos citados anteriormente, ou a acção do sistema é pouco conhecida, ou observa-se apenas uma degradação pelo sistema independente de algum tipo de controlo regulacional. Contudo, há ainda três exemplos que devem ainda ser referidos, quer devido à importância relativa das proteínas envolvidas, quer devido a poderem apresentar aspectos ainda desconhecidos do sistema: o envolvimento da ubiquitina na rede microtubular, e a ubiquitinação do fitocromo e da calmodulina.

O uso de alguns anticorpos antiubiquitina (mas não de todos) em microscopia de imunofluorescência detecta uma estrutura celular complexa identificada com a rede microtubular (Murti *et al.*, 1988). Embora não seja de excluir a existência de uma proteína desconhecida que reaja com o anticorpo, estes dados sugerem que a ubiquitina ou alguns dos seus conjugados possam ser componentes da estrutura microtubular. Os mesmos estudos revelam que a tubulina, principal componente dos microtúbulos, não é ubiquitinada (Murti *et al.*, 1988); todavia, em insectos foi detectada a presença de uma proteína, a artrina, que mais não é do que a actina ubiquitinada (Ball *et al.*, 1987). Esta proteína apenas se encontra presente nos músculos do tórax que actuam indirectamente no voo, e a sua síntese ocorre várias horas após a síntese da actina, não se tratando por isso de um ciclo de ubiquitinação/desubiquitinação da actina (Ball *et al.*, 1987). A ubiquitina é também um dos componentes da rede de filamentos helicais característica de uma doença neurológica, a doença de Alzheimer, um tipo de estrutura que contém ainda outras proteínas microtubulares, como as *tau* (Mori *et al.*, 1987; Mayer *et al.*, 1991; Mescos & Timiras, 1991). Contudo, a função do sistema ubiquitina na rede microtubular da célula é ainda desconhecida, nem foi determinado se a artrina é uma proteína funcional ou apenas um precursor da degradação da actina.

O fitocromo é um fotoreceptor regulatório de plantas que existe em duas formas interconvertíveis: Pr, que absorve luz vermelha, e Pfr, que absorve no vermelho longínquo (Quail, 1991). A forma Pr é convertida à forma Pfr após a absorção da luz vermelha, iniciando, por mecanismos ainda desconhecidos, uma série de respostas que facilitam a adaptação das plantas à disponibilidade de luz fotossintética. A forma Pr é bastante estável, mas a forma Pfr é rapidamente multiubiquitinada e degradada, num processo dependente das condições de radiação luminosa: a transferência de sementes de luz vermelha para o escuro inibe a degradação do fitocromo e promove um decréscimo dos níveis dos conjugados ubiquitina-fitocromo (Shanklin *et al.*, 1987; Jabben *et al.*, 1989).

Tabela II.1. Características dos genes ubiquitina encontrados em vários organismos (págs seguintes).

Organismo	Organização	Nº unid. ^a	aa extra	Poliub.	I ^b	Ub52	I ^b	Ub76	I ^b	Referências
PROTOZOA^c										
<i>Sarcomastigoph.</i>										
<i>Mastigophora</i>										
Cinetoplastidea										
	<i>Trypanosoma brucei</i>	c. 5 genes: 1 ub polimorfica + 2 ub52 agrupados	variável (15/30u)	Leu/Met	2 genes alélicos inc.	N				Wong & Campbell, 1989 Wong <i>et al.</i> , 1992; 1993
	<i>Trypanosoma cruzi</i>	10 genes agrupados: 5 ub/ub52 + 5 ub52	4 a +50 (4u/ub52)	Não tem	1 gene inc. 1 gene	N		Não tem		Swindle <i>et al.</i> , 1988 Kirchhoff <i>et al.</i> , 1988 Z14232
	<i>Leishmania major</i>	3 genes	variável (22/38u)	Arg	1 gene	N		Não tem		Fleischmann & Campbell, 1994
	<i>Leishmania tarentolae</i>	1 gene	1u	?	1 gene inc.	N		Não tem		Krebber <i>et al.</i> , 1994
Diplomonodina	<i>Giardia lamblia</i>									
<i>Sarcodina</i>										
Amoebina	<i>Acanthamoeba castellanii</i>	c. 5 genes ?	? (≥ 1u)	Tyr	1 cDNA inc.					Ahn & Henney, 1994
	<i>Entamoeba histolytica</i>	?	(≤ 3u)		PCR					Wöstmann <i>et al.</i> , 1992
	<i>Quinqueloculina seminulum</i>	?	(≤ 3u)		PCR					Wray & DeSalle, 1994
	<i>Elphidium clavatum</i>	?	(≤ 3u)		PCR					Wray & DeSalle, 1994
	<i>Ammonia beccarii</i>	?	(≤ 3u)		PCR					Wray & DeSalle, 1994
	<i>Affinetrina eburnea</i>	?	(≤ 3u)		PCR					Wray & DeSalle, 1994
	<i>Myxotheca spp.</i>	?	(≤ 3u)		PCR					Wray & DeSalle, 1994
<i>Sporozoa</i>										
<i>Telosporea</i>										
Coccidia	<i>Eimeria bovis</i>	?								Abrahamsen <i>et al.</i> , 1993
<i>Ciliophora</i>										
<i>Ciliata</i>										
Hymenostomatida	<i>Tetrahymena^d pyriformis</i>	≥ 6 genes, agrupados	3 a 5 (5u)	2x SerGln 1x Gln Phe	6 genes inc. 1 gene	N			N	Neves <i>et al.</i> , 1988; 1990; 1991a; 1991b
Hypotrichia	<i>Euplotes eurytomus</i>	≥ 3 genes	3u		1 gene (Ub53) 1 gene macronuclear	N				Hauser <i>et al.</i> , 1991

Organismo	Organização	Nº unid. ^a	aa extra	Poliub.	I ^b	Ub52	I ^b	Ub76	I ^b	Referências
FUNGI										
<i>Myxomycota</i>										
Acridomyceta										
	<i>Dictyostelium discoideum</i>	6-7 genes dispersos	3 a 7 (1x 7u 2x 5u 1-2x 3u)	2x Asn 2-3x Leu	2-3 cDNAs 2 genes	N	1 cDNA	1 cDNA		Giordia & Ennis, 1987 Müller-Taubenberger <i>et al.</i> , 1988b Ohmachi <i>et al.</i> , 1989
<i>Eumycota</i>										
<i>Mastigomycotina</i>										
Oomiceta	<i>Phytophthora infestans</i>	3 - 5 (?) genes	3 a 9 (?) (3u)	Asn	1 gene	N				Pieterse <i>et al.</i> , 1991
<i>Ascomycotina</i>										
Hemiascomyceta	<i>Saccharomyces cerevisiae</i>	4 genes dispersos: 1ub + 2 ub52 + 1 ub76	5u	Asn	1 gene	N	I3	1 gene		Özkaynak <i>et al.</i> , 1987
Pyrenomyceta	<i>Neurospora crassa</i>	2 genes dispersos: 1 ub + 1ub78	4	Gln	1 gene	L8 u1	Não tem	1 gene	T125	Taccioli <i>et al.</i> , 1989; 1991 Wang <i>et al.</i> , 1993
<i>Basidiomycotina</i>										
Hymenomyceta	<i>Coprinus congregatus</i> <i>Phanerochaete chrysosporium</i>	? ?	? (≥ 5u) ? (5u)	? Leu	1 cDNA inc. 1 gene					Foster <i>et al.</i> , 1990 Z24723

PLANTAS	Organismo	Organização	Nº unid. ^a	aa extra	Poliub.	I ^b	Ub52	I ^b	Ub76	I ^b	Referências
Chlorophyta											
<i>Phycomonodina</i>											
Volvocales	<i>Chlamydomonas reinhardtii</i> <i>Volvox carteri</i>	≥ 6 genes ≥ 4 genes	até 8-9 (5u)	? Leu	1 gene	G35 (u1-5)	1 cDNA				Pollmann <i>et al.</i> , 1991 Schiedlmeier & Schmitt, 1994
<i>Chlorophyceae</i>	<i>Acetabularia^d cliftonii</i>	?	? (≥ 6u)	AlaPhe	1 gene	G35 u1?					Z28699
Phaeophyta											
<i>Dinophyceae</i>	<i>Symbiodinium spp.</i>	?	(≤ 3u)		PCR						Wray & DeSalle, 1994
Zooxanthellales											
Rhodophyta											
<i>Rhodophyceae</i>	<i>Aglaothamnion neglectum</i>	?	? (6u)	Met	1 cDNA						M88684
Ceramiales											
Chrysophyta	<i>Stephanodiscus yellowstonensis</i>	?	(≤ 3u)		PCR						Wray & DeSalle, 1994
<i>Bacillarioficeae</i>											
Magnoliophyta											
<i>Liliopsida</i>											
Poales	<i>Hordeum vulgare</i> <i>Avena fatua</i> <i>Triticum aestivum</i> <i>Oryza sativa</i> <i>Zea mays</i>	5-12 (?) genes: 2-3 ub79, 2 agrup. ≥ 7 genes ? ≥ 4 genes ≥ 4 genes ≥ 2 genes	3-4 a 7-8 3 a 8-9 (4u) ? (> 1u) até 7-8 (≥ 3u; 6u) 4 a 6 (2x 6u) ? (≥ 3u)	Lys Gln Gln Leu e Gln Gln AspMet	1 cDNA inc. 1 cDNA 1 cDNA inc. 1 cDNA inc. 2 genes 1 cDNA inc.	(0) 0	1 cDNA (ub53)		2 genes (ub79) 1 gene (ub79) 1 cDNA (Ub79)		Gausling & Barkardottir, 1986 Gausling & Jensen, 1990 Reynolds & Hooley, 1992 X56601 Nishi <i>et al.</i> , 1993 X76064; D12776 Chen & Rubenstein, 1991 Christensen <i>et al.</i> , 1992 Davies & King, 1993
Liliales	<i>Asparagus officinalis</i>	≥ 2 genes	?		1 cDNA inc.						

Magnoliopsida											
Caprales		?	5 a 7-8 (1x5u)	1x SerPhe	1 cDNA 3 genes 1 ψ gene	0	1 cDNA 2 genes	G35 E64 K98 Q120	2 genes (ub81)	L21898	
Fabales	<i>Brassica rapa</i>	9 a 11 genes dispersos: ≥ 3 ub + 1 ψ gene + 3 ub52 + 2 ub81	5 a 7-8 (1x5u)	1x SerPhe	1 cDNA 3 genes 1 ψ gene	0	2 genes	G35 E64 K98 Q120	2 genes (ub81)	Burke <i>et al.</i> , 1988 Callis <i>et al.</i> , 1990 Sun & Callis, 1993 Norris <i>et al.</i> , 1993 X17008	
	<i>Pisum sativum</i>	≥ 2 genes	? (5u)	Phe	2 cDNAs	(0)				D16248; D28123	
	<i>Glycine max</i>	≥ 2 genes	? (2x 4u)	2x Phe	2 gene	(0)				Jacinto <i>et al.</i> , 1994 Jacinto, 1992 X54381	
	<i>Lupinus albus</i>	≥ 3 genes	?	?	1 gene inc.	(0)				Agarwal & Cullis, 1991	
	<i>Lupinus polyphyllus</i>	?	? ($\geq 2u$)	Phe	1 cDNA inc.					Kawalleck <i>et al.</i> , 1993	
Geraniales	<i>Linum usitatissimum</i>	8 a 11 genes: ≥ 3 ub(52-78)	4 a 6	?	5-6 genes		? (até 3 genes)		? (até 3 genes)	Garbarino <i>et al.</i> , 1992	
Umbellales	<i>Petroselinum crispum</i>	≥ 2 genes	6 (2x 6u)	AspPhe	1 cDNA 1 gene	0			1 cDNA (ub80)	Hoffman <i>et al.</i> , 1991	
Polemoniales	<i>Solanum tuberosum</i>	≥ 7 genes: ≥ 3 ub, 1-2 c/ 2 cópias + 2-3 ub80	6 a 7 (1x 6u 2x 7u)	1x AspPhe 2x Phe	3 cDNAs	(0)			1 gene (ub80)	Genschik <i>et al.</i> , 1992	
Scrophulariales	<i>Lycopersicon esculentum</i>	≥ 2 genes	c. 6		4 cDNAs	(0)				X53011	
	<i>Nicotiana sylvestris</i>	6 a c. 10 genes: 2 ub52	5 a 7-8 ($> 6u$, 6u, 5u, $> 4u$)	2x Phe						X67957	
	<i>Nicotiana tabacum</i> ^e	?	? ($\geq 4u$)	Phe	1 cDNA inc.			Y100		X14333 Binet <i>et al.</i> , 1991	
Rubiales	<i>Antirrhinum majus</i>	?									
	<i>Helianthus annuus</i>	5 a c. 10 genes	5 a 6 3x 6u	1x Phe 2x Asn	1 cDNA 2 genes	0					

ANIMAIS	Organismo	Organização	Nº unid. ^a	aa extra	Poliub.	I ^b	Ub52	I ^b	Ub76	I ^b	Referências
<i>Poriphera</i>											
<i>Demospingiae</i>											
<i>Choristida</i>	<i>Geodia cydonium</i>	1 gene (?)	6u	Phc	1 cDNA		Não detectado		Não detectado		Pfeifer <i>et al.</i> , 1993
<i>Platyhelminthes</i>											
<i>Cestoda</i>											
<i>Cyclophyllidea</i>	<i>Echinococcus multilocularis</i>	?	? (1u)	Não tem	1 cDNA						L23315
<i>Aschelminthes</i>											
<i>Nematoda</i>											
<i>Rhabditoidea</i>	<i>Caenorhabditis elegans</i>	2 genes	11	AspIle	1 gene	-5 G47 (u1,4,7,10)	1 gene		Não tem ^f		Graham <i>et al.</i> , 1989 Jones & Candido, 1993
<i>Echiuroidea</i>											
<i>Echinoidea</i>											
<i>Echinoidea</i>	<i>Strongylocentrotus purpuratus</i>	1 gene (?)	c. 8 (≥ 2u)	Gln	1 cDNA inc.		Não detectado		Não detectado		Nemer <i>et al.</i> , 1991
<i>Arthropoda</i>											
<i>Insecta</i>											
<i>Diptera</i>	<i>Drosophila melanogaster</i>	≥ 6 genes: ≥ 4 ub + 1ub80 + 1ub52	até c. 18	IleGlnAla	1 gene	-11	1 gene	-13	1 gene (ub80)	-20 K107	Barrio <i>et al.</i> , 1994 Arribas <i>et al.</i> , 1986 Lee <i>et al.</i> , 1988 Cabrera <i>et al.</i> , 1992
<i>Lepidoptera</i>	<i>Manduca sexta</i>	≥ 4 genes: 1 ub79	até c. 18	?					1 cDNA (ub79)		Bishoff & Schwartz, 1990
	<i>Bombyx mori</i>	?	? (3u)	Asn	1 cDNA						X62484
	<i>Spodoptera frugiperda</i>	?	c. 12	?	1 gene						Guarino, 1990
<i>Crustacea</i>											
<i>Branchiopoda</i>	<i>Artemia franciscana</i>	?	(9u)	?S	1 gene						X74405

Chordata												
Amphibia												
Anura		<i>Xenopus laevis</i>	≥ 6 genes	4 a > 10 (≥ 3u)	Não tem	1 cDNA						Dworkin-Rastl <i>et al.</i> , 1984
Aves												
Galliformes		<i>Gallus gallus</i>	≥ 3 genes dispersos: 2 ub + 1 ub52	3 e 4	Asn e Tyr	2 genes	-6	1 cDNA				Bond & Schlesinger, 1986 Mezquita <i>et al.</i> , 1988; 1993
Mammalia												
Artiodactyla		<i>Sus scrofa</i>	?	? (≥ 3u)	Phe	1 cDNA inc.						M18159
		<i>Bos taurus</i>	?	? (4u, ≥ 3u)	Cys e Val ^h	2 cDNAs	(-6)					M37793 M37794/Z18245
Rodentia		<i>Mus musculus</i>	≥ 3 genes	4u e 10u+1	Tyr	2 cDNA ⁱ	(-6)	1cDNA inc.				X51703; Finch <i>et al.</i> , 1992
		<i>Rattus rattus</i>	≥ 3 genes	4u e 10u+1	Tyr	2 cDNAs ⁱ	(-3)					St.John <i>et al.</i> , 1986
		<i>Cricetulus griseus</i>	≥ 3 genes	5u e 8u+1)	Tyr	2 genes ⁱ	(-6)					Hayashi <i>et al.</i> , 1994
Primata		<i>Homo sapiens</i>	c. 16 genes dispersos, 4 funcionais: 2 ub + 4 ψgenes 1 ub52 + c. 8 ψgenes 1 (2?) ub80	3 e 9	Cys e Val	2 genes, 4 ψgenes	-6 -3	1gene 2 ψgenes				Nenoi <i>et al.</i> , 1992; 1994
												Lund <i>et al.</i> , 1985 Wiborg <i>et al.</i> , 1985 Baker & Board, 1987; 1991; 1992 Cowland <i>et al.</i> , 1988 Y00332

(a) Entre parênteses, o número de unidades dos genes poliubiquitina conhecidos.

(b) Intrões (I); N, ausentes. Quando presentes, são indicados os aminoácidos e unidades interrompidas ou a posição a montante do ATG (em valores negativos). Entre parênteses, locais cuja homologia com outros genes torna possível a presença de um intrão.

(c) Consideradas as classificações em **REINO**, **phylum**, **classe** e ordem.

(d) Organismos que não seguem o código universal.

(e) Provavelmente um fungo contaminante de *Nicotiana* (# 4.1.2).

(f) A proteína de fusão de 79 aminoácidos encontra-se fundida neste organismo a um homólogo da ubiquitina (# 4.2 e Tabela II.2).

(g) O gene contém 13 resíduos extra, mas podem ser consequência de uma unidade mutada.

(h) Supondo um erro de sequência, detectável por comparação com o gene humano homólogo; no original, existem 6 resíduos extra.

(i) Os genes homólogos do gene humano UbC (9u) contém uma unidade terminal com uma deleção interna que origina uma curta proteína não-ubiquitina.

Todavia, a irradiação contínua com luz vermelha das sementes promove uma acumulação contínua dos conjugados, uma observação inesperada pelo facto de que a degradação do fitocromo ser iniciada imediatamente após a irradiação (Jabben *et al.*, 1989). É possível que a taxa de degradação seja diminuída por alterações secundárias após tempos longos de irradiação, ou então que exista uma população heterogénea de conjugados ubiquitina-fitocromo, uns mais susceptíveis a uma rápida degradação do que outros (Hershko & Ciechanover, 1992). A forma Pfr sofre alterações conformacionais que, por exemplo, a tornam mais susceptível a agregações, sendo as formas ubiquitinadas encontradas predominantemente na fracção agregada (Jabben *et al.*, 1989). Isto sugere que o processo que promove a agregação poderá igualmente promover a degradação, mas não é de excluir a activação de uma ligase da ubiquitina, E3, específica do fitocromo seja uma das variadas respostas celulares induzidas pela fotoconversão deste receptor (Hershko & Ciechanover, 1992).

A calmodulina é uma das muitas proteínas cuja degradação se sabe ser mediada pelo sistema ubiquitina (Ziegenhagen & Jennissen, 1990). Contudo, essa degradação é dependente da presença de Ca^{2+} , o que torna a calmodulina a única proteína conhecida cuja ubiquitinação parece ser regulada por um segundo factor (Laub & Jennissen, 1991; Jennissen *et al.*, 1992). A purificação de extractos de lisado de reticulócitos em colunas de ubiquitina-sefarose permitiu identificar duas fracções capazes de promover a degradação da ubiquitina: a fracção retida, Syn F1, capaz de ubiquitinar e degradar a calmodulina independentemente da presença de Ca^{2+} , conterá provavelmente as enzimas E1 e E2 do sistema ubiquitina; a fracção eluída, Syn F2, é incapaz de degradar a calmodulina, mas restaura a dependência de Ca^{2+} quando adicionada à primeira fracção (Majetschak *et al.*, 1993). A fracção eluída conterá provavelmente uma ligase da ubiquitina (Parag *et al.*, 1993), restando determinar se reconhece preferencialmente Ca^{2+} -calmodulina ou se a sua actividade é ela própria dependente da presença de Ca^{2+} .

4. Os genes ubiquitina.

4.1 Estrutura dos genes ubiquitina.

A estrutura dos primeiros genes ubiquitina isolados, em levedura e *Xenopus*, constituiu uma verdadeira surpresa (Özkaynak *et al.*, 1984; Dworkin-Rastl *et al.*, 1984). De facto, em vez de um simples região de 228 pares de bases necessária à codificação de uma proteína de 76 aminoácidos, os genes encontrados apresentavam várias repetições de 228 pb fundidas em série, sem sequências espaçadoras. No ano seguinte, Lund e colaboradores (1985) encontravam em humanos um gene ubiquitina em que a região codificante de 228 pb se encontrava fundida na região 3' a uma sequência nucleotídica não relacionada, capaz de codificar para um péptido de 52 aminoácidos. Os genes ubiquitina apresentavam então uma organização praticamente única entre os genes eucariotas. De facto, apesar de a presença de genes codificando para poliproteínas ser já amplamente conhecida (para revisão, ver Douglass *et al.*, 1984), os péptidos resultantes após o processamento são usualmente distintos entre si, ou no caso dos dois únicos genes conhecidos que codificam para péptidos idênticos, a preprocaeruleína (Hoffmann *et al.*, 1983) e o preprofactor de α -mating (Kurjan & Herskowitz, 1982), as repetições encontram-se separadas por sequências espaçadoras. Desde então, os genes ubiquitina têm sido isolados em todos os organismos eucariotas onde têm sido pesquisados (Tabela I1.1). Alguns genes têm sido também encontrados em vírus, embora codificando para proteínas mais divergentes do que as ubiquitinas eucariotas,

Tabela II.2. Características dos genes homólogos aos genes ubiquitina.

Organismo	Classe	Genes	Observações	Referências
AcMNPV	A	1 V-ubi	aa extra: Tyr	Guarino, 1990
OpMNPV	A	1 V-ubi	aa extra: NGLRK GKRRCLSL LQFI	Russell & Rohrmann 1993
FBR-MuSV	B	1 V-fau	Transcrito como antisenso	Michiels <i>et al.</i> , 1993
<i>Saccharomyces cerevisiae</i>	B	? (1 clonado) RAD23	Fusão não processado Envolvido na reparação do DNA	Watkins <i>et al.</i> , 1993
<i>Caernohabditis elegans</i>	B	2 (1 clonado) Ubl	Fusão processado Equivalente à Ub79 Intrão G45 (= G47)	Jones & Candido, 1993
<i>Caernohabditis briggsae</i>	B	2 (1 clonado) Ubl	Fusão processado Equivalente à Ub79 Intrão G45 (= G47)	Jones & Candido, 1993
<i>Xenopus laevis</i>	C	2 An1a; An1b	Homologia aa 28/103 Fusão não processado	Linnen <i>et al.</i> , 1993
<i>Mus musculus</i>	A	1 (1 cDNA) Nedd-8	Extra aa: GGLGQ	Kumar <i>et al.</i> , 1993
	B	≥ 3 genes M-fau	Fusão não processado? Codifica proteína ribossomal	Michiels <i>et al.</i> , 1993
<i>Rattus rattus</i>	B	≥ 3 genes R-fau	Fusão não processado? Codifica proteína ribossomal S30	X62671
<i>Homo sapiens</i>	C	1 BAT3	Homologia aa 18/93 Intrões: (=S20, =N60)	Banerji <i>et al.</i> , 1990
	A	1 UCRP	=diubiquitina não processada Extra aa: GTEPGGRS Intrão: M1	Reich <i>et al.</i> , 1987
	B	? (2 cDNAs) HHR23A; HHR23B	Não processados Envolvidos na reparação do DNA	Masutani <i>et al.</i> , 1994
	B	1 GdX	Fusão não processado Intrões. (=E16, =D52), R121	Toniolo <i>et al.</i> , 1988
	B	≥ 3 genes H-fau	Fusão não processado? Codifica proteína ribossomal Intrões: -8, (=K27, =G76), K93	Kas <i>et al.</i> , 1992

pelo que têm sido normalmente considerados como genes homólogos aos das ubiquitinas e não como verdadeiros genes ubiquitina (Tabela I1.2). A ubiquitina aparenta ser uma proteína exclusivamente eucariota, já que não tem sido detectada nem em procariotas nem em vírus com eles relacionados. Contudo, foi parcialmente sequenciada na arqueobactéria *Thermoplasma acidophilum* uma proteína idêntica à ubiquitina (Wolf *et al.*, 1993). Este resultado carece todavia ainda de confirmação, já que anticorpos antiubiquitina não detectam nenhum sinal neste organismo.

De acordo com a sua organização, os genes ubiquitina foram classificados por Warner (1989) em duas classes (ou em três, segundo Jentsch *et al.* (1991)). A classe I (classe III segundo Jentsch), ou genes poliubiquitina, é constituída por várias repetições directas de 228 pb fundidas em série, e de um a três codões extra após a última unidade. A classe II, ou ubiquitinas de fusão, é constituída por unidades monoméricas de 228 pb fundidas a sequências não relacionadas de 52-53 codões (classe I segundo Jentsch) ou de 76-81 codões (classe II segundo Jentsch).

4.1.1. Classe I: genes poliubiquitina.

A ubiquitina é uma das proteínas mais conservadas ao longo da escala filogenética que se conhecem. De facto, dos seus 76 aminoácidos, 49 são absolutamente invariantes, mesmo quando são comparadas sequências de eucariontes tão distantes como o primitivo *Entamoeba histolytica* e o homem (Fig.I1.4). Ainda assim, se forem excluídos das comparações alguns grupos de organismos muito divergentes, como os protozoários foraminíferos, diplodominos e as algas, a homologia encontrada não é inferior aos 91% encontrados entre o protozoário *Trypanosoma brucei* e o fungo *Phytophthora infestans*. Esta extrema conservação da sua estrutura primária sugere que a maioria dos seus aminoácidos são requeridos para funções próprias.

A análise das sequências de ubiquitina encontradas não permite concluir por uma correlação entre determinadas substituições e grupos de organismos. De facto, à excepção de alguns *phylum*, como os animais cordados ou as plantas magnoliófitas, que codificam para uma mesma proteína, todos os outros grupos apresentam uma variação interna das sequências. Se certas substituições podem ser associadas a determinados organismos, como a Ala13 aos protozoários cinetoplastídeos, outras, como a Ala57, encontram-se associadas a organismos tão distintos como as plantas e algas verdes, ou alguns protozoários ciliados, foraminíferos e cinetoplastídeos. Estas características da ubiquitina têm, aliás, impedido o seu uso para a construção de árvores filogenéticas (Wöstmann *et al.*, 1992; Wray & DeSalle, 1994).

Os genes poliubiquitina são a principal forma de manutenção da informação genética da ubiquitina em eucariotas, já que constituem a única classe de genes ubiquitina que tem sido inquestionavelmente encontrada em todos os organismos onde procurada (# 4.3). Estes genes são repetições organizadas em série de sequências de 228 pb, cada uma codificante para a ubiquitina, sem sequências espaçadoras. O gene codifica, então, para uma poliproteína que é imediatamente hidrolisada em monómeros ubiquitina. O número de repetições varia com o gene, e pode ir de 3 a mais de 50, apesar de alguns organismos como *G.lambliia* ou *E.multicularis* possuírem genes com uma única unidade (Tabela I1.1). A sequência codificante é raramente interrompida por intrões, incluindo as únicas excepções conhecidas dois fungos, *N.crassa* e *P.chrysosporium*, duas algas, *V.carteri* e *A.cliftonii*, e um nemátodo, *C.elegans*. Apenas as algas apresentam uma localização comum do intrão,

	1	10	20	30	40	50	60	70
H. sapiens^a	MQIFVKTLTGKTTITLEV	PSDPTIENVKAKIQDKEG	IPPDQQRLLIFAGKQLEDG	RTLSDYNIQKESTLHLVLR	RG			
<i>S. purpuratus</i>	S.....							
<i>C. elegans^b</i>	A.....							
<i>E. multilocularis</i>	S.....				R.....			
<i>A. thaliana^c</i>	S.....D.....					A.....		
<i>S. yellowstonensis</i>	D.....H.N..T.....		N.....V.....			N.....		
<i>A. neglectum</i>	S.....T.....					N.....		
<i>S. spp.</i>	D.N.S.....					A.....		
<i>C. reinhardi^d</i>	S.....					A.....		
<i>C. congregatus^e</i>	S.....D.....							
<i>N. crassa</i>	S.....D..Q.....							
<i>S. cerevisiae</i>	S.....D..S.....							
<i>P. infestans</i>	D.....S.....Q.....							
<i>D. discoideum^f</i>	G.N.....							
<i>E. eurystomus</i>	D.Q.....D..T.....					A.....		
<i>T. pyriformis</i>	D.A.....							
<i>E. bovis</i>	D.....							
<i>M. spp.</i>	S.....RS.....					A.....		
<i>A. eburnea</i>	D.....D..T.....							
<i>A. beccari</i>	D.N..Q.....			E.....				
<i>E. clavatum</i>	D.AN..T.....			E.....	D.....			
<i>Q. seminulum</i>	S.....D.....							
<i>E. histolytica</i>	N.S.DAI...E.....				E.K.....			
<i>A. castellani</i>	S.....Q.....					A.....		
<i>G. lamblia</i>	I.....T.N.I.....			S.....N.Q.S..DA.....				
<i>L. major^g</i>	A.....				E.....			
<i>T. cruzi</i>	A..S.....					A.....		
<i>T. brucei</i>	A...A.....				E.....E.....A.....			
<i>A. castellani</i>	S.....					A.....		

Fig. 11.4. Sequências de ubiquitinas encontradas em vários organismos. Apenas os resíduos divergentes da sequência humana foram indicados; os pontos representam resíduos idênticos. ^a Inclui todos os cordados e artrópodes conhecidos; ^b Inclui todos os asquelminas e poríferos conhecidos; ^c Inclui todas as magnoliófitas conhecidas; ^d inclui todas as clorofitas conhecidas; ^e Inclui todos os basidiomicotinos conhecidos; ^f Três das cinco unidades do gene conhecido apresentam uma Thr na posição 28; ^g Inclui todas as *Leishmania* conhecidas.

no resíduo Gly35, mas enquanto em *V.carteri* todas as unidades são interrompidas, em *A.cliftonii* apenas uma das seis conhecidas o é. O número de unidades interrompidas varia igualmente nos outros três organismos: *N.crassa* possui apenas a sua primeira unidades interrompida (em Leu8), mas *P.chrysosporium* possui três em cinco (em Val70) e *C.elegans* quatro em onze (em Gly47) (Tabela II.1). Apesar disso, os genes poliubiquitina são frequentemente interrompidos por intrões, embora localizados preferencialmente na região 5' transcrita mas não traduzida ("leader"). Todos os animais parecem possuir intrões, localizados a -6 ou a -3 do ATG em vertebrados e a -11 em insectos. O nemátodo *C.elegans*, para além dos intrões já referidos, possui igualmente um localizado a -5 do ATG, mas que, como é característico neste organismo, é eliminado por *trans*-splicing (Graham *et al.*, 1989). As plantas magnoliófitas também possuem um intrão na "leader", localizado imediatamente antes do ATG. Já noutros reinos a situação é diferente: enquanto os intrões parecem não existir em protozoários, em fungos são encontrados nas regiões codificantes de *N.crassa* e *P.chrysosporium*, como referido, mas não em *S.cerevisiae* ou *P.infestans*.

Praticamente todos os genes poliubiquitina possuem ainda de um a três codões extra após a sua última unidade. As únicas excepções conhecidas são em *Xenopus laevis* e em *Echinococcus multilocularis*, que aparentemente não possuem nenhum, embora sejam resultados que necessitem de confirmação. No boi, Meyers e colaboradores (1991) referem também um cDNA com seis resíduos extra; contudo, pode tratar-se de um erro de sequência, que após correcção conduz a um único resíduo extra codificante para Val, tal como o gene homólogo em humanos. A presença destes codões extra poderá constituir um sistema de bloqueio da região C-terminal, impedindo assim o envolvimento prematuro da poliubiquitina no processo de conjugação. Dos vinte aminoácidos existentes, doze já foram encontrados nesta posição, o que sugere a inexistência de qualquer conservação nos resíduos extra. Contudo, a verdade é que se observa uma clara preferência por cinco aminoácidos, Phe, Gln, Asn, Leu e Tyr, que representam 80% da mais de meia centena de resíduos extra conhecidos. Pode-se igualmente observar alguns padrões dentro de algumas classes, embora não seja claro se isso corresponde a uma homologia interna entre as ubiquitinas de cada grupo ou a diferentes propriedades funcionais das hidrolases da ubiquitina. Assim, os fungos e as plantas liliopsidas (monocotiledóneas) parecem preferir resíduos amidados, como Gln e Asn, enquanto as plantas magnoliopsidas (dicotiledóneas) possuem um ou dois resíduos extra habitualmente terminados em Phe (Tabela II.1). O insecto *D. melanogaster* é o único a possuir três codões extra, o que pode não ser uma característica comum a outros insectos. Já os animais cordados possuem um único resíduo extra, usualmente Tyr ou Cys para um dos seus dois genes, e Val para o outro. Em roedores, este último gene sofreu uma deleção interna de 64 nucleótidos sua última repetição, o que originou uma unidade truncada e com uma região C-terminal de 30 aminoácidos não relacionada com a ubiquitina (Nenoi *et al.*, 1994). Contudo, é curioso notar que, caso a deleção não tivesse ocorrido, estes genes terminariam igualmente em Val.

Todos os genes poliubiquitina conhecidos são homogéneos, isto é, todas as suas unidades codificam para uma proteína idêntica. Foi no entanto encontrada uma excepção num cDNA de *Dictyostelium*, em que três das suas cinco unidades codificam para uma ubiquitina diferente das outras duas ou das de outros genes do mesmo organismo (Fig.II.4). Esta homogeneidade deve-se possivelmente à ocorrência de uma evolução concertada neste tipo de genes (Sharp & Li, 1987a). Segundo este modelo evolutivo, as mutações ocorridas numa unidade ubiquitina são capazes de se propagar rapidamente às outras unidades e a outros genes por recombinação, ou são eliminadas pelas unidades preexistentes. Este

processo é bastante evidente em alguns organismos, onde a homogeneidade se estende inclusive às sequências nucleotídicas (Sharp & Li, 1987a; Neves *et al.*, 1991b). Este modelo justificará então a presença de três Thr28 em vez de Ala28 em *Dictyostelium* como sendo uma mutação recente em fase de propagação dentro do gene (Sharp & Li, 1987a).

4.1.2. Classe II: genes ubiquitina-proteínas ribossomais.

A outra classe de genes ubiquitina consiste na fusão de um monômero ubiquitina com uma sequência não relacionada com 52-53 (Fig.II.5A) ou 76-81 codões (Fig.II.5B). Estas sequências são também muito conservadas ao longo da escala filogenética, sendo cerca de 44 a 46% dos seus resíduos comuns a todos os genes encontrados. Ambas as sequências codificam para proteínas de características extremamente básicas com dois motivos comuns: um domínio "zinc finger" de ligação a ácidos nucleicos (Klug & Rhodes, 1987) determinado pela presença de 4-5 Cys de posições conservadas, e uma sequência de resíduos básicos semelhante à sequência de translocação nuclear de proteínas (Kalderon *et al.*, 1984), situada na região C-terminal da proteína de 52-53 aminoácidos e na N-terminal da proteína de 76-81 aminoácidos.

Apenas os genes codificantes para a proteína de fusão de 52-53 aminoácidos (Ub52) de protozoários parecem não conter intrões (Tabela II.1). A levedura apresenta um intrão a interromper o codão Ile3 enquanto o gene de *Drosophila* apenas possui um na região "leader" a -13 do ATG. Nos outros dois genes de eucariotas superiores estudados, os do homem e da planta *Arabidopsis*, existem quatro intrões, três dos quais têm uma localização comum: nos codões Gly35 e Glu64 da metade ubiquitina e Lys22 da proteína de extensão. A planta possui ainda um outro no codão Gln120, enquanto o homem possui um localizado a -8 do ATG. Não é de excluir, contudo, a presença de intrões na "leader" de outros organismos, já que não foram realizados estudos conducentes à sua detecção. Contrariamente à *Arabidopsis*, a planta *Nicotiana tabacum* possui um intrão localizado no codão Tyr24 da proteína de extensão; contudo, este gene deverá pertencer a um fungo contaminante e não a uma planta, já que a sua proteína de extensão é extraordinariamente divergente da da planta *Nicotiana sylvestris*, mas bastante homóloga à da levedura (ver Fig.II.5A). É curioso ainda notar que, das 15 sequências conhecidas, apenas três codificam para uma proteína de 53 aminoácidos em vez dos 52 habituais, pertencendo duas a protozoários não cinetoplastídeos e a outra à única monocotiledónea estudada, o arroz. Por outro lado, a proteína de extensão de *D.discoideum*, embora possua 52 aminoácidos, está "deslocada" em relação ao consenso dos outros organismos.

Curiosamente, os genes codificantes para as proteínas de extensão de 76-81 aminoácidos (Ub79) nunca foram detectados em protozoários, sendo inclusive significativo o facto de os estudos genómicos efectuados em três protozoários mastigoforos indicarem a sua provável ausência nestes organismos. A sua presença tem sido contudo assinalada nos outros reinos, embora em *Caernohabditis* os genes encontrados não pertençam verdadeiramente à classe II dos genes ubiquitina (# 4.2). Apenas dois dos genes encontrados possuem intrões, um localizado no codão Thr50 da proteína de extensão do fungo *N.crassa*, e dois localizados a -20 do ATG e no codão Lys31 da proteína de extensão do insecto *D.melanogaster*. Esta proteína de extensão apresenta uma razoável variação no seu tamanho, que se deve fundamentalmente à variabilidade da sua região C-terminal. Contudo, é de notar a presença de um aminoácido extra na posição 35 em monocotiledóneas e o

"deslocamento" da proteína em *D.discoideum* em relação ao consenso dos outros organismos.

O papel biológico das ubiquitinas de fusão é ainda muito pouco claro. Inicialmente, devido à presença de motivos de "zinc finger" e de translocação nuclear, pensou-se que as proteínas de extensão se poderiam associar ao DNA; contudo, foi já demonstrado na levedura, em *Arabidopsis*, em *Dictyostelium* e no rato que estas proteínas co-sedimentam com a fracção ribossomal (Finley *et al.*, 1989; Callis *et al.*, 1990; Müller-Taubenberger *et al.*, 1989; Redman & Rechsteiner, 1989), indicando que são constituintes ribossomais. Assim, a Ub79 é constituinte da subunidade 40S do ribossoma, sendo identificada como a proteína ribossomal S27a do rato (Redman & Rechsteiner, 1989) e S37 da levedura (Finley *et al.*, 1989). Todavia, a localização da Ub52 é contraditória: em levedura e em *Arabidopsis* co-sedimenta com a subunidade 60S do ribossoma (Finley *et al.*, 1989; Callis *et al.*, 1990), mas em *Dictyostelium* cosedimenta com a subunidade 40S (Müller-Taubenberger *et al.*, 1989). O papel destas proteínas no ribossoma não é conhecido, mas, em levedura, a Ub52 é essencial à sobrevivência da célula, enquanto a Ub79 parece ser necessária mas não essencial, já que a sua completa remoção afecta o crescimento mas não a viabilidade das células (Finley *et al.*, 1989).

A função da ubiquitina no gene de fusão é ainda menos clara. Mutantes de levedura em que as metades ubiquitina foram removidas dos genes de fusão são perfeitamente viáveis, sendo a ubiquitina fornecida pelo gene poliubiquitina (Finley *et al.*, 1989). Todavia, um mutante em que a ubiquitina foi removida do gene Ub79 só adquire taxas de crescimento equivalentes às da estirpe selvagem se a proteína de extensão existir em multicópia. Estes dados parecem indicar um possível papel da ubiquitina como "chaperone" da proteína ribossomal na montagem do ribossoma (Finley *et al.*, 1989). Contudo, a proteólise da proteína de fusão Ub79 ou de outras proteínas fundidas à ubiquitina é muito rápida em leveduras e em células animais (Redman & Rechsteiner, 1988; 1989; Monia *et al.*, 1989; Ecker *et al.*, 1989), o que parece entrar em desacordo com esta hipótese. É possível que a ubiquitina sirva apenas como um meio eficiente de se conseguirem proteínas ribossomais sem Met inicial, evitando um bloqueamento N-terminal (Warner, 1989), ou, uma vez que as proteínas podem ser estabilizadas por fusão à ubiquitina, funcione como um meio de estabilizar as proteínas ribossomais antes da sua incorporação em ribossomas (Butt *et al.*, 1989; Ecker *et al.*, 1989). Estas hipóteses justificariam a aparente ausência de genes ubiquitina da classe II em alguns organismos, pelo menos ligados a ubiquitinas canónicas (Jones & Candido, 1993; # 4.2). Contudo, não é de excluir um papel importante para a ubiquitina neste tipo de genes: em *Dictyostelium*, a proteína de fusão Ub52 não é processada nas suas metades ubiquitina-proteína ribossomal, e encontra-se associada por inteiro aos ribossomas (Müller-Taubenberger *et al.*, 1989). Também em levedura a proteólise é incompleta, mas só a proteína ribossomal é encontrada nos ribossomas (Finley *et al.*, 1989; Monia *et al.*, 1989; Müller-Taubenberger *et al.*, 1989). Já em plantas a informação é contraditória: em *Arabidopsis* a proteólise é completa, mas no trigo existem resultados opostos de laboratórios diferentes (Müller-Taubenberger *et al.*, 1989; Callis *et al.*, 1990). Esta inibição da proteólise, pelo menos na Ub52, permite levantar a hipótese de que, afinal, poderá ser a proteína ribossomal a funcionar como "chaperone" no transporte da ubiquitina para o núcleo, onde se conjugará às histonas H2A e H2B (Goldknopf & Busch, 1977; Thorne *et al.*, 1987).

A

	1	10	20	30	40	50
<i>T. brucei</i>	VMEPTLEALAKYNNWEKKVCRRCYARLPVRATNCRKKGCGHCNSLRMKKKLR					
<i>T. cruzi</i>		S.....A.....				
<i>L. major</i>	V.....	A.....				
<i>T. pyriformis</i>	G...IA.....C.....D.....PK.....RK.....SNS..L...PKE					
<i>E. bovis</i>	II..S.AL..S..C..I..K.....P.....T.Q.....PKN					
<i>N. tabacum</i>	II..S.K...S.F.CD.MI..K.....P.....RK...TNHV.P...K					
<i>S. cerevisiae</i>	Y.....SV.....	QL.....				
<i>D. discoideum</i>	-...VI..R.YK.....	H..V..K...S.N.....LK				
<i>A. castellanii</i>	V.....QV..R.....VV.....	H..V..K...S.W.....				
<i>C. reinhardtii</i>	Q...R.Y.QE.....	H..K...KS...Q.....				
<i>O. sativa</i>	Q...R.Y.Q.....	H..V..K...S.Q...I.N				
<i>A. thaliana</i>	MM..R.Y.Q.....	H..V..K...S.Q...I.				
<i>N. sylvestris</i>	M...R.Y.Q.....	H..V..RK...S.Q...I.				
<i>B. rapa</i>	M...R.Y.QA.....	H..V..K...S.Q...I.				
<i>D. melanogaster</i>	RI..Q.Y.....	H.....K...N.....				
<i>H. sapiens</i>	RQ..Q.Y.....	H..V..K...N.....V.				

B

	1	10	20	30	40	50	60	70	79
<i>Z. mays</i>	AKRRKKKTYTKPRKIKHKHKVKVLAVLFQYKVDATGKVTRLRKECPNTECGAGVFMANHFDHRHYCGKGLTYVEQKS								
<i>H. vulgare</i> (i)		Q.....	S.....	AD...T.....					
<i>H. vulgare</i> (ii)		Q.....	S.....	AD...T.....					
<i>L. esculentum</i>		K.....	Q.....	A...T.....	GGD				
<i>A. thaliana</i> (i)			I.G-S..Q.....	AT...T...S.....	E.VEA				
<i>A. thaliana</i> (ii)			G-S..Q.....	AT...T...S.....	E.AQE				
<i>L. albus</i> (i)			-S..Q.....	T.....K.....	ADA				
<i>L. albus</i> (ii)			-S..Q.....	T.....Q.....	ADA				
<i>S. cerevisiae</i>	G.....V..T.....	SY.....	-E...K..R..S.PT.....	L...K..L.....	HSV.KVNA				
<i>N. crassa</i>	G.....V..T.....	KY.....	-SD..IE...R...ET.....	AMQ..Q...R.H....	F...S				
<i>D. discoideum</i>	GG..K.....	AT..VL.R.LR.....	KY..F.EN-.IK.VLR.....	AET.....Q.AN.Q.....	HS.L.K--..K				
<i>C. elegans</i>	V..T...N.R.P.....	KY..I.EN-.I.....	QQPS..G.....Q.AN.....	R.HD.L.VDTATAAA*					
<i>C. briggsae</i>	V..T...N.R.P.....	KY.....I.....	QQPS..G.....Q.AN.....	R.HD.L.VDTATAAA*					
<i>M. sexta</i>	N.ST.....	K...R.....	EN-.IH...R..TGEQ.....	VME.....HS.M.F--.DDDK					
<i>D. melanogaster</i>	N.ST.....	R.....KY.....	EN-.IH...R..GEN.....	A.E.....N..F-F-S.PEEK					
<i>H. sapiens</i>	S.T...N..R.....	KY.....IS...R..SD.....	S.....C...CF-N.PEDK						

Fig.11.5. Sequências das proteínas de extensão dos genes de fusão encontrados em vários organismos. Apenas os resíduos divergentes são apresentados; os pontos representam resíduos idênticos. Os travessões (-) representam deleções. (A) proteínas de extensão de 52-53 aminoácidos. As cinco primeiras sequências foram comparadas com a sequência de *T.brucei*; as restantes foram comparadas com a de *N.tabacum*. (B) proteínas de extensão de 76-81 aminoácidos.

4.2. Genes relacionados aos genes ubiquitina.

Em vários organismos, desde nemátodos ao homem, mas também em vírus, têm sido encontrados vários genes codificando para proteínas relacionadas com a ubiquitina (Tabela I1.2). A função destes genes, ou qual o papel do homólogo ubiquitina, é na generalidade desconhecida. Entre todas as sequências conhecidas (Fig.I1.6), apenas dois aminoácidos são conservados com a ubiquitina (Lys27 e Asp52), embora um grupo de seis apenas apresentem substituições conservadas (Ile3, Val5, Ile23, Ile30, Ile36 e Leu43). De acordo com a sua estrutura, estes genes podem ser classificados em três classes: uma classe A, com uma organização equivalente à classe I dos genes ubiquitina; uma classe B, equivalente à classe II; e uma classe C, que conterà genes codificantes para proteínas contendo uma região interna com homologia à ubiquitina.

A classe A compreende quatro genes estudados, dois de vírus da classe Baculovirus, um de ratinho, Nedd8, e um humano, UCRP (Tabela I1.2). Os três primeiros codificam para uma unidade equivalente à ubiquitina com uma pequena extensão C-terminal de 1 a 5 resíduos que são ou podem ser hidrolisados. Pouco se sabe sobre as suas funções, embora a conservação da região C-terminal e da Lys48 nestes genes permita antever uma possível actuação via conjugação a outras proteínas. Os genes virais são essenciais, uma vez que a sua remoção parece impossível, e são expressos na fase tardia de infecção do vírus (Guarino, 1990). Já o gene Nedd8 parece ser quase tão universal como a ubiquitina, uma vez que é possível detectar por hibridação em Southern a sua presença em vários mamíferos, galinha e levedura, pelo que poderá possuir alguma função essencial (Kumar *et al.*, 1993).

O gene codificante para a UCRP é o que se encontra mais bem estudado. A UCRP é uma proteína de 17 kDa com dois domínios ubiquitina separados por uma Pro, que é processada a 15 kDa por perda de uma extensão C-terminal de 8 resíduos imediatos à sequência ArgGlyGly (Knight *et al.*, 1988). Este gene é especificamente induzido pelo interferão (Haas *et al.*, 1987; Reich *et al.*, 1987), não sendo expresso na sua ausência. A sua expressão está intimamente correlacionada com a resposta antiviral (Haas *et al.*, 1987), actuando através da sua conjugação a uma "pool" de proteínas intracelulares de espectro mais reduzido do que o obtido com a conjugação à ubiquitina (Loeb & Haas, 1992). Esta "pool" é preexistente em células não incubadas na presença de interferão, mas é induzida na sua presença. Contudo, a indução de UCRP não implica necessariamente um aumento dos seus conjugados, o que parece implicar a necessidade de outros factores para uma conjugação eficiente (Loeb & Haas, 1992). Não é claro se a UCRP compete com a ubiquitina no uso do sistema de conjugação ou se usa um sistema próprio, mas estes dados parecem indiciar a existência de modificações pós-traducionais que representem um mecanismo novo de regulação intracelular.

A classe B compreende vários genes representados por quatro grupos: um em *Caenorhabditis*, UbL, um em humanos, GdX, um outro encontrado em vírus e vários mamíferos, *fau*, e um último em levedura e humanos, RAD23 (Tabela I1.2). Todos estes genes são expressos constitutivamente, ditos "housekeeping", e, à excepção do GdX, são constituídos por uma fusão entre uma unidade homóloga à ubiquitina e proteínas ribossomais. Os genes de nemátodos codificam para proteínas ribossomais homólogas à Ub79 da Classe II dos genes ubiquitina. Na verdade, Jones & Candido (1993) indicam que esta é a única forma existente para estas proteína ribossomais em *C. elegans* e *C. briggsae*, não existindo genes Ub79 canónicos. Estes dados relançam a discussão sobre o papel da ubiquitina nos genes de fusão. De facto, caso a ubiquitina funcione como "chaperone" na

<i>H. sapiens</i>	1	10	20	30	40	50	60	70
	MQIFVKTLTKGTTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG							
AcMNPV	..I.....	..A.T...EE.VADL.Q..A....V.V.....	SK.MA.....					
OpMNPV		..V.T...G..VGQ..Q..A....V.V.....	AQ..A.....					
Nedd-8	..L.K.....	..E.EIDI...T.KV.RI.ERVEE.....Q.....	..YS...MN.EK.AA..K.LGG.V.....					
UCPR-A	*DLT..M.A.NEFQVSLSS.MSVSEL..Q.TQ.I.VHAF.....	*PS.VA.Q.RVP.ASGGLPG..VL..VDKCDEP						
UCPR-B	LS.L.RNNK.RSS.Y..RLTQ.VAHL.QQVSGL..VQD.LFW.T.E..P...QLP.GE.GLKPL..VFMN.....							
V-Fau	..L..RAQELH.--..TGQG.VAQI.DHVTSL..A..D.VLL..SPQ..EA..GQCGVEALT..EVAG.ML..							
M-Fau	..L..RAQELH.--..TGQE.VAQI.DHVASL..A.ED.VLL..SP...EA..GQCGVEALT..EVAG.ML..							
R-Fau	..L..RAQELH.--..TGQE.VAQI..HVASL..A.ED.VLL..SP...EA..GQCGVEALT..EVAG.ML..							
H-Fau	..L..RAQELH.--F..TGQE.VAQI..HVASL..A.ED.VLL..AP...EA..GQCGVEALT..EVAG.ML..							
Ubl (<i>C.briggsae</i>)	..V....NR-.LY...AANEDVLSI.Q..EAA....SAE...V..R...SDCGL.A----	A.IYVN.E.L..						
Ubl (<i>C.elegans</i>)	..V....HR-.LF...AANEDVLSI.Q..EAA....AEE...CY..R...SDCGL.S----	A.IYVN.E.L..						
RAD23 (<i>S. cerevisiae</i>)	VSLTF.NFKKEKVP.DL...N..LET.T.LAQISCEES.IK..YS..V.Q.SK.V.ECGLKGDQVVFVMSQKKS							
HHRAD23A (<i>H.sapiens</i>)	VT.TL...QQQ.FKIRM..DE.VKVL.E..EAEK.*.VAG.K.Y..I.S.DVPIR..R.DEKNFVVVMTTKTKA							
HHRAD23B (<i>H.sapiens</i>)	..VTL...QQQ.FKIDID.EE.VKAL.E..ESEK.*.VAG.K.Y..I.N.DTA.KE.K.DEKNFVVVMTKPKA							
GdX	..LT..A.Q.RECS.QVPEDELVSTL.QLVSE.LNV.VR....L.K..A.A..KR....S.GPN.K.N..VKPLEK							
Anla (<i>Xenopus</i>)	..EL.IE....TCFE.R.S.YE.VTS..S...RL....VA..H..RNNME...ECS..G...SEGC..KM..AM...							
Anlb (<i>Xenopus</i>)	..EL.IE....TCFE.R.S.YE.VTS..S...RL....VA..H..WNNME...ECS....SEGC..KM..AM...							
BAT3	LEVL....DSQ..R.FI.GAQMNVKEF.EH.RASVS..SEK....YQ.RV.Q.DKK.QE..VG-GKVI...E.APPQ							

Fig.11.6. Sequências das regiões homólogas à ubiquitina encontradas em diversas proteínas. As diversas regiões homólogas foram comparadas com a sequência de ubiquitina humana; apenas as diferenças foram assinaladas; os pontos representam resíduos idênticos. Os travessões (-) representam deleções, e os asteriscos (*) inserções.

montagem dos ribossomas, como proposto por Finley e colaboradores (1989), a região de 70 resíduos de aminoácidos homóloga à ubiquitina que se encontra nestes genes manteria esta função, apesar da sua extrema divergência. É curioso notar que, tal como no gene poliubiquitina de *C. elegans*, estes genes possuem um intrão na Gly45. Contudo, a existência de outra proteína ribossomal, S30 de rato, que possui também uma região N-terminal homóloga à ubiquitina (genes *fau*; Kas *et al.*, 1992) pode apoiar as teorias de que a ubiquitina funcione apenas como uma maneira eficiente de se possuir proteínas ribossomais sem Met inicial (Warner, 1989), ou como indutora/estabilizadora da sua expressão (Finley *et al.*, 1989). Um possível esclarecimento poderá ser dado pela presença de um gene *fau* no vírus FBR-MuSV, que é transcrito como um mensageiro antisenso dos genes equivalentes das células de mamífero (Michiels *et al.*, 1993). A presença deste mRNA antisenso duplica a eficiência de transformação do vírus, mas permanece por esclarecer se se deve à inibição da proteína ribossomal do hospedeiro ou da sua metade ubiquitina (Michiels *et al.*, 1993). Por seu lado, o gene humano GdX codifica para uma proteína não processada de 157 resíduos expressa constitutivamente. A sua presença é detectada em outros organismos, como répteis, insectos e fungos, o que lhe parece conferir uma função essencial (Toniolo *et al.*, 1988). Não é impossível que, tal como os outros genes desta classe, codifique também para uma proteína ribossomal, já que pelo menos em *Dictyostelium* é possível detectar outras proteínas ribossomais que reagem a anticorpos antiubiquitina (Müller-Taubenberger *et al.*, 1989). Finalmente, as RAD23 são proteínas que parecem estar envolvidas na reparação do DNA danificado por UV (# 3.1).

Da classe C são conhecidos três genes: dois em *Xenopus*, An1a e An1b, e um em humanos, BAT3 (Tabela I1.2). Este último, descoberto por se localizar no "loci" dos genes do complexo de histocompatibilidade maior, codifica para uma proteína de 120 kDa que contém, entre os resíduos 18 e 93, uma região de 75 resíduos de aminoácidos semelhante à ubiquitina. Os dois genes de *Xenopus* correspondem a dois cDNAs de mensageiros maternos característicos do hemisfério animal de ovos não fertilizados. Codificam para proteínas de 77-78 kDa com uma região homóloga à ubiquitina entre os resíduos 28 e 103. Ambas as regiões possuem a sequência C-terminal ArgGlyGlyXaa, que é reconhecida por um número de proteases celulares (Lopez-Otin *et al.*, 1989), entre as quais a hidrolase da ubiquitina. Contudo, não há processamento destas proteínas por ao aminoácido Xaa corresponder uma Pro, que se sabe inibir a hidrólise da junção C-terminal (Bachmair *et al.*, 1986). Estas proteínas possuem também uma região C-terminal com um motivo "zinc finger".

4.3. Organização genómica e expressão dos genes ubiquitina.

Apesar de estudados em cerca de meia centena de organismos, a organização genómica dos genes ubiquitina é na generalidade muito pouco conhecida. Aparentemente, todos os organismos possuem pelo menos um gene poliubiquitina, que é o principal responsável pela manutenção da informação genética da ubiquitina. Já a presença de ubiquitinas de fusão não parece ser universal, pelo menos na sua forma canónica. Resta por esclarecer se a correspondente proteína ribossomal se encontra igualmente ausente, presente mas não fundida a uma ubiquitina, ou se, como em *Caernohabditis*, se encontra fundida a uma proteína relacionada com a ubiquitina (Jones & Candido, 1993). O número total de genes ou a sua disposição no genoma é variável, mas grupos próximos parecem possuir organizações semelhantes. Os genes ubiquitina são expressos constitutivamente,

independentemente de serem poliubiquitinas ou genes de fusão; contudo, os primeiros são preferencialmente regulados por agentes de "stress", o que não sucede com os segundos, regulados preferencialmente pelo desenvolvimento celular. Uma descrição mais detalhada da organização e expressão dos genes ubiquitina por grupos taxonómicos é indicada a seguir.

Protozoários - O reino *Protozoa* é constituído por organismos muito variados cuja relação filogenética é por vezes muito distante. Esta variação reflecte-se no número de genes ubiquitina encontrados em cada espécie, desde um único gene em *G.lambliia* até cerca de 10 em *T.brucei* (Tabela I1.1). O número de unidades por gene é igualmente muito variável, tendo sido inclusive encontrado um único gene com uma única unidade em *G.lambliia*. Para além dos genes poliubiquitina, os protozoários parecem possuir sempre pelo menos um gene de fusão Ub52, mas nenhum gene de fusão Ub79 foi ainda encontrado. Aliás, nas duas espécies que possuem uma organização dos seus genes ubiquitina bem caracterizada, *Trypanosoma cruzi* e *Leishmania tarentolae*, este gene parece estar mesmo ausente. Em *T.cruzi* existem 10 genes ubiquitina, todos fundidos à proteína ribossomal de 52 aminoácidos, apesar de metade serem poliubiquitinas (Swindle *et al.*, 1988). Curiosamente, esta organização não parece ser comum em *Trypanosoma*, porque na espécie aparentada *T.brucei*, para além de aparentemente existir um menor número de genes ubiquitina, o gene poliubiquitina termina com um aminoácido extra e não com uma proteína de fusão (Wong & Campbell, 1989; Wong *et al.*, 1992). Todavia, ambos apresentam características comuns, como um gene poliubiquitina com um elevado número de monómeros, e um *loci* contendo genes ubiquitina, calmodulina e EF-H5 (Wong *et al.*, 1993; Chung & Swindle, 1990; Ajioka & Swindle, 1993). Em *Trypanosoma cruzi*, todos os genes ubiquitina estão organizados em série numa única região do genoma (Swindle *et al.*, 1988). Este tipo de organização em série parece ser também seguido no ciliado *Tetrahymena pyriformis* (Neves *et al.*, 1988; 1990), mas não em *L.tarentolae*. Neste organismo, foram encontrados três genes, um poliubiquitina e dois de fusão Ub52, que se encontram localizados em cromossomas distintos. À semelhança do observado em *Trypanosoma*, os alelos do gene poliubiquitina possuem um número distinto de unidades, e um dos genes de fusão encontra-se igualmente associado a genes de calmodulina e EF-H5 (Fleischmann & Campbell, 1994).

Em protozoários, o gene de fusão Ub52 é expresso constitutivamente e, pelo menos em *Trypanosoma*, *Eimeria* e *Acanthamoeba*, expresso diferencialmente nas diversas fases do ciclo de vida destes organismos (Abrahamsen *et al.*, 1993; Kirchhoff *et al.*, 1988; Swindle *et al.*, 1988; Ahn & Henney, 1994). A acção de agentes de "stress" aparentemente não afecta a expressão deste gene em *Trypanosoma* e *Tetrahymena* (Swindle *et al.*, 1988; Neves *et al.*, 1991a), mas induz especificamente a expressão de genes poliubiquitina em *Tetrahymena*, embora não em *Giardia* ou *Trypanosoma*, onde apenas se observa uma indução durante a fase de crescimento estacionário (Neves *et al.*, 1988; 1991b; Swindle *et al.*, 1988; Krebber *et al.*, 1994).

Fungos - À excepção do fungo primitivo *Dictyostelium*, os fungos parecem possuir um número reduzido de genes ubiquitina (Tabela I1.1). Todavia, só em Ascomicetos há um conhecimento aprofundado da organização genómica dos genes ubiquitina, que é distinta para os dois organismos onde foi estudada, *S.cerevisiae* e *N.crassa* (Özkaynak *et al.*, 1987; Taccioli *et al.*, 1989; 1991). Ambos possuem um único gene poliubiquitina com um reduzido número de unidades e um gene de fusão Ub79, mas enquanto em *S.cerevisiae* se encontram dois genes de fusão Ub52, estes estão ausentes em *N.crassa*. Todavia, sendo a proteína ribossomal codificada por estes genes essencial em levedura (Finley *et al.*, 1989), é

possível que um gene codificando para esta proteína venha também a ser encontrado em *N.crassa*, eventualmente não fundido a uma ubiquitina canónica. Em levedura, existem apenas quatro genes ubiquitina, mas existem dados que apontam para a existência de genes relacionados aos genes ubiquitina (Kumar *et al.*, 1993).

A expressão dos genes ubiquitina encontra-se bem estudada em fungos. Os genes de fusão são expressos constitutivamente e, em *Dictyostelium* e *Neurospora*, regulados diferencialmente durante os diversos ciclos de vida celulares, sendo preferencialmente expressos durante as fases de crescimento e proliferação celular (Westphal *et al.*, 1986; Taccioli *et al.*, 1991). Em *S.cerevisiae*, estes genes são os principais fornecedores de ubiquitina durante o seu crescimento vegetativo (Özkaynak *et al.*, 1987). Neste organismo, como em *Dictyostelium* e *P.infestans*, os genes poliubiquitina são fortemente induzidos em várias situações de "stress" celular (Özkaynak *et al.*, 1987; Müller-Taubenberger *et al.*, 1988a; Pieterse *et al.*, 1991). Todavia, tal já não ocorre em *N.crassa*, onde as diversas situações de "stress" celular testadas não alteram a expressão do seu único gene poliubiquitina; este é apenas induzido na presença do inibidor da síntese proteica cicloeximida (Taccioli *et al.*, 1989).

Algas - Apesar de serem conhecidos genes ubiquitina em cerca de meia dúzia de algas, apenas em duas, *C. reinhardtii* e *V.carteri*, foi efectuado um estudo mais aprofundado da sua organização. Ambas possuem aparentemente um número elevado de genes ubiquitina, contendo pelo menos um gene de fusão. Em *Volvox* é possível observar-se polimorfismo para os genes ubiquitina, sendo dois dos genes co-segregados durante cruzamentos de estirpes polimórficas (Schiedlmeier & Schmitt, 1994).

A acção de agentes de "stress" regula diferentemente os genes ubiquitina em algas, reduzindo a expressão dos genes de fusão mas aumentando a dos genes poliubiquitina (Pollmann *et al.*, 1991; Schiedlmeier & Schmitt, 1994). Em *Volvox*, os diferentes genes ubiquitina são igualmente regulados durante o ciclo de vida da alga (Schiedlmeier & Schmitt, 1994).

Plantas liliopsidas (monocotiledóneas) - As plantas monocotiledóneas parecem caracterizar-se por possuírem um número elevado de genes ubiquitina e por um número de monómeros bastante variável nos seus genes poliubiquitina (Tabela II.1). Os genes parecem encontrar-se dispersos, mas podem encontrar-se algumas associações pontuais, causadas por duplicações recentes, como em dois genes de fusão Ub79 de *H.vulgare* (Gausling & Jensen, 1990). Outra característica comum parece ser a predominância de Gln como codão extra dos genes poliubiquitina.

A expressão dos genes ubiquitina em monocotiledóneas é bastante homogénea, pelo menos nas espécies estudadas. Os genes de fusão são expressos constitutivamente nos diversos tecidos, mas os níveis de transcritos estão coordenados com as taxas de divisão e síntese proteica de cada tecido (Gausling & Barkardottir, 1986; Nishi *et al.*, 1993; Chen & Rubenstein, 1991). Os genes poliubiquitina são transcritos constitutivamente, mas pelo menos um é induzido em situações de "stress" celular ou na presença de hormonas (Gausling & Barkardottir, 1986; Reynolds & Hooley, 1992; Chen & Rubenstein, 1991; Christensen & Quail, 1989). Um "stress" hipertérmico severo pode também afectar o processamento dos intrões dos pré-mRNAs poliubiquitina, mesmo dos induzidos por "stress" (Christensen *et al.*, 1992).

Plantas magnoliopsidas (dicotiledóneas) - As plantas dicotiledóneas, tal como as monocotiledóneas, possuem um número elevado de genes ubiquitina, mas diferem destas por possuírem um número mais homogéneo de monómeros por gene poliubiquitina, ao redor de seis (Tabela II.1). Numa planta, *Arabidopsis thaliana*, a organização encontra-se

razoavelmente estabelecida: este organismo possui cinco genes de fusão (2 Ub52 e 3 Ub79) e possivelmente outras tantas poliubiquitinas, dispersas no genoma.

A expressão dos genes ubiquitina é menos homogênea em dicotiledóneas do que em monocotiledóneas. Os genes de fusão são regulados de forma equivalente à observada em monocotiledóneas, mas nas polemoniales observa-se também uma indução destes genes na presença de hormonas (etileno) e algumas situações de "stress" (injúria), mas uma repressão noutras situações de "stress" ("stress" térmico) (Garbarino *et al.*, 1992; Hoffman *et al.*, 1991). Já a expressão dos genes poliubiquitina é distinta: enquanto em polemoniales e rubiales pelo menos um gene é regulado em situações de "stress" (Garbarino *et al.*, 1992; Genschik *et al.*, 1992; Jamet *et al.*, 1990; Binet *et al.*, 1991), em umbelales ou caparales não se observam alterações da expressão nas mesmas condições (Kawalleck *et al.*, 1993; Burke *et al.*, 1988).

Poríferos - Pouco se sabe da organização neste *philum*, onde apenas numa espécie foram estudados os genes ubiquitina (Tabela I1.1). Esta espécie, *G.cydonium*, apenas contém um mRNA codificando para uma poliubiquitina, que é induzido durante a agregação celular deste organismo. Curiosamente, não foi detectado nenhum sinal de presença de genes de fusão (Pfeifer *et al.*, 1993).

Plantelmintes - Um único organismo estudado, *E.multilocularis*, onde foi sequenciado um cDNA que, curiosamente, codifica para uma ubiquitina com uma única unidade sem codão extra.

Ascelmintes - Ambos os nemátodos *Caenorhabditis* estudados possuem apenas dois genes ubiquitina, um correspondente a uma poliubiquitina com onze unidades e o outro a uma ubiquitina de fusão Ub52 (Tabela I1.1). O gene correspondente à proteína ribossomal de 79 aminoácidos apenas existe nestes organismos fundido a uma proteína relacionada à ubiquitina (Jones & Candido, 1993). Ambos os genes são expressos constitutivamente, não se observando nenhum tipo de regulação quer durante o desenvolvimento dos indivíduos quer sob situações de "stress" (Graham *et al.*, 1989).

Equinodermes - Também aqui apenas se conhece uma espécie estudada, *S.purpuratus*, cujo aparentemente único mRNA poliubiquitina é induzido em situações de "stress", mas não é regulado durante o desenvolvimento embriogénico. Parece igualmente não possuir genes de fusão (Nemer *et al.*, 1991).

Artrópodes - Os artrópodes parecem ter um número reduzido de genes ubiquitina, com pelo menos um gene poliubiquitina com um elevado número de monómeros (Tabela I1.1). Pelo menos em *Drosophila* existe também um elevado polimorfismo nos genes poliubiquitina (Arribas *et al.*, 1986).

A expressão dos genes de fusão é constitutiva durante o desenvolvimento dos insectos, embora os seus níveis possam variar consoante o tecido celular (Bishoff & Schwartz, 1990; Cabrera *et al.*, 1992; Barrio *et al.*, 1994). Os genes poliubiquitina são induzidos por situações de "stress" (Niedzwiecki & Fleming, 1993) e, em *M.sexta*, abundantemente expressos em tecidos que sofrem uma morte programada durante o desenvolvimento da pupa a indivíduo adulto (Schwartz *et al.*, 1990).

Cordados - À excepção do batráquio *Xenopus laevis*, os vertebrados parecem possuir apenas quatro genes funcionais, embora a família ubiquitina possa ainda conter diversos pseudogenes (Tabela I1.1). A organização no homem, com dois genes de fusão (um Ub52 e um Ub79) e dois genes poliubiquitina, um com um baixo (3-4) e outro com um elevado número de unidades (8-10) parece ser a mais comum, pelo menos em mamíferos. Todavia, pode-se observar um elevado polimorfismo no número de unidades (Baker & Board, 1989). Apesar disso, é possível encontrar uma elevada homologia para cada um dos dois genes

poliubiquitina ao longo da árvore filogenética dos vertebrados, que se estende às regiões não traduzidas (Nenoi *et al.*, 1992). É de notar também que em roedores, a última unidade do gene de elevado número de unidades sofreu uma deleção parcial, produzindo uma proteína híbrida provavelmente não funcional. Contudo, o gene é expresso e as unidades restantes codificam para ubiquitinas funcionais (Nenoi *et al.*, 1994).

A expressão dos genes ubiquitina não é muito diferente da observada na generalidade dos organismos. Ambos as classes de genes são expressas constitutivamente, embora os níveis de expressão possam variar bastante consoante o tecido celular (Mezquita *et al.*, 1993; Wiborg *et al.*, 1985). Os genes poliubiquitina são induzidos por situações de "stress" (Bond & Schlesinger, 1986; Nenoi *et al.*, 1992; Nenoi, 1992), e em galinha, os dois genes poliubiquitina são expressos diferencialmente durante a embriogénese (Mezquita *et al.*, 1993). Tal como em plantas, o "stress" térmico severo afecta o processamento dos intrões nos genes poliubiquitina (Bond & Schlesinger, 1986).

Capítulo I

INTRODUÇÃO 2

O ciliado *Tetrahymena*

1. Caracterização do ciliado *Tetrahymena*.

1.1. Introdução ao phylum Ciliophora.

O *phylum* Ciliophora é um dos mais antigos ramos eucariotas, tendo sido provavelmente um dos primeiros a emergir do tronco comum (Baba *et al.*, 1981). As cerca de 7000 espécies actualmente conhecidas são constituídas por células unicelulares possuindo um elevado grau de diferenciação e organização. Apesar da sua diversidade, apresentam várias características comuns que os permitiram ser considerados desde há longo tempo um grupo monofilético, como sejam o dimorfismo nuclear, um processo de reprodução sexual envolvendo conjugação e um córtex que consiste numa estrutura pelicular e subpelicular complexa que compreende os cílios. Estes encontram-se muitas vezes organizados em linhas longitudinais anteroposteriores, com os corpos basais associados a uma rede típica de fibras citoesqueléticas.

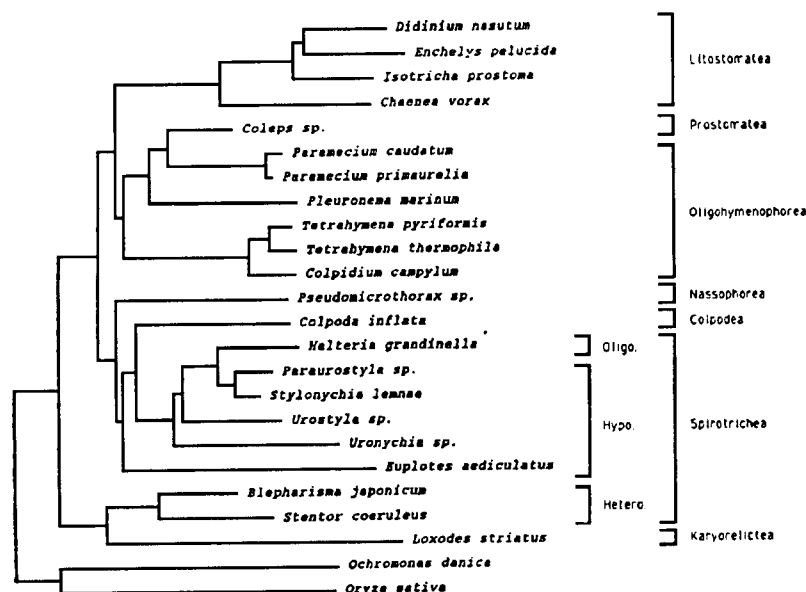


Fig.I2.1. Árvore filogenética dos ciliados gerada pelo método de "neighbor-joining". A análise foi efectuada por alinhamento das sequências nucleotídicas dos rRNAs 28S. À direita, os grupos foram designados de acordo com a sistemática de Lynn & Corliss (1991). As ordens Heterotrichia, Hypotrichia e Oligotrichia são igualmente especificadas (Corliss, 1979). As distâncias horizontais são proporcionais à distância evolutiva. Adaptado de Baroin-Tourancheau e colaboradores (1992).

Dentro do *phylum*, a diversificação é manifestada pelo padrão de implantação dos cílios na superfície celular e na região especializada pela ingestão de alimentos, o aparelho oral. Esta disposição ou, mais recentemente, a análise comparativa da organização ultraestrutural dos corpos basais, é a base da classificação taxonómica do grupo, para a qual o mais recente tratado reconhece oito classes (Lynn & Corliss, 1991). A Fig.I2.1 representa uma árvore filogenética com representantes de sete das oito classes, calculada com base nas sequências dos rRNAs 28S, que permite analisar o seu interrelacionamento (Baroin-Tourancheau *et al.*, 1992). Aparentemente, o primeiro grupo a emergir nos ciliados é constituído pelos Cariolícteos e, curiosamente, pelos Heterotricos, habitualmente classificados dentro dos Spirotricos. Contudo, ambos formam um grupo monofilético bem caracterizado, o que pode indicar uma classificação errónea dos Heterotricos pela sistemática clássica. Nas outras classes, algumas são bem suportadas como grupos monofiléticos por este estudo, como a classe Litostomatea, enquanto outras aparecem associadas a outras classes mais representativas, como a classe Prostomatea à classe Oligohymenophora. Da filogenia representada na Fig.I2.1, deve-se salientar ainda dois

factos importantes: o primeiro prende-se com a posição ambígua do género *Euplotes*, um grupo de ciliados muito estudado em Biologia Molecular. Habitualmente classificado dentro dos Hipotricos, apresenta apesar disso uma grande divergência relativamente a outras espécies da classe (exemplificada, p. ex., pelo diferente uso de codões; # 3.1) e diversas análises sustentam a possibilidade de uma emergência independente na base da tricotomia oligomenofera-litostomatea-hipotrica (Baroin-Tourancheau *et al.*, 1992; Lynn & Sogin, 1988). O segundo facto prende-se com as distâncias filogenéticas observadas. Apesar da sua aparente proximidade, a divergência molecular observada entre, por exemplo, *Euplotes* e *Paramecium* ou *Tetrahymena* é superior à observada entre o arroz e o anfíbio *Xenopus*, enquanto a divergência entre a *Paramecium* e a *Tetrahymena* é superior à observada entre o arroz e o limoeiro (Baroin-Tourancheau *et al.*, 1992).

1.2. Características e distribuição do género *Tetrahymena*.

O género *Tetrahymena* Furgason 1940 é o mais conhecido das Tetrahymenidae e engloba actualmente cerca de três dezenas de espécies, tendo como espécie-tipo a *T. pyriformis* (Ehrenberg 1830) Lwoff 1947 (Fig. I2.2). Esta espécie possui forma de pêra (dando origem ao nome) de dimensões aproximadas $60\mu \times 30\mu$ e recoberta por 15-25 fiadas meridionais de cílios. No seu pólo anterior localiza-se o aparelho oral, constituído por cerca de 160-170 corpos basais associados a cílios e a fibras intracelulares, que se agrupam em quatro estruturas ciliares compostas ("membranas ondulatórias") que originaram o nome do género, *Tetrahymena*. A ingestão de alimentos é permitida pela formação de vacúolos alimentares no aparelho oral, e a excreção é efectuada pelo citoprocto, localizado no pólo posterior, através de vacúolos contrácteis posicionados na região médio ventral. As *Tetrahymena* encontram-se distribuídas por quase todo o globo terrestre, habitando praticamente todos os tipos de águas. São organismos habitualmente bacteriófagos, embora possam ser igualmente carnívoros e, em alguns casos, parasitas facultativos de hospedeiros invertebrados. Em laboratório, a maioria cresce facilmente em meio axénico, ou em meios que forneçam todas as vitaminas, aminoácidos e sais indispensáveis.

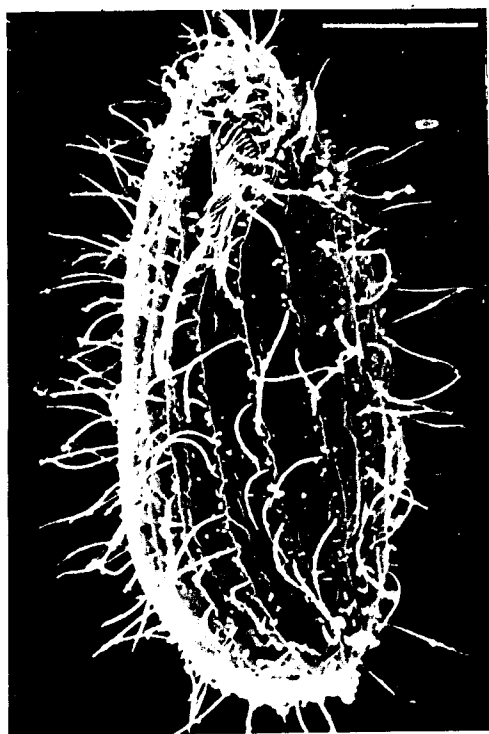


Fig. I2.2. O ciliado *Tetrahymena pyriformis*.

A classificação e taxonomia das diversas espécies do género *Tetrahymena* tem encontrado diversas dificuldades ao longo dos tempos, devido à extrema semelhança morfológica que observam entre si. Assim, há cerca de vinte anos eram consideradas cerca de 10 a 14 espécies dentro do género (Corliss, 1973), baseadas apenas em alguns caracteres morfológicos, como p.ex. largos aparelhos orais (que permitem a alimentação carnívora),

ou ecológicos. Contudo, não eram considerados caracteres como a capacidade de reprodução sexuada das células (# 2.2.2), o que conduzia p.ex. a que espécie *T. pyriformis* incluísse diversos organismos amiconucleados (e por isso incapazes de se cruzarem) e doze "syngen", i.e., conjuntos de organismos capazes de se cruzarem entre si mas não com organismos de outros conjuntos. Esta "espécie" (ou complexo) só foi resolvida em 1976, quando Nanney & MacCoy apresentaram razões morfológicas, ecológicas e bioquímicas que permitiram caracterizar como espécies independentes cada um dos "syngen" do complexo. Assim, o "syngen" 1 veio a ser a *T. thermophila* devido à sua tolerância a temperaturas mais elevadas, o "syngen" 2 veio a ser a *T. americanis* por só se encontrar no continente americano, etc. Por seu lado, as estirpes amiconucleadas foram caracterizadas sobretudo por grupos fenotípicos de isoenzimas, e à *T. pyriformis* juntaram-se a *T. ellioti*, *T. furgasoni*, *T. lwoffii* e, mais recentemente, *T. mimbres*.

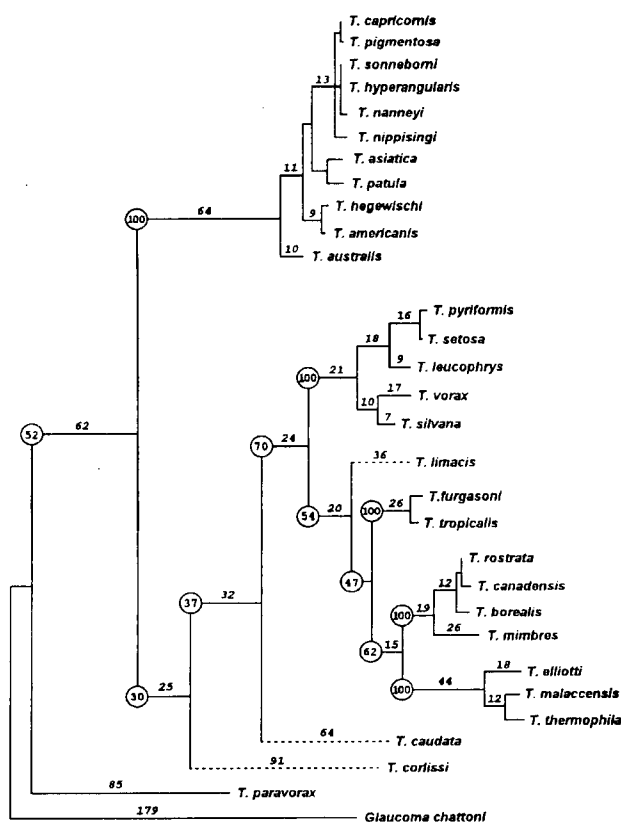


Fig.I2.3. Árvore filogenética das *Tetrahymena* gerada pelo método de "maximum parsimony". A análise foi efectuada por alinhamento das regiões intergénicas entre os genes das histonas H3II e H4II. A posição das espécies com linhas em pontado é ambígua. Os números em itálico correspondem às substituições de nucleótidos, e os números dentro de círculos são valores de confiança (em percentagem) dos grupos descendentes de cada nó. Adaptado de Sadler & Brunk (1992).

Apesar do aperfeiçoamento dos critérios de classificação, a catalogação de um organismo recém-recolhido é ainda um processo moroso. O simples critério da capacidade de cruzamento é muito falível, pois este nem sempre é fácil de ser conseguido em laboratório, e deve-se tomar em conta que as *Tetrahymena* são organismos que possuem períodos de imaturidade e senilidade no seu ciclo de vida (# 2.2.2.1). Além disso, o critério é inaplicável para espécies amiconucleadas, que representam cerca de 30 a 50% dos organismos encontrados na Natureza. Apenas com o desenvolvimento da Biologia Molecular moderna será de esperar um desenvolvimento da taxonomia, sobretudo com o advento das técnicas de sequenciação e de PCR. A Fig.I2.3 representa uma árvore filogenética onde estas técnicas foram aplicadas para determinação da sequência da região intergénica entre as histonas H3 e H4 (Sadler & Brunk, 1992), e que permite uma melhor compreensão das relações entre as diversas espécies de *Tetrahymena*. A principal característica é a presença de dois grupos principais, um muito homogéneo com onze

espécies muito próximas entre si, e o segundo mais heterogêneo. Outra característica importante advém da existência de uma distribuição em grupos distintos das três espécies amiconucleadas, *T.pyriformis*, *T.furgasoni* e *T.mimbres*, demonstrando que apesar da incapacidade de cruzamento são realmente espécies distintas. Finalmente, é de notar ainda que as espécies anteriormente classificadas como "syngen" do complexo *T.pyriformis* se encontram distribuídas por diversos grupos: p.ex., *T.americanis* está mais próxima do anteriormente considerado complexo *T.patula*, e *T.canadensis* está muito relacionada com a principal espécie do anterior complexo *T.rostrata*. Mais ainda, nenhuma das quatro espécies filogeneticamente mais próximas à própria *T.pyriformis* pertencia anteriormente ao seu próprio complexo.

2. O ciclo de vida em *Tetrahymena*.

2.1. O dimorfismo nuclear em ciliados.

Uma das características mais marcantes dos ciliados, para além dos cílios, é a existência de um dimorfismo nuclear. Com poucas exceções (p.ex., *T.pyriformis*, # 1.2), os ciliados possuem dois tipos de núcleos estruturalmente e funcionalmente diferentes, chamados, devido ao seu tamanho relativo, micronúcleos e macronúcleos. O número destes núcleos varia com a espécie: por exemplo, o género *Tetrahymena* possui um micronúcleo e um macronúcleo, enquanto *Oxytricha* possui quatro micronúcleos e dois macronúcleos.

Os micronúcleos são núcleos diplóides cuja principal função é a de transmitir a informação genética de uma geração sexual à seguinte, por um processo denominado conjugação (# 2.2.2), sendo transcricionalmente inactivos durante a vida vegetativa da célula. O DNA encontra-se fisicamente organizado em cromossomas, replicando-se por mitose de uma maneira equivalente à observada noutros organismos. Esta organização é contudo mais clara em *Tetrahymena*, cujos cinco cromossomas metacêntricos são facilmente visualizáveis durante a meiose (Ray, 1956), do que em outros organismos, como *Paramecium* ou *Stylonychia*, que possuem de 50 a mais de cem cromossomas micronucleares de difícil visualização (Sonneborn, 1974; Ammerman, 1971). A facilidade de isolamento dos micronúcleos em *Tetrahymena* (Gorovsky *et al.*, 1975) permitiu caracterizá-los melhor neste ciliado do que em outros. Apesar da sua aparente inactividade e da existência de estirpes amiconucleadas naturais, a eliminação do micronúcleo é normalmente letal para a célula. Contudo, é possível isolar estirpes nulissómicas viáveis, em que o micronúcleo não possui de um a três dos seus cinco cromossomas, desde que os cromossomas eliminados não incluam o braço esquerdo do cromossoma 1 (Bruns *et al.*, 1983). A incapacidade de isolar estirpes nulissómicas do cromossoma 1 sugere que este contém genes inactivos ou ausentes no macronúcleo. Efectivamente, apesar de ser possível isolar algumas estirpes amiconucleadas (Kaney & Speare, 1983), estas parecem ter sofrido rearranjos prévios que permitiram a inclusão de sequências específicas do micronúcleo, possivelmente essenciais, no macronúcleo (Karrer *et al.*, 1984). A cromatina micronuclear encontra-se organizada em nucleossomas típicos, em que, curiosamente, a unidade de repetição obtida após digestão com nuclease micrococcal é inferior à do macronúcleo (Gorovsky *et al.*, 1977). As histonas constituintes do nucleossoma são do mesmo tipo das encontradas no micronúcleo, à excepção da histona de ligação H1: esta é substituída por três péptidos aparentemente processados a partir de uma poliproteína única que desempenharão um papel equivalente (Allis *et al.*, 1984; Wu *et al.*, 1988). O DNA micronuclear é, tal como o macronuclear, extremamente rico em AT (25% GC; Gorovsky,

1980), e estudos de renaturação genética sugerem que contém poucas sequências muito repetidas, e cerca de 10-20% de sequências moderadamente repetidas (Yao & Gorovsky, 1974).

Os macronúcleos podem ser considerados os núcleos somáticos do organismo, uma vez que é neles que ocorre a maioria, senão a totalidade, da transcrição celular. São núcleos extremamente poliplóides, com $45n$ em *Tetrahymena* (Seyfert & Preparata, 1979), mas podendo atingir acima de $15000n$ em *Stylonychia* (Steinbruck *et al.*, 1981). Apesar do seu elevado conteúdo de DNA, a complexidade das sequências macronucleares é inferior à das micronucleares. De facto, o macronúcleo é gerado de novo em cada geração sexual, por amplificação de um micronúcleo num processo complexo que envolve rearranjos e eliminação de DNA micronuclear (# 2.3). Em *Tetrahymena*, 10 a 20% de sequências micronucleares são eliminadas (Yao & Gorovsky, 1974), mas em ciliados Hipotricos esse valor pode atingir os 95% (Lauth *et al.*, 1976). O DNA macronuclear não se encontra organizado em cromossomas, mas antes em fragmentos subcromossómicos que se dividem amitoticamente. Aparentemente, a fragmentação do DNA macronuclear é uma característica geral dos ciliados. Em *Tetrahymena*, existem cerca de 200 elementos subcromossómicos, de cerca de 21 a 1100 kb com um tamanho médio de 600 kb (Preer & Preer, 1979; Conover & Brunk, 1986), e em *Paramecium* vão de 50 a 800 kb, com 300 kb de tamanho médio (Preer & Preer, 1979; Caron & Meyer, 1989). Mas a fragmentação atinge os seus valores extremos em ciliados Hipotricos, onde os valores médios dos fragmentos vão de 2 a 3.5 kb. O pequeno tamanho do DNA macronuclear nesta ordem de ciliados levantou a hipótese de cada molécula representaria um gene individual e independente, o que tem sido apoiado quer pela sequenciação directa de fragmentos macronucleares, quer pela observação de que cerca de 90% do DNA macronuclear híbrida com o RNA celular (Nock, 1981; Klobutcher & Prescott, 1986). A cromatina macronuclear organiza-se igualmente em nucleossomas que, em *Tetrahymena*, contém para além das cinco histonas H1, H2A, H2B, H3 e H4, as duas variantes exclusivas Hv1 e Hv2 (Allis & Wiggins, 1984; Allis *et al.*, 1986). Ao contrário do DNA micronuclear, que não contém bases modificadas, cerca de 0.8% das adeninas do DNA macronuclear deste ciliado são N-6 metiladas, embora não seja detectável nenhuma metilação em citosinas, como observado em organismos superiores (Gorovsky, 1973).

2.2. O ciclo de vida em *Tetrahymena*.

2.2.1. O ciclo vegetativo.

Na sua fase vegetativa, os ciliados dividem-se por fissão originando duas células-filha praticamente idênticas à célula-mãe. A divisão nuclear ocorre antes da divisão celular, e os estudos efectuados demonstraram que é independente e distinta para cada um dos núcleos (Cameron, 1973). Em *Tetrahymena*, a divisão micronuclear ocorre antes da divisão macronuclear, através de uma mitose equivalente mas não idêntica à observada em eucariotas superiores: durante a mitose não há destruição da membrana nuclear, sendo o fuso mitótico intranuclear, e não se observa uma condensação típica da cromatina nos cinco cromossomas observados durante a meiose. A síntese de DNA micronuclear é praticamente imediata à divisão, pelo que as células praticamente não apresentam fase G_1 e passam a maior parte da sua vida vegetativa em fase G_2 micronuclear.

A divisão do macronúcleo é amitótica e inicia-se praticamente de seguida à divisão micronuclear. O macronúcleo sofre uma constrição e separa-se em dois sem quebra da membrana nuclear, e na ausência de uma condensação cromossômica ou formação do fuso mitótico. Cada uma das células-filha recebe uma quantidade diferente de DNA (Cleffman, 1980), e análises genéticas demonstram que cromátídeas irmãs não são igualmente herdadas pela progenia, podendo, eventualmente, perder-se um alelo por cada *locus* enquanto outro é fixado (Nanney & Preparata, 1979). Este "sortido fenotípico" deveria levar à perda periódica do próprio *locus*, o que todavia é contrariado pelos estudos efectuados com a viabilidade de linhas celulares, sugerindo a existência de um mecanismo regulatório. Uma das possibilidades sugeridas indica que o macronúcleo pode não controlar a segregação cromossômica, mas controlar o número de cópias cromossômicas (Doerder, 1979; Preer & Preer, 1979). Caso os cromossomas macronucleares se comportem como plasmídeo linear, cada qual possuindo o seu próprio elemento controlador do número de cópias, este não distinguiria entre diferentes alelos do mesmo *locus*, pelo que seria mantido um número estável de cópias para cada *locus*, mas não necessariamente para cada alelo (Bruns, 1986). Após a divisão, e contrariamente ao micronúcleo, o macronúcleo passa por uma fase G₁, ocorrendo a síntese de DNA a meio do crescimento vegetativo.

A divisão macronuclear não ocorre, contudo, em todos os ciliados. Nos Cariolíteos, cujo macronúcleo tem uma quantidade de DNA diplóide da G₂, não há divisão macronuclear, sendo gerado um novo macronúcleo a partir de um micronúcleo em cada ciclo celular. Nestes organismos, o macronúcleo preexistente nunca é destruído, nem durante a reprodução sexuada (Raikov, 1982).

2.2.2. O ciclo sexual ou conjugação.

2.2.2.1. A conjugação normal.

O ciclo sexual dos ciliados, ou conjugação, é constituído por uma série de acontecimentos particularmente fascinantes e únicos, sobre os quais apenas se possui um pálido conhecimento. Esses acontecimentos, que incluem fusões membranares, meioses e mitoses, gametogéneses, fertilizações, diferenciações e apoptoses nucleares, ocorrem de uma forma coordenada durante cerca de 20 horas após o encontro de duas células, resultando no final em quatro células-filhas (carionides) com um genótipo nuclear idêntico.

Para que duas células de *Tetrahymena* iniciem uma reprodução sexuada, precisam de reunir uma série de condições iniciais, nomeadamente serem sexualmente compatíveis (diferentes "matings"; # 2.2.2.3), sexualmente maduras (ver adiante) e encontrarem-se em condições ambientais favoráveis, i.e., num meio pobre em nutrientes, arejado e de baixa força iónica, mas sem agitação. Quando células nestas condições são colocadas em contacto, entram num período co-estimulatório, onde são induzidas algumas transformações morfológicas, como a transformação do extremo anterior da célula ("tip"; Wolfe & Grimes, 1979) e o aparecimento na zona anterior de receptores sensíveis à concanavalina A (Casci & Hufnagel, 1988). A co-estimulação em *Tetrahymena* requer o contacto físico entre as células, e não é induzida por eventuais moléculas excretadas para o meio (ferohormonas), ao contrário do que sucede em alguns ciliados, como *Euplotes* ou *Blepharisma*, onde este tipo de moléculas tem sido isolado (Luporini & Miceli, 1986). O processo do reconhecimento do "mating" em *Tetrahymena* não é conhecido, mas parece envolver, para

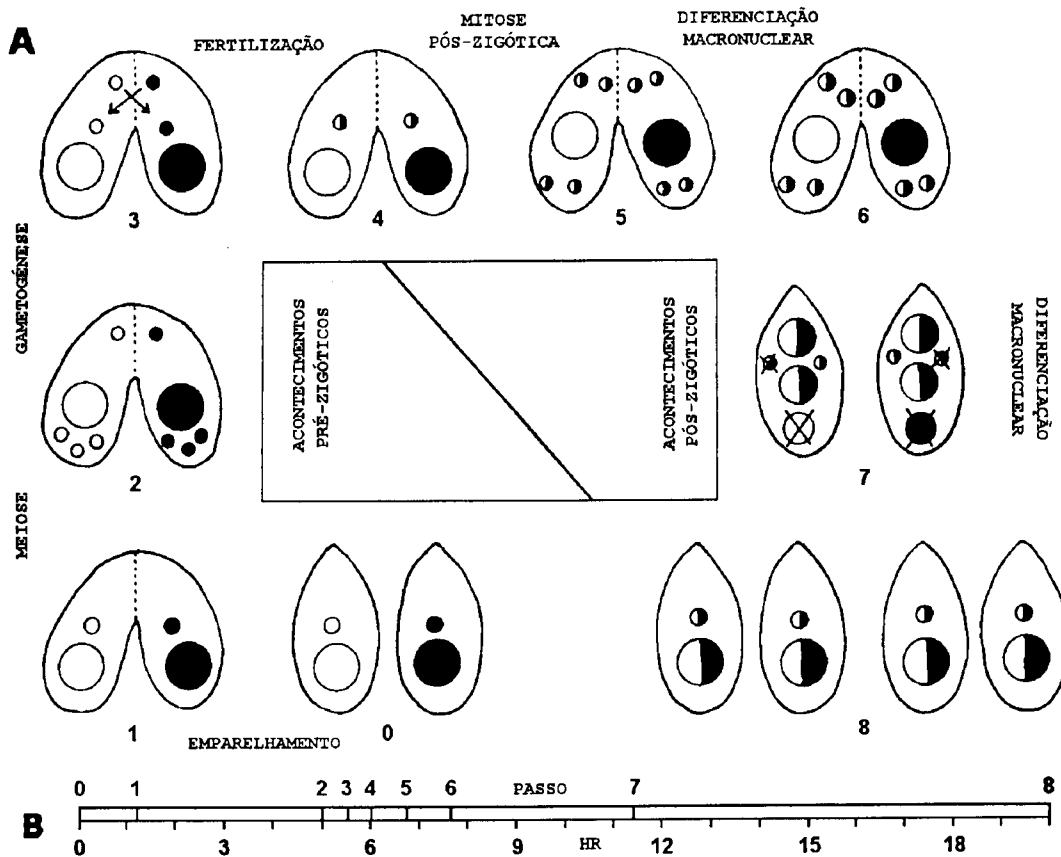


Fig.12.4. A conjugação em *Tetrahymena*. (A) Principais passos da conjugação. 0. Células vegetativas. 1. Células emparelhadas. 2. Meiose, com produção de quatro núcleos haplóides, dos quais apenas o anterior permanece funcional. 3. Mitose do núcleo haplóide, com troca de um dos produtos (núcleo migratório) entre as células. 4. Fusão do núcleo migratório com o núcleo estacionário formando um núcleo diplóide. 5. Dupla mitose originando quatro núcleos. 6. Início da maturação em macronúcleo dos dois micronúcleos anteriores. 7. Separação dos exconjugantes com destruição do macronúcleo original e de um dos micronúcleos. 8. Divisão celular (após realimentação), em que cada carionide recebe um dos macronúcleos e um produto mitótico do micronúcleo funcional. (B) Escala de tempo dos principais acontecimentos da conjugação em *T. thermophila*. Adaptado de Orias (1986).

além dos previamente referidos receptores sensíveis à concanavalina A, outras proteínas presentes nas membranas dos cílios (Wolfe *et al.*, 1993).

Ao fim de cerca de uma hora, as células formam pares fortemente unidos pelos "tips", onde as membranas e o citoesqueleto de cada célula se fundem para formar uma estrutura complexa conhecida por junção de conjugação (Wolfe, 1982). Esta estrutura é interrompida por diversos poros que permitem a troca de pequenas moléculas entre ambos os citoplasmas, mas não a de grandes organelos, como as mitocôndrias. Nesta altura, ocorrem uma série de eventos nucleares, sumariados na Fig.12.4, que se iniciam com a meiose dos micronúcleos, cujos estádios básicos são idênticos aos de outros eucariotas mas sem o acompanhamento de uma divisão celular ou da destruição da membrana nuclear (Martindale *et al.*, 1982; Orias, 1986). A indução da meiose é assinalada pela saída do micronúcleo da sua posição normal, uma espécie de bolsa situada na superfície macronuclear, seguida por um assinalável alongamento nuclear originando uma estrutura designada por crescente, que no seu máximo atinge o dobro do comprimento celular e toma a forma de um U. Quando o crescente encolhe, podem-se observar as tétradas típicas da metáfase I, seguidas em rápida sucessão da anáfase I e das diversas fases da meiose II. Um dos quatro produtos,

aleatoriamente seleccionado, permanece junto da junção de conjugação, enquanto os outros três migram para a região posterior da célula e são desintegrados (Fig.I2.4, passo 2). O núcleo sobrevivente origina, por mitose gametogénica, dois núcleos haplóides idênticos: um, denominado pronúcleo migratório, permanece junto da junção, enquanto o outro, denominado pronúcleo estacionário, migra para uma posição mais próxima do macronúcleo (Fig.I2.4, passo 3). A distinção entre os diversos núcleos que coabitam no citoplasma durante as diversas fases da conjugação, mesmo quando são absolutamente idênticos como é o caso de ambos os pronúcleos, é um dos fenómenos mais impressionantes e pior compreendidos dos que ocorrem nas células de ciliados. Aparentemente, a distinção é espacial, ou seja, apenas dependente da localização dos núcleos, mas nada se sabe sobre a natureza dos sinais bioquímicos que direccionam a função nuclear nem como são segregados no local correcto nas diferentes regiões da célula conjugante. No caso do pronúcleo migratório, sabe-se contudo que, após a sua chegada à vizinhança da junção, é envolvido por uma rede hemisférica microtubular denominada cesto de fertilização, cuja formação requer a presença do pronúcleo e é independente para cada conjugante (Orias *et al.*, 1983; Orias, 1986). O pronúcleo é então transferido para o par conjugado num processo que requer aparentemente a força motora dos microtúbulos, já que o bloqueio da despolimerização destes bloqueia igualmente a transferência (Orias, 1986). A passagem do pronúcleo através da junção de conjugação parece ser conseguida através de uma deformação programada desta, tornando-se os poros mais largos e formando como que uma cortina de membrana tubular que permite a passagem (Orias & Orias, 1983). Após a transferência, cada pronúcleo migratório funde-se com o pronúcleo estacionário residente, numa fertilização recíproca que origina um núcleo diplóide em cada célula conjugante (Fig.I2.4, passo 4).

O desenvolvimento pós-zigótico inicia-se com a dupla divisão mitótica do micronúcleo, dando origem a quatro núcleos diplóides. A segunda divisão pós-zigótica é orientada longitudinalmente, com fusos muito longos, pelo que os núcleos resultantes ficam colocados em lados opostos das células (Fig.I2.4, passo 5). Os dois núcleos anteriores irão originar dois macronúcleos, através de um processo bastante complexo (§ 2.3), enquanto os dois núcleos posteriores permanecem micronucleares (Fig.I2.4, passo 6). Tal como acontece com os núcleos pré-zigóticos, também aqui parece ser determinante a localização de cada núcleo para a definição do seu futuro papel, sendo todos os núcleos totipotentes. De facto, a recolocação, em *Tetrahymena*, dos núcleos pós-zigóticos por centrifugação origina macronúcleos a partir dos núcleos recolocados na região anterior, enquanto os núcleos recolocados na região posterior permanecem micronúcleos (Nanney, 1953). Igualmente, quando a formação do segundo fuso mitótico é impedida, a diferenciação macronuclear é atrasada significativamente, e podem-se observar produtos intermédios na sua aparência entre micro e macronúcleos (Hamilton *et al.*, 1988). Os macronúcleos e um dos micronúcleos migram então para o centro da célula, enquanto o macronúcleo original migra para a região posterior onde, conjuntamente com o micronúcleo restante, é destruído (Fig.I2.4, passo 7). A destruição do macronúcleo original possui várias semelhanças com os processos de apoptose observados em organismos superiores. Em *Tetrahymena*, a eliminação possui duas fases distintas: uma primeira fase de condensação ou picnose acompanhada pela degradação do DNA em fragmentos oligonucleossómicos (que podem ser maiores noutros ciliados: *Paramecium* produz 30-50 fragmentos, e *Euplotes* produz 4), e uma segunda fase de absorção dos fragmentos (Davis *et al.*, 1992). A degradação é aparentemente controlada pelo novo macronúcleo: a remoção deste permite a regeneração do macronúcleo original mesmo que a degradação já se tenha iniciado (Mikami &

Hiwatashi, 1975), e a conjugação de uma estirpe nulissômica do cromossoma 3 de *Tetrahymena* origina células capazes de realizar a picnose e fragmentação do macronúcleo original mas que falham a sua absorção (Davis *et al.*, 1992). Estes resultados sugerem a presença de um actividade repressora persistente mas reversível por parte do novo macronúcleo à qual só o macronúcleo original é sensível. Em *Paramecium*, a existência de um factor citoplasmático repressivo foi confirmada por microinjecção, mas a sua natureza permanece desconhecida (Berger, 1973).

Após a sua separação, cada uma das células resultantes, denominadas exconjugantes, contém dois macronúcleos e um micronúcleo (Fig.I2.4, passo 7). Logo que são realimentadas, estas células sofrem a sua primeira divisão assexuada, a única do seu ciclo em que não há divisão macronuclear. Cada uma das células filha, denominadas carionides, recebe então um macronúcleo diferenciado independentemente, e um produto mitótico do micronúcleo exconjugante (Fig.I2.4, passo 8). Uma vez que os núcleos todas as quatro células-filhas resultam de núcleos zigóticos idênticos, elas são genotipicamente idênticas e constituem um sinclone (um clone uniforme). Todavia, a conjugação não é isenta de consequências genéticas, visto que ambos os exconjugantes adquirem uma herança citoplasmática e mitocondrial distinta, podendo por isso apresentar diferentes fenótipos (Roberts & Orias, 1973). Para além disso, a independente diferenciação do novo macronúcleo pode originar igualmente diferentes fenótipos associados a cada carionide, como seja a atribuição do "mating" em *T.thermophila* (# 2.2.2.3; Nanney, 1964).

Cada carionide é agora uma célula capaz de um crescimento vegetativo normal. Todavia, em *Tetrahymena*, há ainda um período em que a progenia é incapaz de formar pares conjugantes. Este período de imaturidade é variável consoante a espécie estudada, o seu genótipo ou o ambiente de crescimento, mas pode ir de 40 a 100 gerações em *T.thermophila* (Allen & Gibson, 1973). A este período de imaturidade segue-se um período de adolescência, em que as células são incapazes de conjugar entre si, mas capazes de o fazer com células sexualmente maduras (Rogers & Karrer, 1985). Só após este período, que vai de 20 a 25 gerações, as células se encontram sexualmente maduras. Os mecanismos biológicos que determinam cada um destes períodos são ainda desconhecidos, mas parecem estar relacionados com a presença do novo macronúcleo, visto que quando este não se desenvolve e o macronúcleo original é retido, as células não passam por um período de imaturidade (# 2.2.2.2).

2.2.2.2. Desvios importantes à conjugação normal.

Em certas situações, que podem ser expontâneas ou induzidas artificialmente, a conjugação não se processa normalmente e ocorrem desvios, que para além da sua contribuição para o conhecimento da genética dos ciliados, têm mesmo sido usados como ferramentas de trabalho neste tipo de organismo. Estes desvios podem ser classificados em quatro tipos principais: citogamia, fertilização abortada, retenção macronuclear e conjugação tripla.

Na citogamia (Fig.I2.5.A), o pronúcleo migratório não é transferido para o par conjugante e funde-se com o pronúcleo estacionário, seguindo-se um desenvolvimento pós-zigótico normal. As células-filhas resultantes são homocariontes e totalmente homozigóticas, embora cada um dos exconjugantes seja geneticamente diferente (Orias & Hamilton, 1979). A citogamia ocorre com uma frequência de 1% em *Tetrahymena*, mas pode ser induzida por choque osmótico ou agentes anti-microtubulares (Orias *et al.*, 1979).

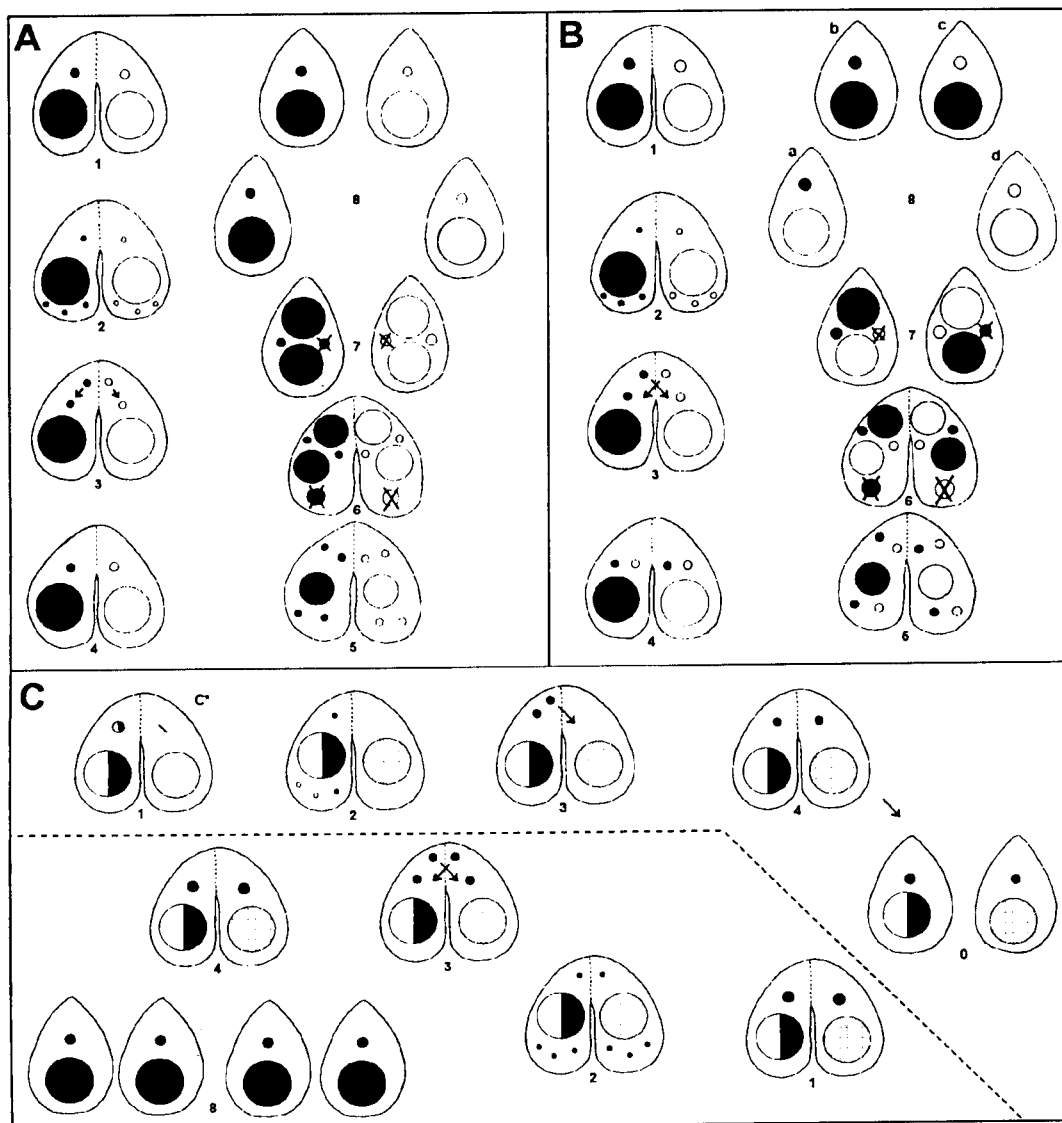


Fig.I2.5. Desvios importantes à conjugação normal. Os acontecimentos nucleares são equivalentes aos indicados na Fig.I2.4. (A) Citogamia. (B) Fertilização abortada. (C) Exclusão genômica. Neste caso, na primeira conjugação ocorre uma retenção macronuclear, e as células resultantes podem regressar à vida vegetativa (parte superior do tracejado) ou ser novamente cruzadas numa conjugação normal originando carionides homocariontes e homozigóticas (parte inferior do tracejado). C* representa uma estirpe "star". Adaptado de Orias (1986) e Bruns (1986).

Uma variante da citogamia, a autogamia, é uma parte integrante do ciclo de vida de alguns ciliados, como algumas espécies de *Paramecium*. A autogamia tem consequências genéticas idênticas à citogamia, mas ocorre em células não conjugadas (Miyake, 1981).

A fertilização abortada (Fig.I2.5.B) é bastante menos frequente (<0.1%), mas pode ser induzida com um tratamento apropriado com vimblastina (Hamilton *et al.*, 1988). Nesta situação, o pronúcleo migratório transfere-se normalmente para o par conjugante, mas não se funde com o pronúcleo estacionário residente. Ambos os núcleos sofrem uma endoreduplicação, tornando-se diplóides, e seguem para a segunda mitose pós-zigótica. Tal como na citogamia, as células-filhas resultantes são homozigóticas, mas como só um dos micronúcleos potenciais é retido, cada exconjugante origina dois tipos de carionides: um homocarionte, com ambos os núcleos derivados de um dos conjugantes, e um heterocarionte, em que o macronúcleo deriva de um dos conjugantes e o micronúcleo do outro.

Na retenção macronuclear, o novo macronúcleo falha o seu desenvolvimento e o macronúcleo original permanece funcional (Allen, 1967). As células-filhas resultantes são fenotipicamente idênticas às conjugantes, e contrariamente à verdadeira progenia são sexualmente maduras. A retenção macronuclear é geralmente pouco frequente, mas é predominante durante o fenómeno de exclusão genómica (Fig.I2.5.C), que ocorre aquando do cruzamento de células normais com algumas estirpes de *Tetrahymena*, chamadas "star", que se caracterizam por perderem o micronúcleo durante a meiose (Allen, 1967). Assim, apenas o pronúcleo migratório da célula normal é transferido, e cada núcleo haplóide sofre seguidamente uma endoreduplicação. Contudo, não há posterior desenvolvimento pós-zigótico, pelo que cada uma das células conjugantes se separa retendo os macronúcleos originais. Curiosamente, as células resultantes comportam-se de maneira diferente: enquanto a célula derivada da conjugante normal permanece fértil e viável, a outra reverte a uma condição "star" ao fim de cerca de 40 gerações, o que sugere uma destabilização do novo micronúcleo por parte do macronúcleo original (Weindruch & Doerder, 1975). As células podem ser todavia imediatamente cruzadas, desenvolvendo-se uma conjugação normal em que as carionides resultantes são homocariontes e homozigóticas (Fig.I2.5.C).

Outro fenómeno pouco frequente envolve uma conjugação entre três células, e não duas como na conjugação normal (Preparata & Nanney, 1977). Todas as três passam por uma fase pré-zigótica normal, mas enquanto duas procedem a uma troca recíproca dos seus pronúcleos migratórios, a terceira transfere o seu a uma das outras mas não recebe nada em troca, pelo que, após a fertilização, o núcleo zigótico das células é respectivamente haplóide, diplóide e triplóide. Curiosamente, e ao contrário da exclusão genómica, todas as três prosseguem com o desenvolvimento pós-zigótico normal, possuindo os macronúcleos resultantes uma quantidade de DNA equivalente mas que reflecte o conteúdo alélico de cada um dos núcleos zigóticos.

É possível retirar alguns dados importantes com a observação destes desvios à conjugação normal. Aparentemente, a meiose pré-zigótica é relativamente independente para cada célula conjugante, já que prossegue mesmo com a ausência de um dos micronúcleos; contudo, é preciso notar que, durante a citogamia, é a migração de ambos os pronúcleos migratórios que é afectada, e não de apenas um. Por outro lado, há uma independência entre o desenvolvimento pré-zigótico e pós-zigótico, não sendo este automaticamente desencadeado pela ocorrência do primeiro. Na verdade, a origem do sinal que desencadeia o desenvolvimento pós-zigótico é a que se apresenta mais obscura. É possível que seja necessária a presença de dois pronúcleos haplóides, não necessariamente fundidos, em pelo menos uma das células conjugantes, o que justificaria a sua indução em células haplóides durante conjugações triplas mas não durante exclusões genómicas.

2.2.2.3. Os tipos de "mating" em *Tetrahymena*.

A genética da determinação do "mating" em *Tetrahymena* e outros ciliados é típica da sua condição de células binucleadas, e cuja compreensão é essencial para o entendimento de outros fenómenos como a conjugação. Grosso modo, o "mating" em ciliados define o seu sexo, um termo que não pode ser correctamente aplicado nestes organismos porque, como descrito previamente, a reprodução sexual não envolve a formação de gâmetas. No género *Tetrahymena*, conhecem-se dois tipos de determinação de "mating": por controlo genético directo, como em *T.pigmentosa* (Orias, 1963; Simon & Orias, 1987), ou por herança carionidal, como em *T.thermophila* (Nanney, 1956; Allen & Gibson, 1973).

Em *T. pigmentosa* existem três tipos de "mating" que, quando cruzados em várias combinações, originam quatro carionides manifestando o mesmo tipo de "mating", que depende do genótipo das células conjugantes de uma forma Mendeliana. Os resultados dos cruzamentos sugerem a presença de um único *locus* de "mating" com três alelos, $mt^A > mt^B > mt^C$. Uma célula será então de "mating" I caso possua o alelo mt^A , de "mating" II caso possua o alelo mt^B mas não o alelo mt^A , e "mating" III caso seja homozigótica para mt^C , o que evidencia a natureza regulatória do *locus* *mt*. Existe, todavia, alguma instabilidade genética na definição do "mating" que afecta cerca de 8% da progenia, causada quer "selfing" ("mating" intraclonal) ou por expressão ilegítima do alelo mais recessivo (Simon & Orias, 1987).

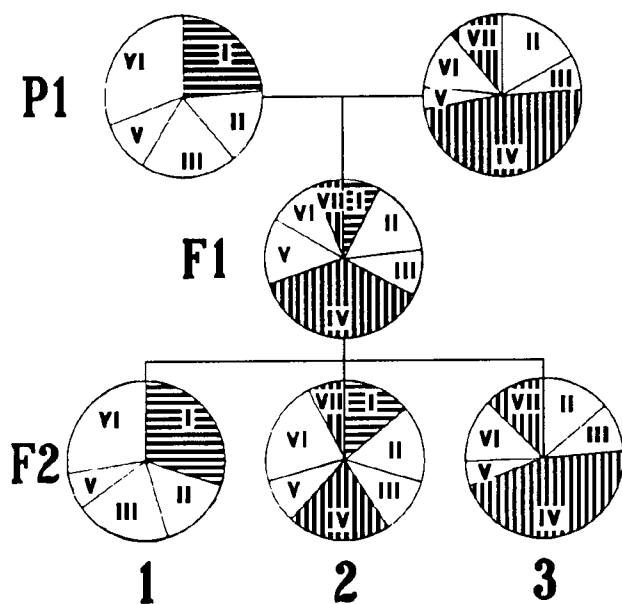


Fig.I2.6. Frequências dos tipos de "mating" em cruzamentos de *T. thermophila*. As frequências em P1 são derivadas dos dados acumulados para as famílias A (mt^A) e B (mt^B); as frequências em F1 são derivadas directamente de F1; as frequências das três classes em F2 são derivadas de testes de progenia em que carionides irmãs de pares F2 foram cruzadas. As barras horizontais representam o tipo exclusivo da família A, e as barras verticais os tipos exclusivos da família B. Adaptado de Allen & Gibson (1973).

A determinação do tipo de "mating" em *T. thermophila* é bastante mais complexa. Este organismo possui sete tipos de "mating" cujo cruzamento entre si origina carionides com "matings" não correlacionados entre si ou com os das células conjugantes, mas com algumas frequências características. Os resultados sugerem a existência de um *locus* com dois alelos, mt^A e mt^B , que apenas definem os tipos de "mating" que as carionides podem adquirir e a sua frequência relativa. Assim, o cruzamento de estirpes homozigóticas mt^A não pode produzir tipos IV e VII, enquanto o cruzamento de estirpes homozigóticas mt^B não é capaz de produzir o tipo I, aparecendo os restantes tipos de "mating" em frequências características (Fig.I2.6, P1). O cruzamento entre ambas as estirpes origina uma progenia heterozigótica em que os sete tipos de "mating" estão presentes, e o cruzamento dentro desta origina três classes de progenia, homozigóticas para mt^A ou mt^B e heterozigótica (Fig.I2.6). Assim, existe um controlo genético directo mendeliano quanto aos tipos de "mating" possíveis para cada carionide e respectivas frequências em que ocorrem na descendência, mas o tipo de "mating" efectivamente obtido por um dado carionide é determinado pelo desenvolvimento do seu macronúcleo. A maneira como este fenómeno ocorre é desconhecida, mas a hipótese mais simples é a do(s) factor(es) codificado(s) pelo *locus* de "mating" sofrer(em) rearranjos alternativos durante a maturação do novo

macronúcleo (# 2.3), originando assim novos factores que determinarão o tipo de "mating" de cada carionide.

2.3. O desenvolvimento macronuclear.

2.3.1. Em *Tetrahymena*.

Durante o desenvolvimento pós-zigótico da conjugação, as células de ciliados geram novos macronúcleos a partir de núcleos zigóticos diplóides, enquanto degradam os macronúcleos originais (# 2.2.2). A formação do novo macronúcleo é um processo ordenado e reprodutível que, em *Tetrahymena*, envolve pelo menos três processos distintos: a fragmentação dos cromossomas em várias centenas de fragmentos subcromossómicos, a eliminação de aproximadamente 15% do genoma micronuclear e a amplificação do DNA genómico. Não é de excluir outro tipo de rearranjos neste ciliado, como a reorganização de sequências, inversões ou inserções, que todavia não estão descritos até à data. A facilidade de isolamento, em *Tetrahymena*, do DNA dos macronúcleos em formação (Allis & Dennison, 1982) associada à clonagem e transformação de fragmentos susceptíveis de sofrer rearranjos veio permitir uma melhor compreensão dos processos envolvidos e de alguns dos seus mecanismos. Assim, observa-se que a indução do desenvolvimento do novo macronúcleo é assinalada por uma amplificação inicial do genoma zigótico a $4n$ (Yokoyama & Yao, 1982). Nesta altura, e durante um período aproximado de duas horas em que o genoma se amplifica de $4n$ a $8n$, decorre a maioria, senão a totalidade, de todos os arranjos genómicos observados (Austerberry *et al.*, 1984; Brunk & Conover, 1985; Allito & Karrer, 1986). Perto do fim deste período, o novo genoma é já transcricionalmente activo (Orias, 1986; Mayo & Orias, 1986), e sofre amplificação até ao seu valor final de cerca de $45n$.

A fragmentação dos cromossomas ocorre extensivamente durante o desenvolvimento do novo macronúcleo, estimando-se que o DNA se quebra em 50 a 200 locais específicos, e que quer a sua ocorrência quer a sua localização são largamente invariáveis (Altschuler & Yao, 1985; Conover & Brunk, 1986). Dois desses locais flanqueiam os genes do RNA ribossomal, sendo por isso os mais estudados. O rDNA existe em cópia única no micronúcleo de *Tetrahymena*, sendo processado numa molécula linear livre por quebra do cromossoma nas duas extremidades dos genes, com posterior conversão a um dímero palindrómico e amplificação a cerca de 10^4 cópias. A fragmentação do rDNA envolve a eliminação de várias centenas de nucleótidos de DNA de ambos os locais clivados e a adição de telómeros nas extremidades livres obtidas (Yao, 1986). Perto das junções de clivagem foi encontrada uma pequena região homóloga contendo uma sequência de 15 nucleótidos - 5'-AAAGAGGTGGTTTA-3' -, denominada Cbs ("Chromosome breakage sequence"), que se demonstrou existir em mais de 50 cópias no micronúcleo, mas estar ausente no macronúcleo (Yao *et al.*, 1987). A clonagem de vários fragmentos de DNA micronuclear contendo sequências Cbs revelou que a maioria destas sequências, se não todas, se encontram associadas a locais onde ocorre fragmentação cromossómica, e foi efectivamente demonstrado que a sua presença é necessária e suficiente para promover a fragmentação (Yao *et al.*, 1987; 1990). A fragmentação ocorre em simultâneo com a perda de um segmento de meia centena de nucleótidos que inclui a sequência Cbs (a perda de centenas de nucleótidos observada no rDNA deverá constituir uma excepção; Yao *et al.*, 1987) e a adição de telómeros nas extremidades livres. Esta adição é realizada *de novo*, apesar de a acção enzimática da telomerase requerer aparentemente a presença prévia de

sequências teloméricas (Yu & Blackburn, 1991; # 3.3); todavia, não é de excluir um papel da sequência Cbs na telomerização, podendo o facto de a subsequência 5'-GGTTGG-3' da Cbs ser uma permutação da sequência telomérica 5'-CCCCAA-3' não constituir uma mera coincidência (Yao *et al.*, 1987).

O desenvolvimento macronuclear envolve igualmente a eliminação precisa de pelo menos 6000 fragmentos de DNA, correspondentes a cerca de 15% do genoma micronuclear (Yao *et al.*, 1984). Vários destes fragmentos têm sido clonados e analisados, revelando-se serem na sua maioria membros de famílias de sequências repetidas específicas do micronúcleo, cujos elementos são na sua quase totalidade eliminados durante a maturação do macronúcleo (Yao, 1982; Karrer, 1983; White *et al.*, 1985; Allito & Karrer, 1986; Katoh *et al.*, 1993). A função destas sequências, que se distribuem pelos cinco cromossomas micronucleares, permanece contudo desconhecida. A sua análise não revela regiões codificantes significativas, e possuem um baixo conteúdo de GC (20-25%) (Allito & Karrer, 1986; Austerberry & Yao, 1987; Katoh *et al.*, 1993). Aparentemente, e de modo diferente ao observado em ciliados Hipotricos (# 2.3.2), os fragmentos eliminados não interrompem genes, embora tenham sido detectados em regiões flanqueantes dos mesmos (Callahan *et al.*, 1984; Kile *et al.*, 1988; Katoh *et al.*, 1993). No único destes casos onde a sequência nucleotídica foi determinada, correspondendo a um fragmento de 1.4 kb localizado a cerca de 3.5 kb a montante do gene de calmodulina de *T. thermophila*, foram detectados dois elementos comuns ao fragmento eliminado e à região promotora do gene (Katoh *et al.*, 1993). O seu significado é desconhecido, tendo sido levantada a hipótese de poderem modular a expressão diferencial deste gene no micro e macronúcleo. Contudo, não sendo embora de excluir a hipótese de que pelo menos parte das recombinações observadas possa estar envolvida na activação ou diferenciação de genes (como p.ex. a definição de "mating"), é possível que a maioria do DNA específico do micronúcleo pode possuir apenas um papel estrutural relacionado com funções típicas deste núcleo, como a mitose e meiose (Karrer, 1986).

O estudo dos mecanismos que levam à excisão de fragmentos tem revelado algumas particularidades pouco usuais. De uma maneira geral, as extremidades dos fragmentos a eliminar são reconhecidas e clivadas, seguindo-se a religação das extremidades livres e a degradação do fragmento interno excisado, um pouco à semelhança do que ocorre durante o "splicing" de RNAs. O reconhecimento é bastante preciso, pelo que o processo é reprodutível em cada maturação do macronúcleo. Contudo, podem existir num determinado segmento de DNA mais de dois locais de clivagem e, assim, a deleção reconhecer indiferentemente dois desses locais, originando dois ou mais fragmentos distintos (Yao *et al.*, 1984; Austerberry *et al.*, 1989). A clivagem e posterior junção das extremidades, apesar de precisa, não é exacta, e podem observar-se microheterogeneidades de 1-5 bases na sequência de junção (Austerberry & Yao, 1987; Austerberry *et al.*, 1989). Estas variações, que deverão observar-se na esmagadora maioria das deleções efectuadas durante a maturação, originarão necessariamente uma heterogeneidade geral extremamente elevada nos macronúcleos resultantes (2^{6000} , para 6000 locais produzindo 2 junções diferentes), o que poderá justificar a diferente herança carionidal observada durante a conjugação (Austerberry *et al.*, 1989). Um dos sinais implicados na deleção do DNA foi identificado como sendo a sequência polipurínica 5'-A₅G₅, que se demonstrou ser necessária e suficiente para determinar as extremidades da deleção (Godiska & Yao, 1990; Godiska *et al.*, 1993). A clivagem ocorre a uma distância de 40-50 bases da sequência, preferencialmente em sequências directas de 1-4 bases desde que estas se encontrem à distância apropriada (Godiska *et al.*, 1993). Curiosamente, esta sequência ocorre igualmente

perto dos locais de excisão dos genes *spoIVBC* e *spoIIIC* de *Bacillus subtilis*, que codificam para os dois segmentos do factor de transcrição σ^k , sugerindo que este novo mecanismo poderá não ser exclusivo de *Tetrahymena* (Stragier *et al.*, 1989). Todavia, nem todas as deleções neste ciliado são susceptíveis de serem explicadas através deste sinal. Um outro sinal poderá ser constituído por palíndromos de cerca de 10 pb com uma base de intervalo, detectados em ambas as extremidades de dois fragmentos excisados, mas cuja implicação na deleção permanece por provar (Austerberry & Yao, 1987; Katoh *et al.*, 1993).

Dos três processos aparentemente envolvidos na maturação do novo macronúcleo, a amplificação do DNA é seguramente o pior compreendido. Nada se sabe sobre o sinal que o desencadeia, nem sobre o que determina o número de cópias a manter. Os poucos dados disponíveis foram determinados no fragmento subcromossómico contendo os rDNAs. Este fragmento possui uma única origem de replicação, que é responsável pela manutenção vegetativa, mas cuja influência sobre a amplificação é desconhecida (Cech & Brehm, 1981). É possível que a sequência Cbs possua igualmente um papel na amplificação, para além do evidenciado anteriormente sobre a fragmentação cromossómica: recentemente foram isolados mutantes, cuja sequência Cbs apresenta uma única substituição, que, para além de uma fragmentação defeituosa, apresenta igualmente uma amplificação do rDNA deficiente mesmo quando a fragmentação é conseguida (Kapler & Blackburn, 1994).

Curiosamente, os mecanismos envolvidos na maturação do novo macronúcleo em *Tetrahymena* não parecem ser idênticos aos observados noutros ciliados, incluindo o parente mais próximo, *Paramecium*. Neste organismo, a fragmentação cromossómica é menos precisa, já que após a clivagem, ambas as extremidades sofrem deleções variáveis, que são na ordem das centenas de pares de bases para uma das extremidades e dos milhares para a outra. Os fragmentos podem igualmente sofrer clivagens internas que também estão sujeitas a uma acção exonucleotídica, sendo os subfragmentos resultantes religados ou, alternativamente, transformados em fragmentos cromossómicos por adição de telómeros. Este mecanismo gera, assim, para cada região do genoma micronuclear, uma heterogeneidade de fragmentos e subfragmentos cromossómicos (Caron, 1992). O macronúcleo original parece possuir um papel fundamental na formação do novo macronúcleo, já que é possível observarem-se locais de clivagem preferenciais em células com micronúcleos idênticos mas macronúcleos distintos. Este fenómeno é muito importante em *Paramecium*, já que é responsável por elevados fenómenos de herança não mendeliana que têm vindo a ser observados (Epstein & Forney, 1984; Harumoto, 1986; Meyer, 1992; Scott *et al.*, 1994). Por outro lado, os fragmentos de DNA eliminados são muito mais pequenos do que em *Tetrahymena* (cerca de 30-800 pb) e interrompem frequentemente regiões codificantes. Cada fragmento é ladeado por dinucleótidos 5'-TA-3', que podem ou não ser parte integrante de uma repetição directa ou invertida mais extensa, de 2 a 11 pb (Steele *et al.*, 1994; Amar, 1994). Este tipo de estrutura é muito semelhante ao encontrado em *Euplotes*, embora o mecanismo de excisão neste ciliado Hipotrico pareça envolver igualmente um elemento 5'-TTGAA-3' que ocorre menos frequentemente em *Paramecium* (# 2.3.2). Finalmente, a excisão dos fragmentos não é completa, e a sua presença ou ausência em fragmentos macronucleares parece estar intimamente correlacionada com locais alternativos de clivagem, o que sugere que ambos os mecanismos possam ter factores comuns e possam estar sujeitos a competição (Amar, 1994).

2.3.2. Em ciliados Hipotricos.

O desenvolvimento macronuclear em ciliados Hipotricos apresenta várias características particulares que o distinguem do de Holotricos. A mais marcante será a extensa eliminação de sequências, que atinge cerca de 95% do genoma micronuclear, mas a elevada fragmentação do genoma, que origina fragmentos macronucleares contendo aparentemente um único gene cada um, e a existência de reorganização de genes durante a maturação são igualmente assinaláveis.

O desenvolvimento em Hipotricos estende-se por um período de tempo cerca de dez vezes superior ao de ciliados Holotricos, o que permitiu determinar com maior rigor a sequência temporal dos diversos processos observados. A maturação do novo macronúcleo inicia-se com a amplificação dos cromossomas, que irão formar estruturas politénicas semelhantes às observadas em células de *Drosophila*, mas cujo superior número de bandas as torna impossíveis de mapear (Klobutcher & Prescott, 1986). O início da fase politénica, que ocorre cerca de 30 horas após o contacto das células, é coincidente com a excisão das sequências repetidas do genoma, semelhantes a transposições (ver adiante; Jahn *et al.*, 1989; Krikau & Jahn, 1991). Perto do fim desta fase (45-60 horas após o contacto) decorre a excisão das sequências únicas e a reorganização dos genes "baralhados" (Tausta & Klobutcher, 1989; 1990; Mitcham *et al.*, 1992). Os cromossomas politénicos apresentam agora estruturas vesiculares que os dividem em várias centenas de compartimentos, e que assinalam o início da fragmentação maciça que os irá transformar em pequenos fragmentos de DNA monogenéticos (Kloetzel, 1970; Baird & Klobutcher, 1989; Tausta & Klobutcher, 1990). Nas cinco horas seguintes são adicionadas sequências teloméricas às extremidades livres dos fragmentos, que são seguidamente amplificados até aos seus valores habituais em células vegetativas (Tausta & Klobutcher, 1990).

O primeiro grupo de fragmentos de DNA eliminado é constituído por famílias de sequências longas e repetidas cuja estrutura se assemelha à de transposições. Estas famílias foram detectadas em *Oxytricha* (TBE1; Herrick *et al.*, 1985) e *Euplotes* (Tec1 e Tec2; Baird *et al.*, 1989; Jahn *et al.*, 1989; Krikau & Jahn, 1991), podendo igualmente existir em *Tetrahymena* (Tel-1; Cherry & Blackburn, 1985). A análise das sequências codificantes, inferidas a partir das suas sequências nucleotídicas, sugerem que estes transposições estão relacionados com um largo grupo que inclui transposições de organismos tão distintos como *Caenorhabditis* e *Drosophila* (Doak *et al.*, 1994). Em *Euplotes crassus*, onde têm sido estudados mais intensamente, os elementos das duas famílias encontradas possuem cerca de 5.3 kb com extremidades constituídas por sequências inversamente repetidas de 700 a 725 pb, estando cada família presente em cerca de 30 000 cópias no genoma micronuclear (Jahn *et al.*, 1989; Krikau & Jahn, 1991). Este elevado número de cópias, igualmente observado para o transposão conhecido de *Oxytricha* (cerca 2000 cópias/genoma), interrompe pelo menos 50% das sequências que irão originar fragmentos macronucleares, pelo que a sua presença só é "tolerada" evolutivamente devido à sua eliminação maciça durante o desenvolvimento macronuclear. O mecanismo de excisão dos transposições constitui um processo novo, que envolve aspectos comuns a outros fenómenos decorrentes durante a maturação do novo núcleo. O reconhecimento dos locais de clivagem parece envolver uma sequência conservada cujo núcleo - 5'-TTGAA-3' - é comum a uma outra sequência conservada encontrada próxima dos locais de fragmentação do cromossoma (ver adiante; Krikau & Jahn, 1991; Jaraczewski & Jahn, 1993). A clivagem ocorre a 17 nucleótidos da sequência TTGAA, coincidindo com as sequências repetidas directas 5'-TA-3' que ocorrem nas extremidades dos transposições. O transposão clivado é então

circularizado, sendo o seu ponto de junção constituído por uma estrutura heteroduplex de 10 nucleótidos flanqueados por duas sequências directas 5'-TA-3'. Esta estrutura sugere um modelo de excisão em que a clivagem é desfasada em ambas as cadeias, à semelhança do que se observa por acção de enzimas de restrição. As extremidades protuberantes do transposão, após circularização, constituiriam a estrutura heteroduplex, enquanto as extremidades protuberantes das sequências micronucleares seriam preenchidas por uma polimerase e religadas (Jaraczewski & Jahn, 1993).

Este modelo de excisão é igualmente observado, com algumas variantes, durante a excisão das sequências únicas. Estas sequências, chamadas sequências internas eliminadas (IES), são muito mais pequenas do que os transposões (10-600 pb) e, ao contrário destes, são únicas no genoma, embora o seu número possa atingir 60 000 em *Oxytricha* (Ribas-Aparicio *et al.*, 1987). Ocorrem indistintamente no genoma, interrompendo sequências micronucleares que irão originar sequências macronucleares, quer estas sejam ou não sequências codificantes, pelo que se assemelham de certa forma a intrões de DNA. Em *Euplotes*, as IES estão flanqueadas por repetições directas 5'-TA-3', e a sua remoção origina formas circulares com um ponto de junção em estrutura heteroduplex, como observado para os transposões (Klobutcher *et al.*, 1993). A análise pormenorizada do ponto de junção revela que este é constituído, em cada cadeia, por 10 nucleótidos rodeados por duas sequências directas 5'-TA-3', sendo 2 nucleótidos originários de uma das extremidades exteriores à IES e os outros 8 da outra (excisão 8+2), constituindo os 6 centrais a estrutura em heteroduplex. Esta estrutura sugere um modelo (Fig.I2.7) em que a clivagem é desfasada de 10 nucleótidos, sendo a circularização das IES acompanhada de uma polimerização parcial (4 nucleótidos em cada cadeia) que originará a estrutura 8+2 (Klobutcher *et al.*, 1993). Esta estrutura poderá igualmente estar presente na excisão dos transposões, mas o seu elevado numero de cópias não permitiu ainda uma análise pormenorizada do ponto de junção. Contudo, a ausência de sequências conservadas próximas dos locais de clivagem das IES sugere a presença de um mecanismo distinto para o seu reconhecimento, o que poderá justificar a diferença temporal observada entre os períodos de excisão das IES e dos transposões.

Em *Oxytricha*, as IES originam igualmente estruturas circulares, mas existem algumas evidências que sugerem poder-se estar na presença de um mecanismo de excisão algo distinto do observado em *Euplotes*. Primeiro, as IES não estão flanqueadas por sequências repetidas 5'-TA-3', nem necessariamente por sequências repetidas; segundo, as sequências

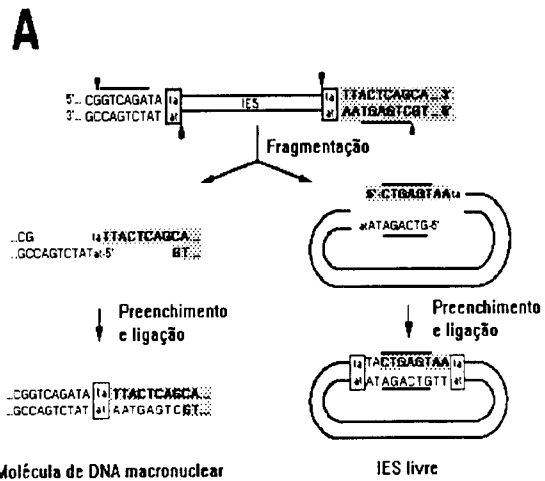


Fig.I2.7. Modelo de excisão do IES. O fragmento contendo o IES é clivado desfasadamente em ambas as extremidades (pontas de setas), e as extremidades livres associam-se por sobreposição de 6 pb que formarão a estrutura heteroduplex do ponto de junção (linhas). Os espaços vazios são preenchidos pela polimerase e ligados, originando uma molécula circular em cujo ponto de junção cada cadeia de DNA contém 8 bases de uma extremidade (a sombreada) e 2 da outra (excisão 8+2). As extremidades que irão formar o DNA micronuclear associam-se pelas sequências repetidas 5'-TA-3' (em caixa) e são preenchidas e ligadas, resultando numa molécula com duas cópias de cada repetição. Adaptado de Klobutcher *et al.* (1993).

destinadas ao macronúcleo (MDS) não se encontram necessariamente dispostas no micronúcleo na mesma ordem, ou mesmo na mesma orientação, em que se irão dispor no macronúcleo, um tipo de organização genómica até agora único de *Oxytricha* (Ribas-Aparicio *et al.*, 1987; Greslin *et al.*, 1989; Mitcham *et al.*, 1992). As repetições directas que se observam localizam-se nas extremidades de MDS que irão ficar adjacentes no macronúcleo, mas a sequência e número de nucleótidos varia para cada repetição, embora pareçam ser tanto maiores quanto mais distantes estão entre si. A excisão das IES deverá então ocorrer por recombinação homóloga entre as repetições, ficando uma das cópias retida na sequência macronuclear. Este mecanismo, que permite a reorganização dos MDS desordenados, originará moléculas circulares que, ao contrário do que ocorre em *Euplotes*, podem ser compostas por várias IES (Greslin *et al.*, 1989; Mitcham *et al.*, 1992).

Em ciliados Hipotricos, os precursores das moléculas macronucleares estão habitualmente muito próximos entre si, podendo inclusivamente sobrepor-se em alguns pares de bases (Baird & Klobutcher, 1989). Estes dados sugerem que a fragmentação cromossómica é um processo extremamente preciso nestes ciliados, e de facto em *Euplotes* apenas um único local de clivagem é usado para gerar as extremidades do DNA macronuclear, embora em *Oxytricha* se possam encontrar vários locais distintos numa pequena região de clivagem de 10-30 pares de bases (Herrick *et al.*, 1987; Baird & Klobutcher, 1989). O reconhecimento do local de clivagem parece envolver uma sequência conservada cujo núcleo - 5'-TTGAA-3' - é idêntico ao encontrada na excisão de transposições (Klobutcher *et al.*, 1981; Baird & Klobutcher, 1989). É possível que o mecanismo envolvido seja semelhante, já que o local de clivagem se situa igualmente a 17 nucleótidos de distância, sendo a clivagem desfasada em ambas as cadeias pelo menos 6 nucleótidos. Contudo, as extremidades livres não são recircularizadas ou religadas e, após a adição de sequências teloméricas, a molécula é convertida num fragmento macronuclear. Constitui ainda uma incógnita porque é que, apesar da aparente semelhança dos mecanismos de excisão e fragmentação, estes processos produzem não só diferentes resultados como ainda ocorrem em diferentes momentos temporais no ciclo do desenvolvimento macronuclear.

3. Outras contribuições nos conceitos da biologia molecular.

3.1. O código genético.

A universalidade do código genético foi posta em causa pela primeira vez em 1979, com a descoberta, em mitocôndrios humanos, de genes em que os codões ATA (Ile) e TGA (terminação) codificavam, respectivamente, para Met e Trp (Barrell *et al.*, 1979). Com excepção dos mitocôndrios de plantas verdes, que usam o código universal, várias alterações ao código têm sido encontradas desde então nos mitocôndrios dos mais diversos organismos, com realce para o uso de TGA como Trp, a única alteração que se tem mantido comum aos diversos códigos encontrados (para revisão, Osawa *et al.*, 1992). A particularidade própria dos mitocôndrios, originárias de eubactérias simbiotes e possuindo um genoma muito reduzido, levou, contudo, a postular-se que os desvios observados, possíveis regressões evolutivas para um código mais primitivo, seriam exclusivos destes organelos e não seriam encontrados em códigos nucleares.

A universalidade do código nuclear só ficou definitivamente posta de parte com a publicação, em 1985, de genes de três ciliados, *Paramecium* (Caron & Meyer, 1985; Preer *et al.*, 1985), *Oxytricha* (Helftenbein, 1985) e *Tetrahymena* (Horowitz & Gorovsky, 1985), onde foi posto em evidência a presença de dois codões de terminação - TAA e TAG -

codificantes para Gln, sendo a terminação da leitura exclusivamente constituída pelo codão TGA. Em *Tetrahymena*, o desvio ao código foi igualmente confirmado pela detecção de dois tRNAs-Gln citoplasmáticos com anticodões UmUA e CUA, pelo que o aminoácido Gln era efectivamente codificado pelos quatro codões YAR (Kuchino *et al.*, 1985). Este desvio ao código foi igualmente encontrado em *Stylonychia* (Helftenbein & Müller, 1988), mas não em dois outros ciliados: em *Euplotes*, um Hipotrico (tal como *Oxytricha* e *Stylonychia*), e em *Blepharisma*, um Heterotrico. Nestes ciliados, TAA é um codão de terminação, enquanto ao codão TAG não foi ainda atribuída nenhuma definição (Harper & Jahn, 1989; Liang & Heckmann, 1993a). Curiosamente, numa espécie de *Euplotes*, *E. octocarinatus*, o codão TGA tem sido encontrado em vários genes codificando para Cys (Meyer *et al.*, 1991; Kaufmann *et al.*, 1992; Liang & Heckmann, 1993b); todavia, não lhe foi atribuída ainda nenhuma definição nas outras espécies de *Euplotes* ou de *Blepharisma*, cujos genes conhecidos usam exclusivamente TAA como codão de terminação.

Após a descoberta em ciliados, a diversidade do código genético nuclear tem sido estendida a outros organismos. Já em 1985 tinham sido encontrados na bactéria *Mycoplasma* codões TGA codificantes para Trp (Yamao *et al.*, 1985), tal como observado nos mitocôndrios. Posteriormente, e à semelhança do que ocorre em ciliados, foram identificados em genes de algas verdes *Acetabularia* codões UAR codificantes para Gln (Schneider *et al.*, 1989), enquanto na levedura *Candida*, o codão CUG (Leu) foi atribuído ao aminoácido Ser (Kawaguchi *et al.*, 1989). Inicialmente, estas divergências na utilização de codões foram justificadas como relíquias de um código genético anterior ao aparecimento do código dito universal (Grivell, 1986). Todavia, a posição filogenética das espécies que usam códigos alterados, calculada através da construção de árvores filogenéticas com sequências ribossomais, mostra a sua separação de um tronco comum onde o código universal é utilizado, pelo que se supõe actualmente que os códigos alterados são recentes e derivados do código universal (Osawa *et al.*, 1992).

3.2. Os genes de rRNA e o conceito de autoexcisão ("self-splicing").

Tal como ocorre em muitos outros eucariotas, os três rRNAs componentes do ribossoma de *Tetrahymena* transcritos pela RNA polimerase I (ditos rRNA nucleolares) são codificados por genes organizados em "tandem" no genoma, ordenados 5'-17S-5.8S-25S-3', com as respectivas regiões codificantes flanqueadas por regiões não codificantes (Yao *et al.*, 1974). A quarta subunidade, o rRNA 5S, é transcrito pela RNA polimerase III e os seus genes encontram-se organizados em pequenos grupos de 280 pb dispersos no genoma (Kimmel & Gorovsky, 1976; Guerreiro *et al.*, 1993). Contudo, ao contrário do rDNA 5S, cujo número de cópias micro e macronucleares parece ser similar (Kimmel & Gorovsky, 1976), o ciliado *Tetrahymena* possui apenas uma única cópia de rDNAs nucleolar no seu micronúcleo (Yao *et al.*, 1974). A sequência de DNA contendo os rDNAs nucleolares é amplificada durante o desenvolvimento do macronúcleo e convertida numa molécula linear extracromossomal livre de 21 kb, autoreplicativa, constituída por um palíndromo de duas unidades transcricionais (Engberg *et al.*, 1974; 1976). No macronúcleo existirão assim cerca de 9000 cópias, cuja transcrição, iniciada numa sequência localizada no centro do palíndromo, origina um rRNA precursor de 35S que, após processamento, origina as moléculas de 25S e 17S (Rodrigues-Pousada *et al.*, 1975; 1979).

Em algumas espécies de *Tetrahymena*, como a *T. thermophila* e a *T. pigmentosa*, o gene codificante para o rRNA 26S é interrompido por uma sequência interveniente - intrão - de cerca de 400 pares de bases que, contudo, não apresenta as sequências típicas de junção observadas em intrões de mRNAs (Cech & Rio, 1979). De facto, os estudos realizados vieram a demonstrar que, pelo menos *in vitro*, o intrão se autoexcisa e os exões se reúnem na ausência de qualquer proteína, ao contrário do que ocorre na excisão de intrões de pré-mRNAs, que exigem a presença de um complexo proteico (Cech *et al.*, 1981; Kruger *et al.*, 1982). O mecanismo da excisão envolve duas reacções de transesterificação consecutivas (Fig. I2.8), que requerem um resíduo de guanosina como cofactor e um catião divalente, Mg^{2+} ou Mn^{2+} (Cech, 1987). A guanosina (ou um nucleótido de guanina, como o GTP) é capturado pelo intrão e actua como um nucleófilo, iniciando a primeira reacção de transesterificação. A guanosina é então substituída pela guanosina presente na extremidade 3' do intrão, e o grupo hidróxilo livre do primeiro exão actua sobre a ligação fosfodiéster 3' da nova guanosina, ligando os exões e libertando o intrão, que irá ser convertido numa molécula de RNA circular (Grabowski *et al.*, 1981). Apesar do RNA actuar como catalisador da sua própria reacção, o facto de não ser regenerado na sua forma inicial faz com que ele não possa ser considerado como uma enzima, sendo designado por ribozima. Contudo, é possível construir versões truncadas do ribozima capazes de clivar especificamente substratos de RNA ou DNA de cadeia simples sem sofrer alterações, agindo assim como verdadeiras enzimas (Herschlag & Cech, 1990).

A presença de intrões capazes de se autoexcisarem veio a ser posteriormente confirmada em outros organismos, constituindo-se assim numa nova classe de intrões (grupo I). Todavia, à excepção dos núcleos de protistas como *Tetrahymena*, o único outro RNA nuclear onde foram detectados intrões do grupo I pertence a cianobactérias. De resto, este tipo de intrões apenas tem sido detectado em mitocôndrios de fungos, cloroplastos e fagos de eubactérias (Cech, 1988; Kuhsel *et al.*, 1990; Xu *et al.*, 1990). Foi demonstrado que muitos destes intrões sofrem autoexcisão *in vitro*, mas em alguns casos a catálise fornecida pelo RNA mostrou ser suplementada *in vivo* por proteínas, presumivelmente de modo a estabilizar a estrutura activa do intrão (Lambowitz & Pearlman, 1990).

3.3. Os telómeros.

Os telómeros, as extremidades naturais dos cromossomas, foram descritos inicialmente mais devido às suas propriedades do que à sua estrutura. A necessidade de protecção das extremidades cromossómicas de ligações e recombinações indesejáveis,

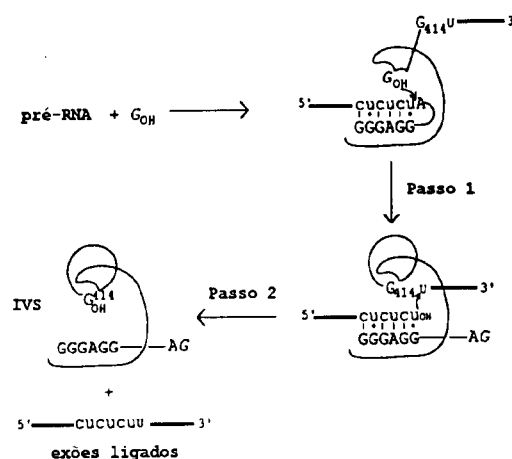


Fig. I2.8. Diagrama esquemático da autoexcisão ("self-splicing") nos intrões do Grupo I. Adaptado de Cech *et al.* (1992).

observadas em cromossomas fragmentados, levou a postular a existência de uma estrutura especializada capaz dessa protecção. Mais ainda, os estudos da replicação do DNA induziram também a necessidade de um mecanismo impeditivo da perda de bases terminais durante a replicação de moléculas lineares, sem o qual sucessivas rondas de replicação levariam a um encurtamento progressivo do cromossoma. O estudo da estrutura dos telómeros foi inicialmente dificultado pelo tamanho dos cromossomas eucariotas, e só foi ultrapassado com os estudos realizados em ciliados. Devido à organização do seu DNA macronuclear, que no caso extremo dos Hipotricos consiste em milhares de pequenas moléculas de DNA cujas extremidades terminam em estruturas que podem ser descritas como telómeros, os ciliados revelaram-se de uma grande utilidade na determinação da estrutura, mecanismos de formação e proteínas associadas aos telómeros.

Os telómeros revelaram ser complexos DNA-proteínas cuja forma, embora não a sequência precisa, é muito conservada entre os eucariotas. Consiste na repetição em "tandem" de sequências curtas de 5-10 pb, CCCCAA em *Tetrahymena* e CCCCAAAA em ciliados Hipotricos, tendo a cadeia 3' mais 12 ou 16 bases que a cadeia 5' (Yao & Yao, 1981; Klobutcher *et al.*, 1981). O número de repetições é normalmente variável com o organismo ou o seu estado fisiológico; contudo, os telómeros de *Oxytricha* e *Euplotes* exibem uma regulação restrita do seu comprimento, possuindo exactamente 20 e 28 pb de cadeia dupla telomérica, respectivamente, com uma extensão de 16 ou 14 bases (Klobutcher *et al.*, 1981). Foi também nestes organismos que se isolaram inicialmente as proteínas que se ligam especificamente à extremidade protuberante do DNA telomérico (Gottschling & Zakian, 1986; Price & Cech, 1987; Price, 1990). A proteína de ligação ao telómero de *Oxytricha* é um heterodímero de 97 kDa, cuja subunidade α , de 56 kDa, determina a especificidade da sequência de interacção, parecendo a subunidade β , de 41 kDa, apenas modificar a ligação da subunidade α (Price & Cech, 1989; Gray *et al.*, 1991). Já a proteína correspondente de *Euplotes* foi isolada como um monómero de 51 kDa, não necessitando aparentemente do envolvimento de uma segunda proteína para a ligação (Price *et al.*, 1992). Qualquer destas proteínas liga-se muito fortemente às extremidades, só sendo possível libertá-las na presença de elevadas forças iónicas.

A enzima responsável pela adição e manutenção das sequências teloméricas,

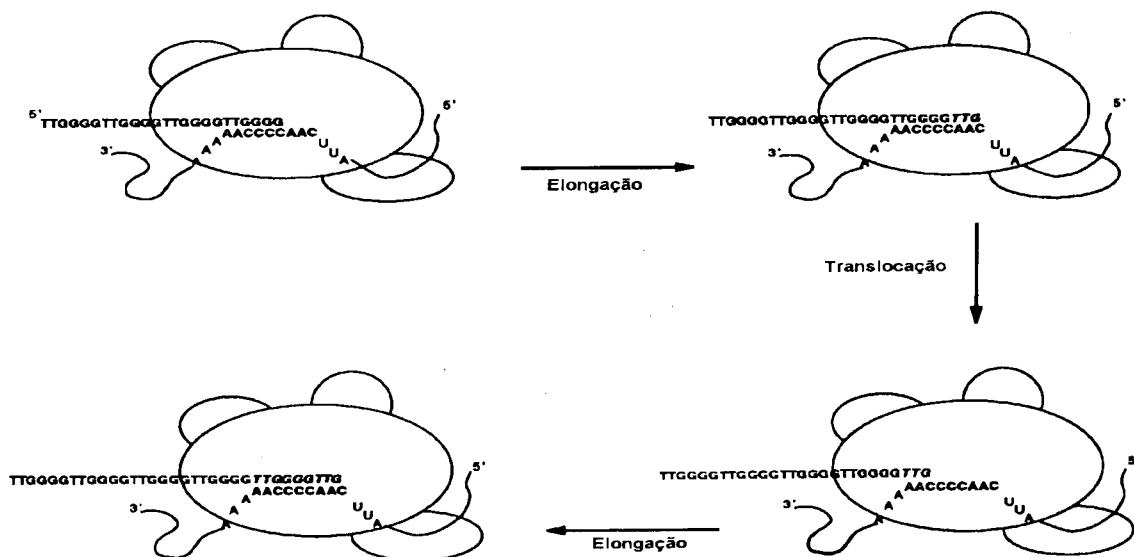


Fig.I2.9. Modelo para a acção da telomerase em *Tetrahymena*. Adaptado de Zakian *et al.* (1990).

telomerase, foi inicialmente isolada em *Tetrahymena*, e posteriormente em *Oxytricha*, *Euplotes* e células HeLa humanas (Greider & Blackburn, 1987, Blackburn, 1991). A telomerase é uma transcriptase reversa pouco usual, já que o RNA que serve de molde é uma parte intrínseca da proteína. Apesar de pouco ser conhecido sobre a componente proteica da telomerase, o RNA componente já foi isolado em *Euplotes* e *Tetrahymena*, consistindo numa molécula de 150-200 bases contendo uma região de 6-12 bases complementar à cadeia 3' do telómero (Shippen-Lentz & Blackburn, 1990; Romero & Blackburn, 1992). Este facto levou à proposição de um modelo de translocação/polimerização na actuação da enzima (Fig.I2.9), no qual o RNA da telomerase, após reconhecer a extremidade telomérica, serve de molde para a elongação da cadeia; a enzima é então translocada ao longo desta, e segue-se uma nova polimerização, prosseguindo o ciclo até à libertação da enzima (Greider & Blackburn, 1989; Blackburn, 1991). Embora este modelo pressuponha a prévia existência de uma extremidade protuberante 3' como iniciadora, os telómeros são efectivamente adicionados *de novo* a extremidades livres durante a maturação do macronúcleo em ciliados, sem que exista aparentemente nenhuma sequência consensual ou pré-telomérica iniciadora (Yu & Blackburn, 1991). É possível que, em determinadas circunstâncias, os requerimentos de sequências por parte da telomerase possam ser relaxados, podendo interagir com sequências da extremidade 3' ou internas do DNA iniciador (Harrington & Greider, 1991; Morin, 1991).

Capítulo II

Material e Métodos

1. Células e condições de cultura.

1.1. Células eucariotas: *Tetrahymena* spp.

1.1.1. Espécies e estirpes usadas.

Na realização deste trabalho utilizou-se como modelo duas espécies de protozoários ciliados do género *Tetrahymena*.

T. pyriformis - É uma espécie amiconucleada que, por conseguinte, se reproduz por fissão binária. Neste trabalho foi utilizada a estirpe **CGL**, que se desenvolve rapidamente em meios axénicos, sendo um organismo auxotrófico relativamente às purinas, pirimidinas e dez aminoácidos. Quando cultivada em meio rico à temperatura óptima de crescimento de 28°C, as células possuem um tempo de geração de cerca de 3 horas, atingindo a fase de crescimento estacionário com valores superiores a 7×10^5 células/ml.

T. thermophila - É uma espécie micronucleada que se reproduz por fissão binária ou por reprodução sexuada (conjugação). É cultivada em meio rico ou em 1% proteose peptona à temperatura óptima de crescimento de 30°C, possuindo nestas condições um tempo de geração de 3-3h30. Foram usadas as seguintes estirpes:

SB210 - Usada para análise de sequências ubiquitina por Southern genómico.

Genótipo: Linha B ("alelo *mt^B*"; ver Capítulo I2, # 2.2.2.3).

SB1917 - Usada para a construção de um banco genómico total.

Genótipo: Linha B; "mating" II; micronúcleo: *Pmr^R*; macronúcleo: *Pmr^S*, 2dgal^R.

SB1918 - Usada como par conjugante na reprodução sexuada.

Genótipo: Linha B; "mating" V; micronúcleo: *Pmr^R*; macronúcleo: *Pmr^S*, 2dgal^R.

B4 - Usada como par conjugante na reprodução sexuada.

Genótipo: Linha B; "mating" ? (III, IV, VI ou VII).

1.1.2. Condições de cultura.

Ambas as espécies de *Tetrahymena* foram cultivadas habitualmente em meio rico PPY [0.75% proteose peptona; 0.75% extracto levedura; 1.5% glucose; 1 mM MgSO_4 ; 50 μM CaCl_2 ; 500 μM KH_2PO_4 ; 100 μM citrato de ferro] às temperaturas óptimas de crescimento em Gallemkamper com agitação moderada (c. 60 rpm/min). No entanto, para certas experiências, como por exemplo a análise de mRNAs, as células foram sujeitas a diversos tratamentos específicos.

Stress hipertérmico - O choque hipertérmico foi aplicado a culturas em crescimento exponencial (cerca de 2×10^5 células/ml) por incubação das células a uma temperatura sub-letal (34°C para *T. pyriformis* e 40°C para *T. thermophila*) durante 15 ou 30 min em banho-maria e com agitação ligeira.

Conjugação - A conjugação de estirpes de *T. thermophila* foi realizada de acordo com o método descrito por Martindale *et al.* (1982).

As estirpes usadas para a conjugação foram previamente cultivadas em meio PPY diluído 1:3 até atingirem uma concentração de cerca de $2-2.5 \times 10^5$ células/ml. De modo a iniciar o período de jejum, que irá induzir as condições celulares apropriadas à conjugação, recolheram-se as culturas por centrifugação a 100g durante 2 min à temperatura ambiente

numa centrífuga de bancada. As células foram então lavadas com 10 mM Tris-HCl pH 7.4, ressuspensas a cerca de $1.5\text{-}2.5 \times 10^5$ células/ml, e deixadas em jejum durante 24h a 30°C. A conjugação foi iniciada pela mistura de igual número de células de cada estirpe e prosseguida com incubação da mistura a 30°C sem agitação. As condições anaeróbicas inibem a conjugação, pelo que a falta de agitação impede o uso de volumes elevados. Em casos onde foi necessário conjugar um elevado número de células, a mistura foi subdividida em pequenos volumes iguais (tipicamente 25-50 ml com um mínimo de 2.5×10^6 células num matrâz de 500 ml) que foram incubados separadamente. Nestas condições, 80-90% das células conjugam ao fim de 3-4 horas, e iniciam a separação ao fim de 10-12 horas. A realimentação, necessária para que as células exconjugantes iniciem a divisão celular originando cada uma dois carionides (ver Capítulo I2, # 2.2.2.1), foi introduzida cerca de 13 horas após o início do contacto por adição de meio volume de PPY.

Os estados citológicos foram seguidos corando-se as células pelo método de Giemsa como descrito por Bruns & Brussard (1981). A cada ponto, cerca de 200µl de células foram centrifugadas a 160g durante 1-2 min, ressuspensas em 200µl de fixador [99:1 de solução de Schaudinn's (solução saturada de HgCl_2 em 2:1 água:etanol) e ácido acético glacial] e incubadas durante 30 min. A suspensão foi então centrifugada e o sedimento ressuspenso em 200µl de 70% etanol, onde pode ser mantido durante alguns dias. Para a coloração, centrifugaram-se 100µl da suspensão, sendo o sedimento ressuspenso no mesmo volume de 3:1 metanol:ácido acético e novamente centrifugado. As células foram ressuspensas no sobrenadante restante, e uma gota deste foi então deixada tombar sobre uma lâmina de microscópio de uma altura de um palmo e seca ao ar. A preparação foi hidrolisada por imersão em 5M HCl durante 1-2 min, lavada com água e seca ao ar. A coloração foi conseguida por imersão da preparação numa solução a 4% de corante de Giemsa em 10 mM tampão fosfato de sódio pH 6.8 durante 10-15 min, após o que foi lavada com água, seca ao ar e montada para visualização ao microscópio.

1.2. Células bacterianas: *Escherichia coli*.

1.2.1. Estirpes usadas.

Como meio de selecção e amplificação de DNA recombinante foram também usadas diversas estirpes da bactéria *Escherichia coli*, escolhidas de acordo com os vectores a usar:

XLI-B - Estirpe usada para selecção de plasmídeos e fagos M13 e amplificação de plasmídeos. Possui um episoma F' modificado que permite a infecção de fagos M13, a selecção azul/branco pelo fenótipo do operão lactose na presença de X-Gal e a resistência ao antibiótico tetraciclina pela presença do transposão Tn10.

Genótipo: *supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1 [lacF' proAB⁺ lacI lacΔ M15 Tn10(tet^R)]* (Bullock *et al.*, 1987)

HB101 - Estirpe usada para a construção de um banco genómico de DNA de *T. pyriformis* em pBR322, de onde se isolaram algumas das sequências usadas neste trabalho (Barahona *et al.*, 1985; Neves *et al.*, 1988).

Genótipo: F⁻ *supE44 hsdS20 [r_B, m_B] recA13 ara-14 proA2 lacY1 galK2 rpsL20 (Sm^r) xyl-5 mtl-1 λ⁻* (Boyer & Roulland-Dussoix, 1969).

JM101 - Estirpe usada inicialmente para a amplificação de fagos M13 e posteriormente substituída pela estirpe TG1, menos recombinante. Permite a selecção azul/branco pelo fenótipo do operão lactose na presença de X-Gal.

Genótipo: *supE thi Δ(lac-pro) F' [traD36 proAB lacI^q ZΔM15]* (Messing, 1983).

TG1 - Estirpe usada para a amplificação de fagos M13. Derivada da estirpe JM101 após modificação do sistema modificador do DNA (*hsdΔ5*).

Genótipo: *supE hsdΔ5 thi Δ(lac-pro) F' [traD36 proAB lacI^q ZΔM15]* (Yanish-Perron *et al.*, 1985).

P2 392 - Estirpe usada para a selecção de fagos λ. Possui o marcador P2 que permite a selecção Spi⁻ para, por exemplo, os fagos λDashII recombinantes usados neste trabalho.

Genótipo: *supE44 supF58 hsdR514 galK2 galT22 metB1 trpR55 lacY1 (P2)*.

LE 392 - Estirpe usada para a amplificação de fagos λ. Idêntica à estirpe P2 392, mas sem o marcador P2.

Genótipo: *supE44 supF58 hsdR514 galK2 galT22 metB1 trpR55 lacY1*.

1.2.2. Meios e condições de cultura.

As células de *E. coli* crescem habitualmente em meio líquido com agitação ou sólido à temperatura de 37°C. Podem ser mantidas em meio sólido a 4°C durante alguns dias, mas a sua conservação por períodos mais longos deve ser feita em meio líquido com 15% de glicerol, a -20°C por algumas semanas ou a -70°C por alguns anos. Os meios de cultura usados dependeram quer da estirpe usada quer do objectivo para o qual foram utilizadas. Os meios habituais de manutenção e crescimento foram:

LB - Usado com as estirpes HB101 e XL1-B, embora seja um meio de crescimento normal que pode ser usado com a quase generalidade das estirpes. Foi suplementado com 25 µg/ml de tetraciclina (**LB+Tc**) para a estirpe XL1-B ou com 100 µg/ml de ampicilina (**LB+Ap**) para a selecção/amplificação de plasmídeos que confirmam resistência a este antibiótico. Como meio sólido foi acrescido de 1.5% de agar-agar (**LA**), eventualmente suplementado com antibióticos ou, no caso da selecção azul/branco, com 5-bromo-4-cloro-3-indolil-β-galactopirósido (X-Gal; 40 µl/placa de uma solução 2% em dimetilformamida) e isopropiltio-β-D-galactosido (IPTG; 40 µl/placa de uma solução 100 mM).

Composição: 10 g/l bactopectona; 5 g/l extracto de levedura; 5 g/l NaCl.

2xTY - Usado com as estirpes JM101 e TG1, para crescimento e propagação de fagos M13.

Composição: 16 g/l bacto-triptona; 10 g/l extracto de levedura; 5 g/l NaCl.

H - Meio sólido usado como base para o plaqueamento de fagos M13 com células XL1-B, normalmente coberto com um meio semi-sólido com apenas 8 g/l de agar-agar (**H-top**) que contém a mistura de células, fagos e meio selectivo azul/branco (X-Gal e IPTG).

Composição: 10 g/l bacto-triptona; 8 g/l NaCl; 12 g/l agar-agar.

TB - Usado com as estirpes LE 392 e P2 392 durante o crescimento e propagação de fagos λ. Neste último caso, deve ser suplementado com 10 mM MgSO₄ e 0.2% maltose (**TB supl.**) para indução dos receptores do fago.

Composição: 10 g/l bacto-triptona; 5 g/l NaCl.

NZY - Usado com as estirpes LE 392 e P2 392 para crescimento e propagação de fagos λ . Neste último caso, deve ser suplementado com 5 mM CaCl_2 (**NZY supl.**). Como meio sólido é acrescido de 1.5% agar-agar (**NZY plates**), que é normalmente coberto, na preparação de placas fágicas λ , de um meio semi-sólido com 0.7% agarose (**Top NZY**) que contém a mistura de células e fagos.

Composição: 10 g/l hidrolisado de caseína; 5 g/l extracto de levedura; 5 g/l NaCl; 2g/l MgSO_4 .

2. Vectores de clonagem.

2.1. Vectores plasmídicos.

Como vectores plasmídicos foram usados os plasmídeos **pBR322**, **pUC19** e **pBluescript II**.

O **pBR322** é um plasmídeo de baixo número de cópias com 4363 pb (Bolivar *et al.*, 1977) que possui genes que conferem resistência à ampicilina e à tetraciclina. Foi usado para a construção de um banco genómico de DNA de *T.pyriformis* de onde foram isolados alguns dos clones usados neste trabalho (Barahona *et al.*, 1985; Neves *et al.*, 1988).

O **pUC19** é um plasmídeo multicópia de 2686 pb derivado do pBR322 (Yanish-Perron *et al.*, 1985). Este plasmídeo possui apenas o gene que confere resistência à ampicilina, e ainda um local de clonagem com 13 locais únicos de reconhecimento de endonucleases, denominado "polilinker". A ordem das enzimas caracteriza os plasmídeos pUC18 e pUC19. Esta região encontra-se inserida num gene *lacZ* de *E.coli*, o que permite a selecção dos clones positivos por α -complementação com células hospedeiras que não possuam o enzima β -galactosidase funcional. A presença desta enzima pode ser detectada com o uso de X-Gal, que é um reagente incolor mas cujo produto após hidrólise enzimática é azul, sendo também necessário a presença de IPTG para indução do operão da lactose. Quando um fragmento se insere no "polilinker", o gene *lacZ* plasmídico é inactivado e as colónias resultantes são brancas em vez de azuis, constituindo a selecção azul/branca. Foi usado como vector de clonagem e subclonagem de fragmentos, e ainda para a construção de um banco parcial de DNA de *T.pyriformis* de onde foi isolado um dos clones usados neste trabalho (Neves *et al.*, 1991b).

O **pBluescript II** é um fagemídeo multicópia de 2961 pb derivado do pUC19 (Stratagene). Este plasmídeo confere resistência à ampicilina, possui uma origem fágica f1 que lhe permite ser recuperado em cadeia simples com coinfeccção por um fago M13 "helper", e 21 locais únicos de reconhecimento por endonucleases. A ordem das enzimas caracteriza os plasmídeos SK(+/-) e KS(+/-), e a orientação da origem fágica f1 caracteriza a cadeia senso (+) ou antisense (-) recuperada em cadeia simples. Tal como o pUC19 este fagemídeo permite uma selecção azul/branca por α -complementação.

2.2. Vectores fágicos: o fago M13 e seus derivados.

O bacteriófago M13 é um fago filamentosos que infecta células *E.coli* que possuam o epissoma F sem produzir lise celular. É constituído por DNA circular de cadeia simples

com cerca de 7.2 kb, mas a sua forma replicativa intracelular possui cadeia dupla. Os seus dois principais derivados, M13mp18 e M13mp19, possuem uma região "polilinker" equivalente à dos plasmídeos pUC18 e pUC19 que permite igualmente uma selecção azul/branca por α -complementação (Messing, 1983). A forma replicativa de cadeia dupla é usada como vector de clonagem e a forma vegetativa de cadeia simples é usada para sequenciação.

2.3. Vectores fágicos: o fago λ DashII.

O fago λ DashII (Stratagene) é um derivado do bacteriófago λ dito de substituição por permitir a substituição um fragmento central não essencial de 14kb por DNA heterólogo de 9 a 23 kb. As extremidades desse fragmento possuem vários locais únicos de reconhecimento por endonucleases.

3. Preparação e análise de DNA de *Tetrahymena*.

3.1. Preparação do DNA nuclear.

A extracção do DNA nuclear de *Tetrahymena* foi realizada de acordo com o método de Karrer & Gall (1976) adaptado por Barahona *et al.* (1985).

As células de uma cultura (2x 500 ml) de *Tetrahymena* no início da fase estacionária (cerca de 5×10^5 cel/ml) foram recolhidas por centrifugação a 4000g durante 10 min à temperatura de 4°C. As células foram ressuspensas em 20 ml (volume final) de água bidestilada estéril, e à suspensão foram adicionados 36 ml de Solução I [500 mM EDTA; 10 mM Tris-HCl pH 9.5; 1% SDS]. A mistura foi incubada a 50°C durante 20 min, sem agitação, de forma a disromper as membranas celulares e inactivar as DNases libertadas. Seguidamente adicionaram-se 20 ml de solução II [500 mM EDTA; 10 mM Tris-HCl pH 9.5] e 15.2 mg de Proteinase K, prolongando-se a incubação por mais 3 horas com agitação ocasional. Realizaram-se então duas extracções, a primeira com 76 ml de uma mistura de fenol, clorofórmio e álcool isoamílico (50:50:1) e a segunda com 76 ml de uma mistura de clorofórmio e álcool isoamílico (25:1). Os ácidos nucleicos foram precipitados com igual volume de isopropanol na presença de 0.3 M de acetato de sódio durante cerca de 30 min a -20°C.

O DNA obtido foi seguidamente purificado por gradiente de CsCl (Sambrook *et al.*, 1989). Os ácidos nucleicos foram suavemente ressuspensos em água bidestilada estéril até a solução final atingir 7.15g, e à qual se adicionou seguidamente 7.5g de CsCl e 350 μ l de uma solução de 10 mg/ml brometo de etídio. Os gradientes foram centrifugados a 70000 rpm no rotor Ti80 (Beckman) durante 16-18h a 18°C. A banda de DNA foi recolhida com uma seringa e o brometo de etídio foi removido com várias extracções de 1-butanol e/ou 2-butanol. A solução foi seguidamente dialisada contra TE [10 mM Tris-HCl pH 7.5; 1 mM EDTA] a 4°C durante a noite para remoção do CsCl. A concentração final do DNA foi determinada por espectrofotometria a 260 nm, sendo que 1 O.D. $\Leftrightarrow$ 40 ng/ml.

3.2. Análise do DNA nuclear (Southern).

3.2.1. Hidrólise e separação do DNA por electroforese em geis de agarose.

Aliquotas com cerca de 5 a 10 µg de DNA nuclear foram hidrolisadas com 1-3 U/µg das enzimas de restrição apropriadas nas condições adequadas durante cerca de 5 horas. Após este tempo, adicionou-se-lhes 1/5 do seu volume de tampão corante de amostras de DNA ["5x loading buffer": 0.1 M EDTA pH 8.0; 80% glicerol; 0.2% SDS; 0.2% azul de bromofenol].

A separação dos fragmentos de DNA foi feita por electroforese em geis de agarose em tampão TAE [0.4 M Tris-HCl pH 7.9; 1 mM acetato de sódio; 1 mM EDTA] contendo 0.2 mg/ml de brometo de etídio. Os geis foram preparados por dissolução no mesmo tampão de 0.8 a 1.2% de agarose (Sigma tipo II), consoante o tamanho dos fragmentos a separar. A electroforese decorreu em aparelhos horizontais de 25 cm (Hooper) ou 10 cm de comprimento ("minigeis"; BRL) a uma diferença de potencial de 50V durante 14-16 horas em geis de 25 cm ou durante 2-3 horas nos de 10 cm. Como marcadores de massa molecular usaram-se diversas misturas comerciais, como λ BstEII (New England Biolabs), "1kb DNA Ladder" (BRL) ou λ HindIII (Stratagene). No final de cada electroforese, o DNA foi visualizado usando luz ultravioleta transmitida a 302 nm.

3.2.2. Imobilização do DNA.

Após separação por electroforese, os fragmentos de DNA foram transferidos para membranas de nitrocelulose ($\varnothing$ 0.45 µm; Schleicher & Schuell, BA 85) ou de nylon (Hybond N+; Amersham) pelo método de Southern (1975).

Os geis de agarose foram primeiramente submersos numa solução 0.5 M NaOH e incubados durante 30 min com agitação para desnaturação dos fragmentos de DNA. De seguida o gel foi lavado três vezes com água destilada e neutralizado em 1 M Tris-HCl pH 7.5, 10x SSC [SSC: 150 mM NaCl; 15 mM acetato de sódio] durante 30 min com agitação. O sistema de transferência foi então montado. Este sistema consiste, resumidamente, numa tina onde se coloca a solução de transferência (20x SSC), e numa placa de vidro sobre a qual está colocada uma "ponte" de papel 3MM cujas pontas estão imersas na solução. Sobre esta "ponte" é colocado o gel, e sobre este a membrana (previamente humedecida em água e 2x SSC), 2-3 folhas de papel 3MM, uma camada (5-10 cm) de papel absorvente e um peso (0.5-1 kg). A transferência de DNA para a membrana dá-se passivamente por capilaridade de 20x SSC através do gel de agarose para a membrana. No final da transferência, que decorre durante 14-18 horas, a membrana é lavada suavemente com 2x SSC e seca. A imobilização é finalmente conseguida por incubação da membrana em estufa a 80°C durante duas horas ou, alternativamente para membranas de nylon, por exposição a luz ultravioleta de 302 nm durante alguns minutos.

No caso de filtros de nylon realizaram-se também transferências alcalinas. Nestas condições, o gel é desnaturado em 0.4 M NaOH durante cerca de 20 min com agitação, não há uma fase de neutralização e a transferência é realizada como indicado mas usando 0.4 M NaOH e não 20x SSC. A membrana não necessita de ser previamente humedecida, podendo ser colocada directamente sobre o gel. Após a transferência, o tratamento dado à membrana é idêntico.

3.2.3. Preparação de sondas radioactivas.

Como sondas para análise do DNA genómico de *Tetrahymena* usaram-se três tipos de DNA (Tabela M1):

Cadeia dupla - Fragmentos de DNA de clones ubiquitina isolados como indicado em # 6.1. A marcação destes fragmentos foi realizada com $[\alpha\text{-}^{32}\text{P}]\text{dATP}$ ou $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ (3000 Ci/mmol) usando-se o *kit* "Multiprime" (Amersham) segundo as indicações do fabricante. Genericamente, cerca de 25-50 ng de DNA em 25 μl de água bidestilada são desnaturados por fervura durante 2 min. Adicionam-se-lhe depois 5 μl tampão, uma mistura (3x 4 μl) dos três desoxiribonucleótidos não marcados, 5 μl de uma mistura de hexaoligonucleótidos que irão funcionar como iniciadores aleatórios da reacção de polimerização, 2 μl do nucleótido marcado e 2 μl (2U) de "Klenow". A polimerização é efectuada por incubação da mistura à temperatura ambiente durante 3 horas ou a 37°C durante 30 min, e parada por fervura durante dois minutos.

Cadeia simples - Utilização de um fragmento de DNA inserido no fago M13. O processo de marcação é equivalente ao anterior, mas utiliza-se como iniciador um oligómero específico da região "polilinker" ("40"). Este é incubado com o DNA de forma semelhante à usada em reacções de sequenciação (# 7.4), sendo a polimerização efectuada como indicado anteriormente.

Tabela M1. Sondas ubiquitina usadas para análise do DNA de *Tetrahymena*

Designação	Nome	Tipo	Características	Temperatura de hibridação
Ubiquitina	Ub, GpI	fgm ds	<i>HindIII-HindIII</i> de 0.23 kb	55°C ou 65°C
Grupo II	GpII	fgm ds	<i>EcoRI-EcoRI</i> de 0.23 kb	65°C
Gene A	A-5'	ss	-422 (<i>HindIII</i>) a -27	55°C
Gene B	Oli-B3'	oligómero	+91(GAGTAGATACAATG)+78	28°C
	B-3'	fgm ds	<i>HindIII-BamHI</i> de 0.5 kb	65°C
Gene C	Oli-C5'	oligómero	-3(GTTGTTTGGTAAATCG)-18	40°C
	C-5'	fgm ds	<i>BamHI-HindIII</i> de 0.5 kb	55°C
Gene D	D-3'	fgm ds	<i>HincII-ClaI</i> de 1.2 kb	55°C
Gene E	Oli-E5'	oligómero	-26(GAGTTAAGTTGTTGTCTTTC)-46	45°C
Gene F	F-5'	ss	-412 (<i>BglII</i>) a -26	55°C
Gene H	H-3'	fgm ds	<i>BglII-HindIII</i> de 0.7 kb	55°C
Gene X	X-5'	fgm ds	<i>HindII-ClaI</i> de 2.0 kb	55°C
Gene Y	Y-3'	fgm ds	<i>HindIII-HindIII</i> de 0.8 kb	65°C

Oligómeros - A marcação de sondas oligonucleotídicas foi realizada com [γ - ^{32}P]ATP (3000 Ci/mmol). Cerca de 100 ng de DNA são marcados na presença de 7U da enzima T4 polinucleótido cinase e 50 μCi do ribonucleótido marcado, em 10 ou 25 μl de uma solução tampão específica [50 mM Tris-HCl pH 7.6; 10 mM 2-mercaptoetanol; 10 mM MgCl_2], durante 1 hora a 37°C. A reacção é parada por inactivação da enzima a 65°C durante 10 min.

Independentemente do método usado para marcação das sondas, a actividade específica destas foi determinada por contagem de cintilações. Cerca de 5 μl de uma diluição 1:500 da sonda em 200 mM EDTA foi precipitada com 3-5 ml de 10% TCA na presença de 200 μl de 200 mM EDTA e 50 μl de DNA de esperma de salmão 0.5 $\mu\text{g}/\mu\text{l}$ durante 10 min a 4°C. O precipitado foi recolhido por filtração sob vácuo em filtros GF-C (Whatman) e lavado várias vezes com 10% TCA e uma vez com etanol/éter etílico (50:50). Depois de secos, os filtros foram colocados em frascos de contagem de cintilações contendo 5 ml de "cocktail" de cintilação [100 mg/l 2,2'-p-fenileno-bis-5-feniloxazole; 4 g/l difeniloxazole em tolueno/Triton X100 (2:1)] e a medição da radioactividade foi efectuada em contadores de radiação β (Beckman LS100C ou Beckman LS8100).

3.2.4. Condições de hibridação dos filtros.

3.2.4.1. Hibridação com sondas não oligoméricas.

Os filtros contendo DNA genómico de *Tetrahymena* digeridos com as enzimas de restrição apropriadas foram pré-hibridados por imersão num tampão 3x/1x [3x SSC; 1x Denhardt (0.02% Ficoll 400; 0.02% polivinilpirrolidona); 0.1% SDS; 20 mM fosfato de sódio pH 6.5; 100 μg DNA de esperma de salmão sonificado) e incubados durante 2-16 horas num banho com agitação suave a 55°C.

A hibridação decorreu no mesmo tampão (renovado) contendo cerca de 10^6 cpm da sonda apropriada durante 14-18 horas à mesma temperatura. Os filtros foram seguidamente lavados 2x30 min numa solução 2x SSC também a 55°C e 2x15 min numa solução 0.2x SSC, 0.2% SDS à temperatura ambiente. Depois de secas, as membranas foram colocadas em contacto com películas X-OMAT AR5 (Kodak) entre écrans intensificadores de sinal em cassetes próprias, que foram armazenadas a -70°C durante o número de horas considerado necessário à impressão. A revelação foi efectuada nas condições indicadas pela Kodak.

Certas experiências exigiram condições de hibridação mais restritivas do que as indicadas. Estas foram conseguidas hibridando-se os filtros à temperatura de 65°C e aumentando-se o número e tempo das lavagens para 4x30 min com a primeira solução, à mesma temperatura, e 4x15 min com a segunda, à temperatura ambiente.

3.2.4.2. Hibridação com sondas oligoméricas.

A hibridação com sondas oligoméricas foi realizada segundo as condições descritas por Zeff & Geliebter (1987). Os filtros foram pré-hibridados numa solução de pré-hibridação de oligómeros (TPO) [5x SSC; 10x Denhardt; 7% SDS; 20 mM fosfato de sódio pH 7.0; 100 μg DNA de esperma de salmão sonificado] durante 4-16 horas à

temperatura apropriada a cada oligómero (Tabela M1). A hibridação decorreu à mesma temperatura no tampão de hibridação de oligómeros (TO) [TPO; 10% sulfato de dextran] contendo 1-5 ng/ml do oligómero marcado radioativamente. As membranas foram seguidamente lavadas, à temperatura de hibridação, durante 30 min na primeira solução de lavagem [3x SSC; 10x Denhardt; 5% SDS; 10 mM fosfato de sódio pH 7.0] e durante 30 min na segunda solução de lavagem [1x SSC; 1% SDS]. Os filtros foram colocados em contacto com películas X-OMAT AR5 (Kodak) de forma idêntica à indicada anteriormente.

4. Preparação e análise do RNA de *Tetrahymena*.

4.1. Extracção de RNA.

4.1.1. Extracção de RNA citoplasmático.

A extracção de RNA total citoplasmático de *Tetrahymena* foi realizada em pequena escala de acordo com o método descrito por Soares *et al.* (1991).

As células de uma cultura em fase exponencial (cerca de 2×10^5 cel/ml) foram divididas em aliquotas com cerca de 2×10^6 células e, quando necessário, submetidas a alguns tratamentos antes da extracção do RNA (p.ex., stress hipertérmico). Para a extracção, as células foram recolhidas por centrifugação a 6000g durante 2 min a 4°C, lavadas em tampão Pi frio [2.5 M $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 6.8; 1 mM MgCl_2 ; 50 mM NaCl] e ressuspensas em 500µl de tampão Martis de abertura frio [100 mM Tris-HCl pH 8.5; 100 mM β -mercaptoetanol; 1 mM EDTA; 50 µM ácido aurintricarboxílico]. A lise foi efectuada por adição de 25µl de 10% Nonidet NP-40 frio com agitação suave com uma vareta, sendo o lisado obtido centrifugado a 15 000g durante 10 min a 4°C para eliminação dos resíduos celulares e do DNA. Ao sobrenadante foram então adicionados 5µl de 20% SDS, realizando-se então três extracções à temperatura ambiente, as duas primeiras com 500µl de uma mistura de fenol, clorofórmio e álcool isoamílico (50:50:1) e a terceira com o mesmo volume de uma mistura de clorofórmio e álcool isoamílico (24:1). O RNA foi precipitado durante a noite a -20°C por adição de 2.5 volumes de etanol absoluto na presença de 0.3 M de acetato de sódio, sedimentado por centrifugação a 10 000g durante 30 min a 4°C, e ressuspenso em água bidestilada estéril previamente tratada com 0.1% de dietilpirocarbonato (DEPC). A concentração final do RNA foi determinada por espectrofotometria a 260 nm, sendo que 1 O.D. $\Leftrightarrow$ 20 ng/ml.

4.1.2. Isolamento de mRNA poli(A)+.

Para o isolamento de mRNA poli(A)+ são necessárias grandes quantidades de RNA citoplasmático total, que é obtido de forma semelhante à descrita anteriormente (# 4.1.1) a partir de culturas de 1 litro de células a $1.5\text{-}2 \times 10^5$ cel/ml. A cultura é inicialmente diluída em tampão Pi frio (600 ml/litro), e as células recolhidas ressuspensas em 40 ml de tampão Martis, sendo a lise efectuada por adição de 4 ml de 10% Nonidet NP-40. O sobrenadante obtido após a centrifugação de 20 min a 15 000 rpm é transferido para um matrâz contendo 4 ml de 1 M Tris-HCl pH 9.5 e 400µl de 20% SDS, e as extracções efectuadas com 44 ml das misturas de fenol e clorofórmio descritas. O RNA é então precipitado com etanol

absoluto e acetato de sódio, e ressuspensão em cerca de 4 ml (20 O.D./ml) de tampão de retenção [20 mM Tris-HCl pH 7.5; 1 mM EDTA; 0.1% SDS; 0.5 M NaCl].

A separação dos mRNAs poli(A)⁺ é efectuada por cromatografia de afinidade em coluna de oligo(dT)celulose. A coluna é habitualmente montada em pipetas Pasteur vedadas numa das extremidades com um pouco de fibra de vidro. A oligo(dT)celulose é previamente ressuspensa em tampão de retenção (cerca de 0.5g para 100 O.D. de RNA total), empacotada na pipeta e lavada com duas ou três passagens do mesmo tampão, procurando-se simultaneamente conseguir-se um fluxo controlado de cerca de 0.4 ml/min. O RNA total é então passado pela coluna, recolhendo-se o eluido contendo o RNA poli(A)⁻. A coluna é então lavada com tampão de retenção até à remoção total do RNA poli(A)⁻ (cerca de 15 ml, até obter-se uma absorvência do eluido inferior a 0.05), sendo os mRNA poli(A)⁺ eluidos seguidamente com cerca de 5 ml de tampão de libertação [20 mM Tris-HCl pH 7.5; 1 mM EDTA; 0.1% SDS]. Os RNAs obtidos são finalmente precipitados por adição de 2.5 volumes de etanol absoluto na presença de 0.3 M de acetato de sódio, como descrito anteriormente.

4.2. Análise do RNA citoplasmático (Northern).

4.2.1. Separação do RNA por electroforese em geis de agarose.

A separação dos RNAs totais ou poli(A)⁺ foi feita por electroforese em geis de agarose contendo formaldeído. Os geis foram preparados por dissolução em 156 ml de água bidestilada de 5g de agarose (Sigma tipo II) à qual, após arrefecimento a 60°C, se adicionou 50 ml de 5x RB [1x RB ("Running Buffer"): 40 mM MOPS pH 7.0; 10 mM acetato de sódio; 1 mM EDTA] e 44 ml de 37% formaldeído. Obteve-se assim 250 ml de solução de agarose 2% que foi solidificada em tabuleiros de 20x25 cm (Hofer).

A preparação das amostras envolveu a adição a cada 4.5µl de solução de RNA de 2µl de 5x RB, 3.5µl de 37% formaldeído, 10µl de formamida desionizada e 2µl de mistura corada para RNA [50% glicerol; 1 mM EDTA; 0.04% azul de bromofenol; 0.04% verde de xileno-cianol]. As amostras são então desnaturadas a 55°C durante 15 min, e aplicadas de imediato no gel. A electroforese decorreu em RB a uma diferença de potencial de 35V durante 14-16 horas, usando-se como marcadores de massa molecular RNA poli(A)⁻ de *Tetrahymena*, ricos em RNA ribossomais de massa molecular conhecida ou alternativamente uma mistura comercial de RNAs, "RNA ladder" (BRL). Após a electroforese, as faixas correspondentes aos marcadores foram removidas do gel e coradas numa solução a 1 µg/ml de brometo de etídio durante 1 hora, eliminando-se o excesso de corante com lavagens com água, de modo a poderem-se visualizar os marcadores com luz ultravioleta transmitida a 302 nm.

4.2.2. Imobilização do RNA.

Após a migração, os geis de RNA foram lavados abundantemente com água bidestilada, de modo a remover-se o excesso de formaldeído, e incubados durante 1 hora em 20x SSC. A transferência decorreu para filtros de nitrocelulose de modo semelhante ao descrito para os DNAs (# 3.2.2).

4.2.3. Condições de hibridação dos filtros.

As condições de hibridação dos filtros contendo RNA de *Tetrahymena*, bem como a preparação das sondas usadas, são idênticas às descritas para os DNAs (# 3.2.3 e # 3.2.4).

4.3. Análise da transcrição aparente ("run on")

4.3.1. Preparação de núcleos de *Tetrahymena*.

Os núcleos de *Tetrahymena* foram obtidos a partir de culturas de 500 ml de células em crescimento exponencial ou submetidas a alguns tratamentos (p.ex., stress hipertérmico). As células foram recolhidas por centrifugação a 4 000g durante 5 min, lavadas duas vezes com tampão RSB [10 mM Tris-HCl pH 7.4; 1 mM NaCl; 3 mM MgCl₂; 8.5% sacarose] e ressuspensas num volume final de 5 ml em RSB. A lise foi efectuada por adição de 250µl de 10% Nonidet NP-40, e ao lisado foi adicionado tampão RSB até um volume final de 20 ml. Os núcleos foram sedimentados por centrifugação a 1500 rpm durante 3 min a 4°C no rotor SS-34 (Sorvall), e lavados duas vezes em tampão RSB (10 ml + 5 ml), sendo a centrifugação final efectuada a 2500 rpm. Os núcleos foram então ressuspensos em 1 ml de tampão de armazenamento [50 mM Tris-HCl pH 8.3; 5 mM MgCl₂; 0.1 mM EDTA; 40% glicerol], recentrifugados a 4000 rpm durante 3 min e ressuspensos em 100µl do mesmo tampão. Após a adição de 2µl de 1 M DTT, os núcleos foram rapidamente congelados em azoto líquido e armazenados a -70°C, tendo-se previamente retirado pequenas aliquotas que foram usadas para contagem dos núcleos. Estes foram corados com verde de metilo de pirenina e contados em hematocimetro num microscópio óptico.

Tabela M2. Fragmentos de DNA usados em ensaios de "run-on"

Designação	Nome	Localização
Gene A	A-5'	-422 a -27
	A-cod	-422 a 509
Gene F	F-5'	-402 a -26
	F-cod	-402 a 560
Gene G	G-5'	-241 a -22
	G-cod	-22 a 492
Gene H	H-cod	-264 a 314

4.3.2. Imobilização das sondas de DNA.

Por cada filtro usado em experiências de "run-on" foi preparado 1 µg de cada fragmento de DNA a usar como sonda (Tabela M2; # 6.1). A cada solução de DNA foram adicionados 100 µl de 1 M acetato de amônio, e o DNA foi desnaturado durante 10 min na presença de NaOH (0.35 M final). A solução foi neutralizada por adição de 1 volume de 2 M acetato de amônio, e aliquotas contendo 1 µg de fragmento foram então transferidas para filtros de nitrocelulose, previamente embebidos em 1 M acetato de amônio, por filtração sob vácuo num aparelho Minifold (Schleicher & Schull). Os filtros foram rapidamente lavados em 1 M acetato de sódio, secos, e incubados a 80°C durante 2 horas para fixação do DNA.

4.3.3. Transcrição nuclear e isolamento de RNA.

As experiências de "run-on" foram realizadas essencialmente como descritas por Groudine *et al.* (1981). Para cada ensaio foram usados 10^7 núcleos, eventualmente diluídos com tampão de armazenamento até um volume final de 100 µl. A cada suspensão foram adicionados 100 µl de mistura de transcrição [10 mM Tris-HCl pH 8.0; 10 mM $MnCl_2$; 250 mM $MgCl_2$; 250 mM KCl; 5 mM ATP; 2.5 mM CTP; 2.5 mM GTP; 100 µg creatina fosfoquinase; 20 mM fosfocreatina; 0.1 U/µl RNasina; 100 µCi [α - ^3P]UTP (800 Ci/mmol, Amersham)], tendo a mistura sido incubada a 28°C durante 30 min. Como controlo, foi realizado um ensaio com núcleos preparados de células em fase exponencial aos quais se adicionou, em simultâneo com a mistura de transcrição, 20 µl de 100 µg/µl α -amanitina. Os núcleos foram então sedimentados por centrifugação a 2500 rpm durante 4 min, e lisados por adição de 200 µl de um tampão de elevada concentração iónica [12 mM Tris-HCl pH 7.4; 0.6 M NaCl; 60 mM $MgCl_2$] e incubados durante 15 min à temperatura ambiente com 20U de DNase e 20U de RNasina, sendo a reacção parada pela adição de 5 µl de 0.5 M EDTA e 10 µl de 10% SDS. Realizaram-se então duas extracções, uma por adição de 250 µl de uma mistura de fenol, clorofórmio e álcool isoamílico (50:50:1) e outra com igual volume de uma mistura de clorofórmio e álcool isoamílico (42:1). Os RNAs foram precipitados com 2 volumes de etanol absoluto na presença de 0.25 M acetato de amônio, e ressuspensos em 400 µl de tampão de hibridação [3x SSC; 5x Denhardt; 50% formamida; 1% SDS; 20 mM fosfato de sódio pH 7.0; 0.02% albumina bovina; 1 mM pirofosfato de sódio pH 7.0; 165 µg/ml tRNA de levedura], sendo a incorporação determinada como descrito em # 3.2.3.

4.3.4. Condições de hibridação dos filtros.

Os filtros preparados em # 4.3.2 foram pré-hibridados a 42°C durante duas horas em tampão de hibridação (# 4.3.3) e hibridados com os transcritos nucleares durante 4 dias à mesma temperatura no mesmo tampão. Os filtros foram seguidamente lavados durante 30 min numa solução 2x SSC a 50°C e 30 min numa solução 2x SSC, 1 µg/ml RNase A (Sigma) a 37°C. Depois de secos, os filtros foram colocados em contacto com películas X-OMAT AR5 (Kodak) como descrito em # 3.2.4.1.

5. Construção de bancos de DNA de *Tetrahymena*.

5.1. Construção de bancos de DNA de *Tetrahymena pyriformis* em plasmídeos.

5.1.1. Banco total de fragmentos *Hind*III em pBR322.

Algumas das sequências de DNA macronuclear codificantes da ubiquitina foram isoladas de um banco genómico em pBR322 de DNA de *T.pyriformis* digerido com *Hind*III preexistente (Barahona *et al.*, 1985). A selecção dos clones positivos está descrita em Neves *et al.* (1988).

5.1.2. Banco parcial de fragmentos *Cla*I em pUC19.

O banco parcial de DNA genómico de *T.pyriformis* foi construído por inserção em pUC19 de fragmentos entre 1 e 3 kb recuperados por electroeluição a partir DNA digerido com a enzima *Cla*I e migrado em gel de agarose. A construção do banco e a selecção dos clones positivos foi descrita em Neves *et al.* (1991b).

5.2. Construção de bancos totais em λ DashII.

5.2.1. Bancos de DNA total de *Tetrahymena* parcialmente digerido com *Sau*3AI.

Foram construídos dois bancos de DNA de *Tetrahymena* parcialmente digerido com *Sau*3AI em λ DashII (Stratagene), um com DNA macronuclear de *T.pyriformis* CGL e outro com DNA total de *T.thermophila* SB1917.

A concentração de enzima necessária para produzir fragmentos entre 9 a 15 kb, tamanho ideal para a inserção no vector de substituição λ DashII, foi determinada previamente à realização dos bancos. Cerca de 100 μ l de solução contendo 10 μ g de DNA de *Tetrahymena* e 10 μ l de tampão de reacção da enzima *Sau*3AI (Boehringer) foram distribuídos por 8 "eppendorfs", contendo um 20 μ l e os restantes 10 μ l da solução. Os tubos foram colocados a 4°C, e ao tubo contendo 20 μ l foram adicionados 0.3 U de *Sau*3AI. Deste tubo transferiram-se 10 μ l ao tubo seguinte, e assim sucessivamente até ao penúltimo tubo, sendo os 10 μ l finais descartados. O último tubo foi usado como controlo sem enzima. Os tubos foram então incubados 30 min a 37°C, sendo seguidamente transferidos para 4°C e a reacção parada por adição de tampão corante de amostras de DNA. As restrições foram então analisadas em minigeis de agarose, e a concentração enzimática ideal à produção dos fragmentos desejados nestas condições foi determinada a 0.019 U/ μ g para o DNA de *T.pyriformis* e a 0.012 U/ μ g para o DNA de *T.thermophila*.

Para a preparação dos fragmentos a inserir, foram então digeridos cerca de 100 μ g de DNA de *Tetrahymena* num volume final de 1 ml com as unidades de enzima estipuladas. Aos 30 min adicionou-se a um terço do volume igual volume de fenol saturado com TE, de modo a extrair a enzima e terminar a reacção. O segundo terço foi tratado de igual modo aos 40 min, e o volume restante aos 45 min. As três fases aquosas foram então reunidas, e o DNA foi precipitado com etanol absoluto na presença de acetato de amónio e ressuspenso num volume final de 450 μ l. A separação dos fragmentos com tamanho entre 9 e 15 kb foi realizada em gradiente de sacarose. O gradiente foi preparado por um aparelho de gradiente

a partir de duas soluções de sacarose, respectivamente a 10% e 40%, em 10 mM Tris-HCl pH 7.5, 1 M NaCl e 5 mM EDTA. A solução de DNA foi suavemente depositada sobre o gradiente, e o gradiente formou-se por centrifugação em rotor SW41 (Beckman) a 27 000 rpm durante 22 horas a 20°C. O gradiente foi fraccionado por bomba peristáltica em aliquotas de 250 a 300 µl, sendo analisados 15 µl em gel de agarose 0.5%. As aliquotas contendo fragmentos de DNA compreendidos entre 9 a 23 kb foram reunidas, dialisadas contra TE durante 14-18 horas a 4°C, e precipitadas com etanol absoluto na presença de acetato de sódio, sendo o sedimento de DNA ressuspensão num volume final de 30 µl de TE. A integridade das extremidades dos fragmentos foi então testada por adição de 2U Weiss (120 U Biolabs) de T4 DNA Ligase a cerca de 1 µl (600 ng) de fragmentos para um volume final de ligação de 10 µl (# 6.2), sendo a mistura incubada a 4°C durante 14-18 horas. A ligação final ao vector λDashII foi feita com 300 ng de DNA e 500 ng de vector (relação 2:1) nas mesmas condições.

A encapsidação foi realizada usando-se o *kit* Gigapack II Gold (Stratagene) segundo as instruções do fabricante. Foram usadas metade das quantidades indicadas, e após as 2 horas de incubação a 20-25°C, foram adicionados 250 µl de tampão SM [50 mM Tris-HCl pH 7.5; 1 M NaCl; 10 mM MgSO₄; 0.01% gelatina] e 20 µl de clorofórmio. O banco foi então titulado, sendo considerado representativo a 99% com um título acima de 5x10⁴ pfu, e guardado a -70°C na presença de 7% de DMSO.

Um banco representativo de cada espécie de *Tetrahymena* foi igualmente amplificado e guardado. Para a amplificação, cada banco foi totalmente plaqueado em placas de Petri grandes (Ø=15 cm) como descrito a seguir, contendo cada placa acima de 2.5x10⁴ pfu. Após o crescimento, cada placa foi incubada durante 14-18 horas a 4°C com 10 ml de SM. O líquido foi então recolhido, e as placas lavadas com mais 2 ml de SM. A suspensão final de fagos foi incubada durante 15 min a 20-25°C com 5% de clorofórmio, e os restos celulares foram removidos por centrifugação a 2 000g durante 5 min. O sobrenadante foi armazenado a 4°C com 0.3% de clorofórmio, sendo algumas aliquotas guardadas a -70°C com 7% DMSO.

5.2.2. Selecção dos clones recombinantes.

A selecção de clones recombinantes contendo sequências codificantes para a ubiquitina foi feita por hibridação *in situ* em filtros de nitrocelulose para onde foram transferidas as placas fágicas λ.

A primeira pesquisa foi realizada usando-se com estirpe bacteriana células P2, selectivas para fagos com fragmento inserido. As células de uma cultura inoculada no dia anterior em TB suplementado foram recolhidas a 2 000 rpm durante 10 min numa centrífuga clínica Sigma, e ressuspensas em metade do volume original em 10 mM MgSO₄. Três aliquotas de 1.5-2x10⁴ pfu do banco, de modo a constituir-se um banco representativo a 99%, foram incubadas com 600 µl de células P2 durante 15 min a 37°C. Após a adição de 10 ml de Top NZY mantido a 50-60°C, as bactérias infectadas foram plaqueadas em placas NZY (Ø=15 cm) e incubadas 14-18 horas a 37°C.

Antes da transferência para filtros de nitrocelulose, as placas foram secas e armazenadas durante 30-60 min a 4°C, de modo a solidificar devidamente a agarose do Top NZY. Os filtros de nitrocelulose (Ø=0.45 µm; Schleicher & Schuell) foram então postos em contacto com as placas (dois por placa: 30 seg para o primeiro, 2 min para a réplica),

sendo de imediato tratados de modo a fixar o DNA fágico à membrana. Assim, os filtros foram dispostos sucessivamente em papel 3MM (Whatmann) embebidos durante 1 a 5 min nas seguintes soluções: de lise [0.5 M NaOH; 1.5 M NaCl]; de neutralização [1 M Tris-HCl pH 7.5; 1.5 M NaCl]; de lavagem [2x SSC]. Entre cada solução, os filtros foram colocados em papel de filtro para remover o excesso de líquido. Após o tratamento, os filtros foram incubados a 80°C durante 2 horas para fixar o DNA.

A pesquisa de clones contendo sequências codificantes para a ubiquitina foi feita por hibridação dos filtros com o fragmento *Hind*III de 0.23 kb do clone pTU20 marcada como descrito em # 3.2.3. Os filtros foram pré-hibridados a 55°C durante 14-18 horas num tampão 6x/10x [6x SSC; 10x Denhardt; 100 µg/ml DNA esperma de salmão; 0.1% SDS; 20 mM tampão fosfato pH 6.5], e hibridados nas mesmas condições na presença de um mínimo de 10⁶ cpm de sonda. As lavagens foram feitas em 2x SSC durante 10 min, sendo a primeira a 55°C e a segunda a 60°C. Depois de secas, as membranas foram colocadas em contacto com películas X-OMAT AR5 (Kodak) como descrito em # 3.2.4.1.

A zona de Top NZY que continha os fagos que impressionaram a chapa foi então recolhida para um "eppendorf" contendo 200µl de tampão SM e incubada a 20-25°C durante 2 a 4 horas com agitação, ou a 4°C durante 14-18 horas. As pesquisas seguintes decorreram como descrito, usando-se habitualmente uma diluição 1:10 da suspensão para infectar 200µl de células hospedeiras LE, sendo plaqueadas em placas de Petri de Ø=8 cm. O fago de interesse foi considerado purificado quando a zona recolhida continha apenas uma placa fágica e a pesquisa resultante conduzia a um sinal de hibridação positivo para todas as placas fágicas obtidas no plaqueamento, o que era conseguido após 4-5 pesquisas.

5.2.3. Preparação de DNA de fagos λ.

A preparação de lisados de fagos λ de elevado título foi efectuada pelo método de Patterson & Dean (1987). Cerca de 50-100µl da suspensão de fagos foram colocados em contacto com 1 ml de células LE (preparadas como descrito em (# 5.2.2.) durante 20 min a 4°C de modo a processar-se a sua adsorção nas bactérias. Após a adição de 30 ml de meio NZY suplementado, as células foram incubadas a 37°C até a cultura apresentar uma lise completa, obtida habitualmente em 2-3 horas. Caso ao fim deste tempo se observem sinais de lise incompleta, podem ser adicionados mais 30 ml de meio, prosseguindo a incubação até que a lise se complete. Se esta não ocorreu, o título inicial da suspensão de fagos era considerado insuficiente, tendo-se usado então o sobrenadante da cultura para infectar as células LE e reiniciar o processo. À cultura lisada adicionou-se 0.1 ml de clorofórmio, e após mais 10 min de incubação, os resíduos celulares foram removidos do meio por centrifugação a 10 000g durante 10 min.

A preparação de DNA de fagos λ foi adaptada do método de Jofuku & Goldberg (1988). A fim de eliminar os ácidos nucleicos celulares, foram adicionados à suspensão de fagos 30µl de DNase 10 mg/ml e RNase 10 mg/ml. Após 30 min de incubação a 37°C, adicionou-se então 1.8g (6%) de NaCl, ficando a suspensão durante 1 hora a 4°C. Os resíduos resultantes foram eliminados por centrifugação a 7 000g durante 10 min, e os fagos foram precipitados com a adição de 2.1g (7%) de PEG 6000 a 4°C durante 14 a 18 horas. O sedimento de fagos obtido por centrifugação a 5000g durante 2 min foi ressuspensão em 500µl de tampão TM [50 mM Tris-HCl pH 8.0; 10 mM MgSO₄] e, após a adição de 2.5µl de SDS 20% e 15µl de EDTA 500 mM, foi aquecido a 60°C durante 10 min. O DNA

fágico foi obtido por extracção com fenol saturado com TE (2x) e clorofórmio (2x), sendo precipitado com 2.5 volumes de etanol absoluto e ressuspenso em 50µl de TE.

5.3. Análise dos clones recombinantes.

A análise dos clones obtidos foi realizada por hidrólise com enzimas de restrição, hibridação com sondas ubiquitina e sequenciação.

Cerca de 0.5-1.0µg de DNA foram digeridos com 5-10 U das enzimas de restrição apropriadas durante 1-2 horas a 37°C, sendo os fragmentos de DNA separado em geis de agarose (# 3.2.1). A partir destes geis, foi possível construir-se mapas dos clones com a localização aproximada dos locais reconhecidos pelas diversas enzimas de restrição (mapas genómicos). Alguns dos geis foram usados para transferência a filtros de nitrocelulose ou nylon (# 3.2.2) e posterior hibridação (# 3.2.4.1) usando-se como sonda o fragmento de 0.23 kb *HindIII* do clone pTU20 (# 3.2.3). Foi assim possível determinar os subfragmentos que continham sequências codificantes para a ubiquitina, e subcloná-los em plasmídeo (# 6) ou fagos M13 para posterior sequenciação (# 7).

6. Subclonagem em plasmídeo.

6.1. Preparação de fragmento.

O DNA contendo o fragmento a isolar foi digerido com as enzimas de restrição apropriadas de modo a obter-se pelo menos o dobro da quantidade final de fragmento pretendida. A separação do fragmento foi realizada por electroforese em geis de agarose como descrito em # 3.2.1. A recuperação do fragmento foi efectuada por electroeluição em mangas de diálise ou em câmaras Biotrap (Schleicher & Schuell), para grandes quantidades de fragmento (superior a 1-2µg), ou pelo uso de partículas de sílica ("GeneClean") para pequenas quantidades (até 1-2µg).

Electroeluição em mangas de diálise - Após a electroforese, a zona do gel de agarose contendo o fragmento a isolar foi cortada com um bisturi e colocada dentro de uma manga de diálise juntamente com solução TAE, de modo a não deixar bolhas de ar. A manga contendo o fragmento foi então recolocada no aparelho de electroforese, com o fragmento encostado ao lado da manga oposto ao pólo negativo, sendo-lhe então aplicada uma diferença de potencial de 100V durante o tempo necessário à migração do fragmento da agarose para a solução (usualmente cerca de 2 horas). A polaridade foi então invertida durante cerca de um minuto, de modo a deslocar o fragmento eventualmente adsorvido pela manga para a solução, que foi então recolhida para um "eppendorf". O brometo de etídio presente na solução foi removido com 2-3 extracções com 1-butanol, e o DNA seguidamente precipitado com 2.5 volumes de etanol absoluto na presença de 0.3 M de acetato de sódio. Após centrifugação a 14 000 rpm em centrífuga Sigma de "eppendorfs", o DNA foi ressuspenso em TE.

Electroeluição em Biotrap - A recuperação do fragmento é feita de modo semelhante à anterior, mas são usadas câmaras Biotrap (Schleicher & Schuell) em vez de mangas de diálise. A câmara consiste num tanque aberto nas pontas, onde são colocados dois filtros

impermeáveis ao DNA. Antes do filtro orientado para o pólo negativo, é colocado um terceiro filtro permeável ao DNA, mas a agarose, formando uma minicâmara onde irá ser recolhido o fragmento. O pedaço de agarose é colocado antes deste terceiro filtro, e a câmara e a minicâmara são enchidas com TAE de modo a que o pedaço fique coberto, sendo o conjunto recolocado no aparelho de electroforese. O restante tratamento é idêntico ao anterior, mas a inversão da voltagem é realizada por um tempo mais curto (15 seg). Este método permite em relação ao anterior obter fragmentos isentos de contaminação por agarose.

Recuperação pelo uso de partículas de sílica (GeneClean) - Este método baseia-se no princípio proposto por Vogelstein & Gillespie (1979), que mostram que uma solução aquosa de DNA e partículas de sílica na presença de agentes caotrópicos (como o NaI) sofre alterações a nível da estrutura da água que força as moléculas de DNA a serem selectivamente adsorvidas à superfície das partículas de sílica.

O protocolo usado foi o indicado pelos fabricantes do *kit* de purificação de DNA GeneClean (BIO 101). Resumidamente, o pedaço de agarose contendo o fragmento a purificar foi dissolvido a 50°C em 3 volumes de NaI 6 M. À solução foi adicionado 1 µl/µg (mínimo 5µl) de sílica e incubou-se em gelo durante 5 min. As partículas são recolhidas por centrifugação a 14 000 rpm em centrífuga de "eppendorfs" durante 5 seg, e lavadas 3 vezes com tampão de lavagem NEW. O DNA foi eluído por incubação das partículas a 55°C em 5-10µl de TE ou água bidestilada seguido de centrifugação para recolha da solução. A eluição pode ser repetida uma segunda vez para melhor rentabilização do processo.

6.2. Ligação do fragmento ao plasmídeo.

A ligação de fragmentos de DNA a plasmídeos pode ser feita usando-se fragmentos purificados (# 6.1), digestões de fragmentos purificados ou, em certos casos, digestões de vectores contendo o fragmento a clonar. No primeiro caso trata-se de uma clonagem directa, sendo as outras duas designadas por clonagens aleatórias ("shotgun").

A digestão do fragmento, quando necessário, e do plasmídeo a usar na ligação foi feita com enzimas de restrição que gerassem extremidades compatíveis no plasmídeo e nos fragmentos de interesse a clonar. A inactivação da enzima de restrição foi realizada por desnaturação a 65°C ou a 85°C, consoante a enzima, durante cerca de 20 min com descida lenta da temperatura a 30-35°C. Para a ligação, usou-se 50 ng de vector e uma relação de fragmento habitualmente de 10:1 em relação àquele. A ligação foi realizada em 10µl de uma solução contendo 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 20 mM DTT, 50 µg/ml BSA e 1 mM ATP, e ainda 6U (0.1 U Weiss) de T4 DNA ligase, a 14°C durante 14-18 horas ou a 4°C durante 3 dias.

6.3. Transformação de células *E.coli*.

6.3.1. Preparação de células competentes.

Para a preparação de células competentes usou-se habitualmente a estirpe XL1-B, e seguiu-se o método descrito por Cohen & Chang (1972).

A partir de uma cultura inoculada na véspera iniciou-se uma nova cultura em meio líquido a uma A_{600} inicial de 0.075 O.D./ml. Quando as células atingiram 0.3 O.D./ml, foram recolhidas a 7 000g durante 2 min a 4°C, lavadas, sempre a 4°C, em metade do volume da cultura de partida com 0.1 M $MgCl_2$ e ressuspensas, também em metade do volume inicial, em 0.1 M $CaCl_2$. As células permaneceram então em repouso durante 20 min a 4°C, após o que foram recolhidas e ressuspensas em 1/15 do volume inicial em 0.1 M $CaCl_2$, 15% glicerol. As células, tornadas competentes, foram então aliquotadas e armazenadas a -70°C.

6.3.2. Transformação de células competentes.

A transformação foi realizada por contacto do DNA plasmídico (2µl=10 ng de ligação (# 6.2) ou 500 pg de plasmídeo) com 200µl de células competentes (# 6.3.1) durante um mínimo de 40 min a 4°C. As células foram então submetidas a um choque hipertérmico de 1-2 min à temperatura de 42°C, e recolocadas a 4°C por mais 1-2 min. De seguida, foi adicionado 1 ml de meio LB à suspensão, que foi incubada com agitação durante 1 hora a 37°C. As células foram recolhidas por centrifugação a 2 000 rpm em centrífuga de "eppendorfs" durante 3 min, ressuspensas no meio residual e espalhadas, com a ajuda de bolas de vidro estéreis, em placas LA suplementadas com ampicilina e, quando necessário, com X-Gal e IPTG (# 1.2.2). O crescimento foi efectuado por incubação das placas em estufa a 37°C durante 14-18 horas.

6.4. Preparação de DNA plasmídico.

6.4.1. Extracção de DNA em pequena escala (miniprep).

A extracção de DNA plasmídico em baixa escala (miniprep) foi normalmente usada para análise de subclones após transformação. Aliquotas de 1.5 ml de LB+Ap inoculadas com colónias dos transformantes que se pretendeu analisar foram cultivadas durante 14-18 horas a 37°C com agitação. De cada cultura foi então preparada uma aliquota de 500µl com 15% glicerol, que foi guardada a -20°C (ou a -70°C para um armazenamento prolongado), sendo o restante usado para a extracção do plasmídeo, que foi efectuada por dois métodos: o método de extracção por fervura (Holmes & Quigley, 1981) e o método da lise alcalina (Birnboim & Doly, 1979).

Extracção por fervura ("boiling") - As células foram recolhidas por centrifugação a 14 000 rpm em centrífuga de "eppendorfs" durante 2 min e ressuspensas em 150µl de tampão STET [50 mM Tris-HCl pH 8.0; 50 mM EDTA; 8% sacarose; 5% Triton X-100]. Após a adição de 15µl de uma solução de lisozima 10 mg/ml, que actuou durante cerca de 30 seg, a suspensão foi fervida durante 1 min. O sedimento obtido por centrifugação a 14 000 rpm em centrífuga de "eppendorfs" durante 15 min, constituído por proteínas desnaturadas e por DNA cromossómico, foi removido com o auxílio de um palito. O DNA contido no sobrenadante foi precipitado com 15µl de acetato de sódio 3 M e 375µl de etanol absoluto, sedimentado por centrifugação durante 15 min a 4°C, ressuspenso em 50µl de TE e guardado a -20°C para posterior análise.

Extracção por lise alcalina - As células foram recolhidas por centrifugação a 14 000 rpm em centrífuga de "eppendorfs" durante dois minutos e ressuspensas em 200µl de solução de pré-lise [25 mM Tris-HCl pH 8.0; 10 mM EDTA; 50 mM glucose]. Após 30 s de incubação com 15µl de lisozima 10 mg/ml, foram-lhes adicionados 200µl de solução de lise [0.2 M NaOH; 1% SDS] preparada no momento. A suspensão foi agitada vigorosamente e foi-lhe adicionada 200µl de solução de neutralização [5 M Acetato de potássio pH 4.8] seguida de nova agitação vigorosa. O sobrenadante obtido após 5 min de centrifugação foi precipitado com igual volume (600µl) de isopropanol e novamente centrifugado durante 5 min. O sedimento de DNA foi ressuspenso em 100µl de TE, novamente precipitado com 250 µl de etanol absoluto e finalmente ressuspenso em 50µl de TE, sendo guardado a -20°C para posterior análise.

6.4.2. Extracção de DNA em média escala (mediprep).

A preparação de DNA em media escala foi normalmente usada para preparações de quantidades entre 0.2 a 2 mg de DNA de clones ou subclones de interesse para estudos posteriores. O método usado foi o da lise alcalina (Birnboim & Doly, 1979) ou uma versão comercial do mesmo, o CirclePrep (BIO 101).

Extracção por lise alcalina - As células de uma cultura de 50 ml em LB+Ap, crescida durante 12-18 horas a 37°C, que continham o clone de interesse foram recolhidas por centrifugação a 5 000 rpm durante 2 min e ressuspensas em 5 ml de solução de pré-lise. A suspensão foi incubada durante 2-3 min com cerca de 5 mg de lisozima, e após a adição de 10 ml de solução de lise fresca, foi agitada suavemente até se observar o desaparecimento das nuvens de células. Juntaram-se então 5 ml de solução de neutralização, agitando-se até se observar a formação de pequenos farrapos brancos, constituídos pelas proteínas desnaturadas e DNA cromossómico. A suspensão foi centrifugada 5 min, e o sobrenadante filtrado através de uma gaze para novo tubo, onde o DNA foi precipitado com 12 ml de isopropanol. Após 15 min de centrifugação a 7 500 rpm, o sedimento de DNA foi ressuspenso em 500µl de TE+RNase [TE; 0.5 mg/ml RNase livre de DNases] e incubado a 37°C durante 1 hora. A enzima foi extraída com 500µl de fenol saturado com TE e com 500µl de clorofórmio:isoamílico (24:1), e ao sobrenadante aquoso foram adicionados 1.25 ml de etanol absoluto para precipitação do DNA. O DNA obtido após centrifugação a 14 000 rpm durante 15 min em centrífuga de "eppendorfs" foi finalmente ressuspenso em 100µl de TE ou água bidestilada, e a sua concentração determinada por espectrofotometria a 260 nm, sendo que 1 O.D. ⇔ 40 ng/ml.

Extracção por CirclePrep (BIO 101) - O DNA foi extraído segundo as indicações do fabricante. O processo é equivalente ao da lise alcalina, diferindo após a precipitação do DNA com isopropanol. Neste ponto, o RNA e DNA de cadeia simples são precipitados por adição de 300µl de uma solução saturada de LiCl, e o DNA plasmídico é purificado do sobrenadante resultante por partículas de sílica, de forma equivalente à descrita para o *kit* de GeneClean (BIO 101) em # 6.1.

6.4.3. Extracção do DNA em grande escala.

A extracção em grande escala produz um DNA plasmídico de elevada pureza e foi usada sobretudo para a preparação de plasmídeos ou de clones de grande interesse. O método usado foi o da fervura (Holmes & Quigley, 1981).

Cerca de 500 ml de cultura de células contendo o plasmídeo de interesse foram cultivadas durante 12-18 horas, com adição de 100 mg de cloranfenicol a uma A_{600} de 1 O.D./ml) nos casos em que o plasmídeo a preparar existe em baixo número de cópias na célula (p.ex. pBR322). As células foram recolhidas a 7 000 rpm em Sorvall durante 10 min e ressuspensas em 25 ml de tampão STET. Após a adição de 2.5 ml de lisozima 10 mg/ml, que actuou durante 1 min, a suspensão foi fervida durante 1 min e 30 s e centrifugada a 30 000 rpm no rotor 70Ti (Beckman) durante 20 min. O sobrenadante foi recolhido e o DNA precipitado a -20°C com 2.5 ml de acetato de sódio e 25 ml de isopropanol, sendo sedimentado em Sorvall a 8 500 rpm durante 30 min. O DNA foi seguidamente purificado em gradiente de cloreto de cézio como descrito em # 3.1.

6.5. Análise dos subclones obtidos.

Os subclones obtidos após ligação e transformação foram analisados para determinação do seu interesse por digestão com enzimas de restrição. Cerca de 10 a 25 μl do DNA plasmídico obtido por miniprep foi digerido com 10U da(s) enzima(s) de restrição apropriada(s) num volume final de 50 μl a 37°C durante 3 horas. As digestões foram então analisadas por electroforese em minigeis de agarose 0.8-1.0% (# 3.2.1) a 80-100V durante 1-2 horas de modo a determinar se o(s) fragmento(s) de interesse tinha(m) sido efectivamente inseridos no plasmídeo.

7. Determinação da sequência nucleotídica do DNA.

7.1. Subclonagem em fagos M13.

7.1.1. Geração de subclones por clonagem directa ou aleatória.

Os subclones em derivados do fago M13 - mp18 ou mp19 - foram gerados por clonagem directa de fragmento purificado ou por clonagem aleatória de uma mistura de fragmentos como descrito em # 6.2.

7.1.2. Geração de subclones por deleção parcial do fragmento inserido.

Foram também realizadas deleções parciais de fragmentos previamente inseridos em fagos M13 segundo o método de Dale *et al.* (1985), usando-se o *kit* "Cyclone I Biosystem" (IBI). O método usa a forma vegetativa de cadeia simples (ss) de fagos M13mp19 ou M13mp18, desde que este último não possua o fragmento inserido nos locais *HindIII* ou *SphI*, devido à limitação de restrição com a enzima *HindIII* nos extremos de fragmentos.

O primeiro passo do processo consiste na hibridação de um oligómero próprio ao fago ss, de forma a conseguir-se uma zona única de cadeia dupla acessível à enzima de restrição

que irá linearizar o vector. Os oligómeros complementam a região do fago a jusante do "polilinker" e as seis bases reconhecidas pela enzima *EcoRI* dos derivados mp19 (RD22) ou as doze bases reconhecidas pelas enzimas *HindIII* e *SphI* dos derivados mp18 ("restriction 18-mer"). A 5µl de fago ss contendo o fragmento de interesse, preparado em # 7.3.1, juntaram-se 2.5µl de tampão de reacção, 10µl de água bidestilada e 2µl de oligómero RD22 ou "restriction 18-mer". A hibridação foi realizada por aquecimento a 70°C durante 2 min com descida suave da temperatura a 35°C. Após a adição de 1µl de DTT 100 mM, a restrição foi realizada durante 1 hora a 42°C com 20 U de *EcoRI* para derivados mp19 ou durante 2 horas a 50°C com 40 U de *HindIII* para derivados mp18. A enzima foi inactivada por aquecimento durante 10 min a 65°C (*EcoRI*) ou a 85°C (*HindIII*).

A deleção foi realizada com a enzima T4 DNA polimerase, que possui actividade exonucleásica 3'→5'. A taxa de digestão varia com a concentração da enzima por µg de DNA, encontrando-se nas condições usadas entre 40 bases/min para 1 U/µg e 90 bases/min para 3 U/µg. À digestão anterior adicionaram-se então 2.5µl de DTT 100 mM, 1µl BSA 10 mg/ml e 6U de T4 DNA polimerase, prosseguindo a reacção a 37°C. A períodos definidos previamente (dependendo do tamanho do fragmento inserido) retiram-se aliquotas de 3.5µl que foram aquecidas a 65°C durante 10 min para inactivação da enzima e colocadas de seguida a 4°C. As aliquotas foram seguidamente reunidas e reaquecidas novamente.

A recircularização dos fagos é conseguida com o uso de oligómeros que completam a região a jusante do "polilinker", as seis bases reconhecidas pelas enzimas *EcoRI* ou *HindIII* e um polímero (dG)_n para derivados mp19 (RD22), ou um polímero (dA)_n para derivados mp18 (RD29). Esse polímero é sintetizado pela enzima terminal transferase (TdT), que adiciona indiscriminadamente desoxiribonucleótidos 5'→3'. À digestão anterior adicionaram-se então 3µl de dGTP 50µM para derivados mp19, ou 3µl de dATP 50µM para derivados mp18, 1µl de água bidestilada e 10U de TdT, prosseguindo a reacção a 37°C durante 10 min e inactivando-se a enzima por aquecimento a 65°C durante 10 min. Após a adição de 1µl de oligómero (RD22 ou RD29), realizou-se a hibridação por aquecimento a 70°C durante 2 min com descida muito suave da temperatura a perto de 25-30°C. Seguidamente, adicionaram-se 3µl de ATP 10 mM, 4µl de água bidestilada e 120 U (2U Weiss) de T4 DNA ligase, decorrendo a ligação durante 20 horas a 4°C.

7.2. Transformação com o fago M13.

Para a transformação de fagos M13, o DNA fágico (2µl, equivalendo a cerca de 8 ng de ligação ou a 100 pg de DNA purificado) foi colocado em contacto com cerca de 200µl de células competentes durante um mínimo de 40 min a 4°C. As células foram seguidamente submetidas a um choque hipertérmico de 1-2 min à temperatura de 42°C, e recolocadas a 4°C por mais 1-2 min. À suspensão foram então adicionados 270µl de uma mistura contendo células XL1-B, X-Gal e IPTG [por transformação: 200µl células XL1-B a 0.3 O.D.; 40µl X-Gal 2%; 40µl IPTG 100 mM] e 3 ml de H-Top previamente mantido a 50-60°C. A mistura foi rapidamente agitada e vertida para placas H aquecidas. O crescimento foi efectuado por incubação das placas em estufa a 37°C durante 14-18 horas.

7.3. Preparação e análise dos fagos recombinantes.

7.3.1. Preparação da cadeia simples (ss) dos fagos.

As placas fágicas obtidas por transformação de fagos M13 foram repicadas para 1.5 ml de uma diluição 1:100 em 2xTY de células JM101 ou TG1 de uma cultura de véspera. As culturas foram então incubadas a 37°C com agitação durante 5 horas ou, para culturas realizadas a partir de glicerol ou quando necessário, durante 14-18 horas. Após esse tempo, os fagos foram separados das células por duas centrifugações sucessivas de 2 min em centrífuga de "eppendorfs". As células recolhidas foram usadas para preparação da cadeia replicativa ds (# 7.3.2), ou descartadas quando a preparação era desnecessária. Os fagos contidos no sobrenadante foram precipitados com 200µl de PEG/NaCl [20% polietilenoglicol 6000; 2.5 M NaCl] durante 15 min, sedimentados com 5 min de centrifugação e ressuspensos em 100µl de TE. O DNA fágico foi obtido por extracção com 50µl de fenol saturado com TE seguido de precipitação com 10µl de acetato de sódio 3 M e 250µl de etanol absoluto, sendo ressuspensão num volume final de 25µl de TE.

7.3.2. Preparação da cadeia dupla (ds) dos fagos.

A cadeia dupla dos fagos foi, quando necessária, preparada pelo método da fervura (# 6.4.1) a partir do sedimento de células obtido em # 7.3.1.

7.3.3. Análise dos fagos recombinantes.

A análise dos fagos recombinantes foi feita por três métodos.

A análise do DNA de cadeia simples foi feita habitualmente por migração em minigel de agarose 1% a 80V (# 3.2.1), de modo a determinar a pureza do DNA e estimar a presença do fragmento de interesse por comparação com um DNA fágico de cadeia simples sem inserção.

A presença do fragmento de interesse tem contudo de ser confirmada por um método mais eficaz. Nos casos em que o fragmento é inserido pela primeira vez em fagos, preparou-se a cadeia dupla do fago (# 7.3.2) e a inserção foi analisada por restrição enzimática, de modo semelhante à referida em # 6.5.

Nos casos em que o fragmento de interesse, ou um outro fragmento que o contenha total ou parcialmente, se encontrasse já clonado em fago na orientação complementar à pretendida, usou-se um método de complementação e digestão pela enzima S1. Assim, 5µl de ss do fago a testar foram adicionados a 5µl do fago contendo o fragmento complementar e a 1µl de um tampão usado para restrição enzimática contendo entre 60 a 100 mM de NaCl. As zonas complementares dos fragmentos foram hibridadas por aquecimento a 70°C com descida suave de temperatura a 35°C, e as regiões que não hibridaram foram então digeridas a 37°C durante 15 min após a adição de 11µl de água bidestilada, 2.5µl de tampão da enzima S1 [50 mM acetato de sódio; 200 mM NaCl; 1 mM ZnSO₄; 0.5% glicerol] e 10U da S1 nuclease. A digestão foi então analisada em minigeis de agarose 0.8-1% a 80V (# 3.2.1). Este método permite não só determinar correctamente o tamanho da inserção como pesquisar DNAs fágicos contendo fragmentos complementares.

7.4. Sequenciação do DNA de cadeia simples.

A sequenciação foi feita em ss preferencialmente a DNA plasmídico devido à maior facilidade de manuseamento e à maior legibilidade das reacções de sequenciação resultantes. O método usado foi o descrito por Sanger *et al.* (1977) usando-se o *kit* de sequenciação Sequenase ver. 2.0 (USB). Este método baseia-se na incorporação aleatória por uma DNA polimerase de 2',3'-didesoxiribonucleótidos (ddNTPs) e de ^{35}S -dATP na cadeia polimerizada, originando a terminação do processo de síntese nos locais onde o análogo foi incorporado. As reacções são então migradas em geis de poliacrilamida, que permite separar as diferentes cadeias produzidas pelos seus tamanhos com precisão de um nucleótido de diferença.

Resumidamente, a 7 μl de DNA ss foram adicionados 1 μl de oligómero iniciador ("40 primer") e 2 μl de tampão de reacção, e a hibridação foi realizada por aquecimento a 70°C seguida de uma descida suave da temperatura até 30-35°C. A reacção de polimerização foi iniciada com a adição de 0.5 μl de ^{35}S -dATP (>1000 Ci/mmol), 1 μl 0.1 M DTT, 2 μl de uma mistura diluída 1:15 dos nucleótidos dCTP, dGTP e dTTP, e 2 μl da enzima Sequenase diluída 1:16. A reacção prosseguiu durante 5 min a 20-25°C, sendo de seguida distribuída (3.5 μl a cada) por quatro tubos pré-aquecidos a 37°C contendo, cada um, 2.5 μl de uma mistura dos quatro dNTPs e um ddNTP distinto para cada tubo. Estas reacções, ditas de terminação, prosseguiram durante 15 min, altura em que se lhes adicionou 4 μl de solução Stop.

As reacções foram analisadas em geis desnaturantes de poliacrilamida com 0.35 mm de espessura. Os geis são preparados por enchimento de caixas de vidro próprias com 80 ml solução contendo 42% ureia, 5.7% acrilamida desionizada e 0.3% bisacrilamida em TBE [90 mM Tris-HCl pH 8.3; 90 mM ácido bórico; 2.5 mM EDTA], à qual se adicionou 80 μl de TEMED (N,N,N',N'-tetrametilenodiamina) e 480 μl de 10% persulfato de amónio como iniciadores da polimerização. A electroforese efectuou-se 14-18 horas depois em sistemas BRL S2 usando-se TBE como solução de electroforese. O gel foi pré-aquecido durante 30 min com uma diferença de potencial máxima de 1 500V (ou potência máxima de 70W), e as amostras foram então aplicadas no gel depois de previamente fervidas para desnaturação das cadeias de DNA. A electroforese decorreu nas mesmas condições durante cerca de 1h30 a 2 horas para leituras até 200 nucleótidos, ou de 5-6 horas para leituras entre 200 a 400 nucleótidos. Terminada a electroforese, os geis foram fixados por imersão durante 15 min em 10% ácido acético, 10% metanol, transferidos para papel 3MM (Whatmann) e secos durante 60 min a 80°C num secador de geis sob vácuo. Os geis secos foram colocados em contacto com películas X-OMAT AR5 (Kodak), a 20-25°C, durante pelo menos 14-18 horas. A revelação foi efectuada nas condições indicadas pela Kodak.

7.5. Análise das sequências nucleotídicas.

Os geis de sequenciação foram lidos usando-se uma caneta óptica (Pharmacia-LKB), sendo os dados tratados em computador IBM compatível pelos programas de análise DNASIS 5.0 e PROSIS 6.0 (Hitachi Software Engineering). A pesquisa em bancos de dados e algumas análises de sequência foram realizadas em sistema UNIX pelo programa GCG (Wiscosin).

Capítulo III

RESULTADOS 1

Genes poliubiquitina em *T.pyriformis*

1. Isolamento de genes poliubiquitina em *Tetrahymena pyriformis*.

De forma a complementar o estudo iniciado neste laboratório sobre a família de genes ubiquitina (Neves, 1991), foi houve necessidade de proceder ao isolamento de dois clones adicionais. Um destes clones, pTU1, obtido a partir de um banco genômico de fragmentos *Hind*III de DNA macronuclear inseridos em pBR322 (Neves *et al.*, 1988), contém um fragmento de 6.8 kb. Para a obtenção do outro clone, construiu-se em λ DashII um banco genômico de DNA macronuclear parcialmente hidrolisado com *Sau*3AI (ver Material e Métodos). O banco foi pesquisado por hibridação *in situ* com o fragmento *Hind*III de 0.23 kb do clone pTU20 (Neves *et al.*, 1991a), tendo-se obtido diversos clones com sinal positivo que foram seguidamente analisados mais detalhadamente, por determinação do seu mapa de restrição e por hibridação com uma sonda ubiquitina. Um destes clones revelou incluir um fragmento *Eco*RI com dois genes poliubiquitina incompletos e pouco usuais já previamente descrito (Neves *et al.*, 1990). Deste clone fágico foi então subclonado um fragmento *Hind*III de 4.4 kb que contém o referido fragmento *Eco*RI, sendo o clone resultante denominado por pTU λ 19H.

Os mapas de restrição obtidos para os novos clones encontram-se representados na Fig.III.1, juntamente com os mapas de outros clones já referenciados (Neves, 1991). A sequenciação dos novos clones revelou a presença de três e dois genes ubiquitina, respectivamente. De modo a facilitar as análises posteriores, a cada sequência codificante individual foi atribuída uma letra identificadora e cada repetição ubiquitina foi numerada, à exceção das repetições terminais de genes incompletos que foram designadas arbitrariamente como *n*. Aparentemente, os genes poliubiquitina parecem encontrar-se agrupados em série no genoma do ciliado *Tetrahymena pyriformis*, com pequenas regiões intergênicas de 1.5 a 2.5 kb cada. Curiosamente, nos fragmentos TU1 e TU2 foram igualmente encontrados genes codificantes para 5S-rRNAs, embora o significado da sua presença junto de genes ubiquitina seja desconhecido (Guerreiro *et al.*, 1993; ver Apêndice 1).

2. Análise dos genes poliubiquitina.

2.1. Análise das sequências codificantes.

A sequência nucleotídica das regiões codificantes dos diversos genes poliubiquitinas clonados (# 1) e as correspondentes sequências previstas de aminoácidos encontram-se apresentadas nas Fig.III.2C e Fig.III.3, respectivamente. A análise comparativa das sequências sugere que os genes poliubiquitina de *T. pyriformis* podem ser classificados em três grupos distintos de sequências relacionadas.

As repetições dos genes poliubiquitina A (TU20), B, C (TU2), X (TU3) e Y (TU4) codificam para proteínas idênticas entre si, pelo que foram considerados como pertencentes ao Grupo I. Estes genes apresentam uma elevada similaridade entre as suas sequências nucleotídicas, com um mínimo de 85% de identidade entre as diversas unidades. Mais ainda, existem regiões onde mesmo as mutações silenciosas se encontram praticamente ausentes: nas regiões codificantes para os resíduos de aminoácidos 27-37, 42-49 e 58-64 apenas 10 dos 234 códons representados apresentam mutações. Os genes do Grupo I apresentam igualmente um - CAA - ou dois - TCTCAA - códons extra na última unidade ubiquitina de cada gene.

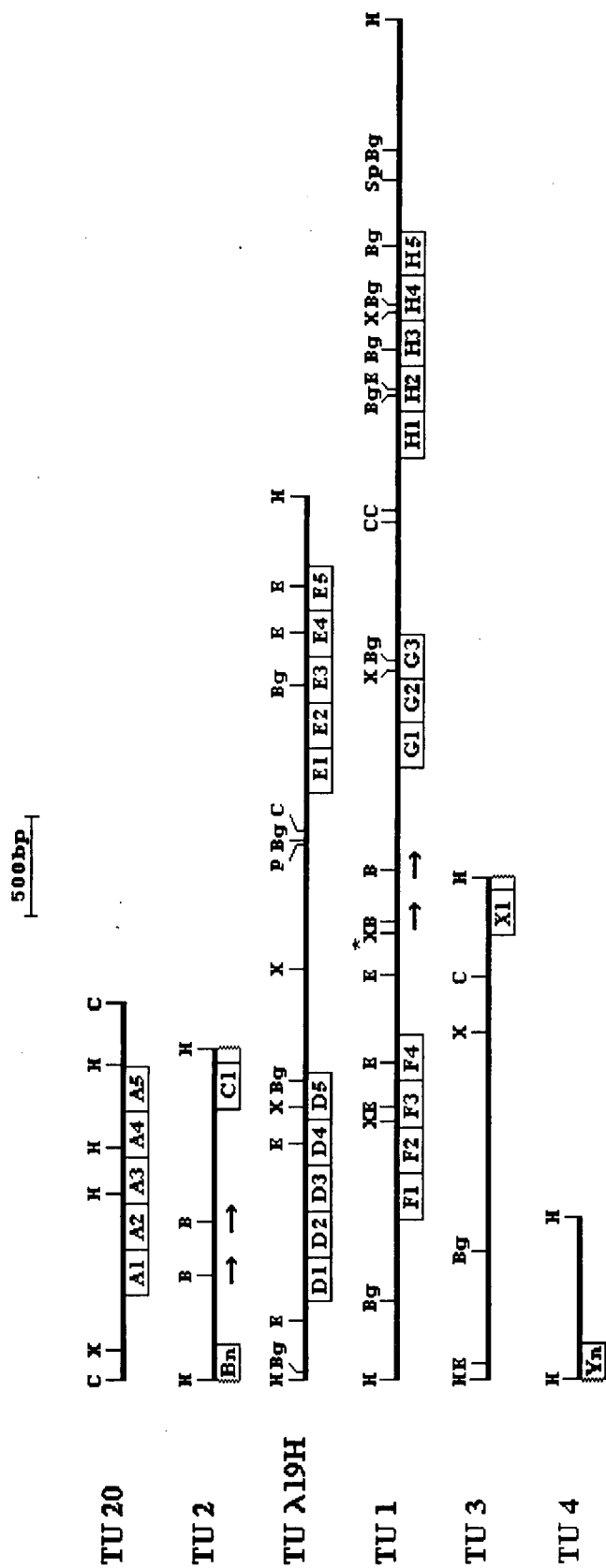


Fig. III.1. Mapas de restrição dos fragmentos de DNA clonados contendo genes poliubiquitina. Os mapas de restrição foram determinados por digestão com as seguintes enzimas de restrição: B-*Bam*HI; Bg-*Bgl*II; C-*Cla*I; E-*Eco*RI; H-*Hind*III; P-*Pst*I; Sp-*Spe*I; X-*Xba*I. X* representa um local de restrição *Xba*I sujeito a uma metilação *dam*. A região codificante foi determinada por hibridação e confirmada por sequencição. As caixas representam unidades ubiquitina e as flechas genes de 5S-rRNA. A cada gene foi atribuída uma letra, sendo as suas unidades numeradas.

Os genes poliubiquitina de *T.pyriformis* D, E (TUλ19H), F e H (TU1) foram considerados como pertencentes ao Grupo II. As suas unidades ubiquitina são diferentes entre si, e apresentam 82 a 93% de identidade com a do Grupo I. Nenhum dos genes apresenta codões extra na sua última repetição, à qual falta igualmente o resíduo Gly76. Os genes do Grupo II estão extensivamente mutados nas suas sequências nucleotídicas, inclusive em regiões onde no Grupo I praticamente não se observam mutações. Contudo, não apresentam codões de terminação internos (TGA em *Tetrahymena*; Kuchino *et al.*, 1985; Hanyu *et al.*, 1986) ou perda da grelha de leitura.

O gene G (TU1) é o único membro do Grupo III. Como descrito para o Grupo II, é constituído por três repetições distintas entre si, sendo contudo muito mais divergentes que as daquele grupo, já que apenas apresentam 60 a 70% de identidade com as sequências de aminoácidos do Grupo I. Para além da ausência do resíduo Gly76 na última unidade, dois outros tripletos estão ausentes nas segunda e terceira unidades, que possuem assim 75 e 74 codões, respectivamente. Nenhuma das repetições possui um resíduo Gly na posição 76, o que impossibilita o produto do gene G de ser convertido em monómeros (Lopez-Otin *et al.*, 1989). Todavia, para além da elevada taxa de mutações e da deleção de três tripletos, não são observados nem codões de terminação internos, nem perda da grelha de leitura.

2.2. Análise das sequências não codificantes.

A ubiquitina é uma proteína de "stress" (Bond & Schlesinger, 1985; 1986; Özkaynak *et al.*, 1987), tendo já sido demonstrado que os genes ubiquitina de *Tetrahymena* são igualmente induzidos na presença de "stress" hipertérmico (Neves *et al.*, 1988, 1991a). Estes resultados são apoiados pela detecção de elementos de choque térmico (HSE - "heat shock elements") nos promotores sequenciados dos genes do Grupo I (Fig.III.2A). Apenas um HSE foi encontrado por gene, encontrando-se o do gene A (TU20) invertido. Outro elemento existente nos promotores dos genes do Grupo I é o motivo AGAGA. Este motivo é encontrado frequentemente em genes de *Tetrahymena*, e pode corresponder ao motivo TATA de outros organismos (Brunk & Sadler, 1990). De facto, devido ao elevado conteúdo de nucleótidos AT nas regiões não codificantes de *Tetrahymena*, os elementos TATA canónicos podem ocorrer quase aleatoriamente, pelo que este organismo poderá usar uma solução diferente. Elementos alternativos ao motivo TATA foram observados, por exemplo, em vários genes ribossomais de *S.pombe*, onde foram substituídos pelo motivo CAGTCACA (Witt *et al.*, 1993). Todavia, a *Tetrahymena* possui uma proteína de ligação ao motivo TATA, TBP ("TATA-binding protein"), cujo gene foi recentemente clonado (Stargell & Gorovsky, 1994). A proteína deduzida da sequência é contudo uma das mais divergentes TBPs encontradas, possuindo inclusive uma extensão C-terminal de 20 resíduos de aminoácidos única, pelo que é possível que não reconheça motivos TATA canónicos, embora a sua ligação a motivos AGAGA também não tenha sido demonstrada. Alternativamente, as sequências AGAGA podem constituir um motivo independente dos motivos TATA e requererem factores de ligação distintos. Em *Drosophila*, o gene *hsp26* possui igualmente zonas ricas em purinas que parecem estar envolvidas na sua transcrição, embora exigindo um factor diferente da TBP (Gilmour *et al.*, 1989).

(A)

-480 -450 -420 -390 -361
A atcgatattatttggattgttggctggtgatttttggttgttggcttggttgta
C tttttattatttttcttcttccataaaatagattcGAAtaTTCtatagaatataataaatttttagaataaaaaaactttttccagaaaatagaattaaagatacaaa
D aagcttgggtgcagctacttctatttattaata
E aattttagacaagttaaatattcatatgaacgtaattatcatcattaaacacataactaattcaatataaaaaatatatttctataagtgaaatggaatcatttggtttttggatttcc
F atctaacaagatttgttaaatatgtaataaaaaataaaaaaaccctttccaaagtttttgggtattaaatgtaagatctttagttcttagtttcaggtaattctctcttggattt
H caaattcagaataaaaagagctcgaacatattaaatagcgctctttaattagagtgcatcttcaattagaaaaggcttaagtttagattttatggtagaaattaaaaataaccattg
G ttttagttatttggaaatatttttaattgggtattatgtagattatgcaatttttagatcattttg
X aggcattttaagttgtaattatgtatttatattagatgaagttatttttttttttttaattcaattcgtattcctaaagtaataataaaaaaacacatatatttattaaaaataattttct

-360 -330 -300 -270 -241
A gcttgccttgattgttttggagagaataaaagcaaaagaagaagaattgtttccattcaaaaaaactcTTCtaGAAGaagaagaataaattatattgttttcttatttatt
C aatataatataattttgatcaaatcttcttcttggatttgggtatcatataatctattcatttaataatagaacaattcttctatcaaattATTGatAGAGAGGcaaaaattatct
D gatctgctctataaataaattatgacttactgaaaaaattattgataataaataagataaatttttggatttccaagatccgctattttagcagaattttatgactcttatttgaaca
E atctcttttataagaatagaattactatttgaagatatatttcttctgtaaaatagataaattcatttatttcaagaactgcagatactcagatttaattagatctgttattta
F tatcattgatttatagaagaaaagatttttaatatagataaatttttggatttcaagatccacacatattcagaattgatttagctcgattaattaaaaaaaattgactttctaatt
H ctagtataagcatttgcattcaattagataaactcgattgttggatttctattatttttagaatagataaatttcaattcaagaatccaaatcgattcagaatttattggattgaa
G ttgatttcaaaaatgggtgaaaattaaagaaaaagaaatatttaataaataatcggttagatttttttagagtttggttctcattttatagaaatcatttttgaataagcatagaatat
X caaatttttGAAtaTTCtaaatattttttaaataattaaaacattatttgcactataaaatcacaatcaatttttttagatcttaggaatagattcttttttcaatgataatAATag

-240 -210 -180 -150 -121
A ttattaaattcttcttcaaaaaattcggcaagcaagcgactttcattcaataaaaaagaagcgtaacaaacgagataacaaaaATTatctGAGAGGtttgcactcagacttaacgga
C tctattttctgatcataatcttcttctttaaattgatcagtttctcttatttatttaggaatagattctgaaattccctctaatttcatagctaatcagattcaatcatatctatttt
D caatactgaatgtaatttcattaaatcaaaagaATGCTCTACTTTCTATTTCCTAcTcTcTataaataaactgggtgactaatttctACTCAtAATCTATGCTATCaatctcttatt
E acacaagacgctgcaatctaatttttaatacaaatcgATGCTCTATTCTCgATTTCCTATTTCCTgcatattgtaaaagtaagattaatcaCCTCAaAATTTATaATATCTaTctatttag
F tcattaaattttcatgaATGCTCTACTTTCTATTTCCTTCTaagaataaagggtttttaaactagatttcaattatactACTCAtAATTCATCaTATAGATTaatcacttgatttt
H taaaggaatctatttaattaaatcATGCTCTACTTTCTATTTCCTATCTCTcaaaagaaaaacagcgaaatctatcaacacaagcACTCATAATTTATCTGATCGaccattttcttggga
G tcaaagtttaattcaccacaaaaataagaacattgaaatttatttttggatttctacttcaaaaaacagatttaattctttctatttataataatcataaaatgatataatACTCAaAA
X acAGAGATccaatcagaattatttctaatttctgaatcgatatgttttcttgaatagatcctttattccgcggtattttatttctcagattcaatcaactctatttttttcttttaa

-120 -90 -60 -30 -1
A tctatttcttccactatttcaaatctcttctcagaccttctcaaaaaaaagcaaaacaattcaaaaaagaatttaatttctaataacccaaaaaacacaaaaaaagcaacaccttatcaa
C attttccactttaaaaagcttttttttaccacagtaaaattcataatcaattcataaaaaaaaataactctaaataagaacactcaaatcaacacattaaacgatttaccacaaacaa
D ataaatcaccattttagaattctacaaaaaaattataaacttcaaaaaaaaatatcaataaaaaataaacatcattcgaaaaccaaataaacaccattcagaagtatccataaatact
E tgtataaatagtttcgaacaacgacttcaaaatgttataataaattcaaaccaatcaaaaattaaataaataagaagacaaacaaacttaactcatttttgcgttcaaacaaaaacac
F aaataaccacttctatttctcaccaaatttttataaaaaaattcaaatcattaaaaaccaaataagaataatataactcaacaataattaaaaactcactaataagtaatacaaaattaca
H taaatagaatcggttcgctttctcataaaaaatataataaattaaaatcaatcattaaatcctaataatcatcattactcaaaaaattcaaaagttaaacaataatatacaataaataat
G TTTATCaaATCTGATtattgttttcttcttttaacagaaaaattttgtgtttgttttcaataaaaaatcaataaattacagagataacttaaaataagttacttaaaagctgaattaaaaata
X aaacagtcgaaaaagatcagaataaatcatataatcaattcatttcaaaaacatcaaaaagtgataatttaatacaaaaactaatacaacaaaaataataaataacttctacataattccta

(B)

+1 +30 +60 +90 +120
A ttaaagcttcaaaaaagagccccctaacccttactatcatttcttacaacaattgaaaattgattatgaataaagttacccttaaaatttatattgattttcttctcttatattatt
B aataaaaaatattagaaaaataaaaaacttattcaatttaataacttataaatactatcaattctattttataaaccattgtatctactcataaaaatgatataaaaaatttcaattat
D ctataaaaaataatcaaggataaataagaataaaaaattatagatatattgtgaattaatatattttctattttgttaataatataataaattgaaatttataaattttaaacaagaa
E ttatataatcatatcattttaaataagattacatatattatcaatttagtacttatataattttatattttccataagttaaattataatctaaatttcaaattaattcagattgaaat
F atatagaatttaattatgaatattaaaaataaacattaaaaatgtatttaattctgtatttatacaaaatttttctcatacaattttaaattgtgatttatcatctgactattaaaagttat
H acataaaataaaaatagatataaaaaagtcataaattgaatttgaatacttgcattgttttaattcataaaaatttgatttgattcttttcaagtaaaattgtttaaataagaattattt
G atctggagaataaaatagatttctattgtcatttaattctttaaataatgtgattgttttgatgaatttttaaaatattttgtttattcttcttcttatgatactttaaattagtgaaac
Y aaaaataaatcaataaaaaaattttaaataatattctataatccatttctatatttcttataaattgttctaaattctataaaataaattccatttgaattaaagtttaaattaaagtcacaa

+121 +150 +180 +210 +240
A aatgttgattgactattaaattttgtagaaatactctacaaactaatcaatcatctctcacattctaattttcttcttaatacaaaattcttttctaaattttcttggttcattccttct
B tctgaattttcataatgatatttcataaaatctcacaataaaaaatcactaatttcaaaaacaaataaataattttaaaaaatttaacagatttaacaaattaaactcatattatgagttt
D tttcatattgttttttaattattttattttttagatatttgaattattctattctattctagatataaacaatatagttgtgttaatacatattgtacatacataataaattatttgattt
E caaagtataacttatcataaatttcaaaaacttattttagtacttcttcaattgattcttattcattcagttataaaaacgtaacagattaaagtattgattgtctttaaatttattc
F tttgcaatgatgaccatagattgattttaaatttctgaaataaccttctcttattctcggtagttaccacaaataatttatggatgaatatctttgttcaaaagtaagattttgattc
H tctattataagttgaagctatttaataatacatacaaaaaagaagattagtaattgaataaataattgtttaaattgagtaaaaaataagtttcaaaaatttcatcaaaaaataactggaaatc
G aaatgaagttaaagatttcatctctcatattcacaataaaactttaaattagtagataaaacattttcaaaatcacatcaataaaacctaatacaattttttagcttcatcttactctttat
Y aatagatttcaaaatcatatttaatttaagcagatttttttttcaattttgtttaaatttttaaaaaataaaactaattcctattatttttctattaaaaatgtaagttgttgaaa

(C)

	M	Q	I	F	V	K	T	L	T	G	K	T	I	T	L	D	V	E	A	S	D	T	I	E	N	V	K	A	K	I	Q	D	K	E	G	I	P	P	D	Q	
A1	ATG	CAA	ATTTT	CGT	TAA	GAC	CCCT	C	A	CTG	GTAA	AACT	ATT	ACT	TTG	GAC	GTG	GA	AGC	CTC	TG	AC	ATT	TG	AA	AC	GTG	CA	AGG	CTA	AG	CA	AGG	A	AGG	TAT	CCCCC	CTG	ATTA		
A2	..T.	..C.	..C.	..T.	G.	..G.	..C.	..C.	..CC.	C.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
A3	..T.	..C.	..C.	..TT.	G.	..C.	..G.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
A4	..C.	..C.	..T.	G.	..G.	..C.	..C.	..CC.	C.	..T.	..T.	..C.	..C.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
A5	..C.	..C.	..T.	G.	..G.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
Bn	..T.	..C.	..C.	..T.	G.	..G.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
C1	..T.	..C.	..G.	..TT.	G.	..G.	..C.	..CC.	C.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
C2	..T.	..C.	..G.	..TT.	G.	..G.	..C.	..CC.	C.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
X1	..T.	..C.	..G.	..TT.	G.	..G.	..C.	..CC.	A.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
X2	..C.	..C.	..T.	G.	..G.	..C.	..CC.	T.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
Yn	..T.	..C.	..C.	..T.	G.	..G.	..C.	..CC.	T.	..T.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.	..C.
D1	..T.	..A.	..T.	G.	..A.	..A.	..G.	..G.	..C.	..C.	..T.	..C.	..CTT.	TA.	..T.	..AC.	..C.	..G.	..TA.	..T.	..TAG.	..A.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.
D2	..G.	..T.	..C.	..TT.	G.	..C.	..C.	..A.	..T.	..C.	..TAAA.	..T.	..AC.	..G.	..TA.	..TAA.	..A.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.	..T.
D3	..T.	..G.	..C.	..TA.	..C.	..TT.																																			

[illegible]

Gp. I (A-C)		1	10	20	30	40	50	60	70	76												
		MQIFVKTLTGKTITLDVEASDTIENVKAKIQDKEGIPPDQORLIFAGKQLEDGRITLSYNIQKESTLHLVLRGG																				
Gp. II	D1	V	AST N	I Q				V		V												
	D2	V	AQT N	I Q				V		V												
	D3	V	DSA N					V		V												
	D4	V	DS N					V		V												
	D5		NST N					SIR	S	*												
E1	E1	V	AST N	I Q		S		E	V	V												
	E2	V	AQT N	I		S		V		V												
	E3	V	ENS N	SI S		S		NI		V												
	E4	V	DS N			S		V		V												
	E5	L	INST N		T			V	S	*												
F1	F1	VA	SST N	DTI Q	Y	S		V		V												
	F2	V	AQT N	I		S		E	V													
	F3	V	IDS N			S		VT														
	F4	M	NSA N	K Q		V		K V	S	N												
	H1	V	AQT SV	I Q				A														
H2	H2		DS N					I														
	H3	LYL	YN	K				V	L	D												
	H4	V	DS EN					V														
	H4							S		*												
	H5		NST N					S		*												
Gp. III	G1	Y	A	I	N	M	E	TH	K	SL	N	LIE	N	NQ	Y	S	IQ	Q	EI	QS	V	
	G2	Q	A		TS	N	NDI	SM	S	S	EL	F	F	EE			L				-I	VS
	G3	K	R	-C	IF	E	DE	N	KT	T		S	LK	S	TK	EK	S	D	N	IN	I	*

Fig.III.3. Comparação das seqüências de aminoácidos. A seqüência de aminoácidos predita para a unidade A1 constitui a seqüência de referência. Todas as seqüências de aminoácidos previstas para as unidades dos genes do Grupo I (genes A, B e C) são idênticas à da unidade A1. Apenas os aminoácidos que diferem da seqüência A1 se encontram representados: os aminoácidos idênticos estão representados por pontos. Os travessões nas unidades do gene G indicam ausência de aminoácidos que foram introduzidas para manter o alinhamento. Os códons de terminação estão indicados por asteriscos. As caixas representam regiões onde se observam a maioria das substituições (ver texto).

Fig.III.2. (Páginas anteriores) Seqüência nucleotídica dos genes poliubiquitina em *T.pyroformis*. As seqüências dos genes A, B e C foram retiradas de Neves et al. (1988, 1991a). (A) Seqüência nucleotídica das regiões 5' não codificantes. A numeração seguida assumiu o valor -1 para o nucleótido imediatamente antes do códon de iniciação. O motivo AGAGA encontra-se em negrito e os possíveis elementos de choque térmico encontram-se em negrito e sublinhados. Os dois elementos de *consensus* estão sublinhados uma - *Consensus* 1 - ou duas vezes - *Consensus* 2. (B) Seqüência nucleotídica das regiões 3' não codificantes. A numeração seguida assumiu o valor +1 para o nucleótido imediatamente a seguir do códon de terminação. (C) Comparação entre as seqüências nucleotídicas das regiões codificantes. A unidade A1 foi tomada como seqüência de referência, sendo a sua seqüência de aminoácidos prevista indicada em cima. Apenas os nucleótidos que diferem da seqüência referência se encontram indicados: os nucleótidos idênticos encontram-se representados por pontos. Os travessões indicam ausência de nucleótidos que foram introduzidas de modo a melhorar o alinhamento entre as seqüências. Os códons de terminação encontram-se em negrito.

Nenhum destes motivos foi contudo encontrado nas promotoras dos genes poliubiquitina dos Grupos II ou III. Todavia, uma pesquisa por computador permitiu encontrar dois elementos *consensus* presentes em todos os genes do Grupo II (Tabela III.1). Estes elementos encontram-se localizados aproximadamente à mesma distância do codão de iniciação, encontrando-se um dos elementos igualmente no gene do Grupo III. Curiosamente, a sequência *consensus* 1 parece possuir um padrão regular, do tipo ATG(TCTN₃)₄. Contudo, o seu significado biológico permanece por esclarecer. Uma região de 70 nucleótidos com 89% de identidade, sem deleções ou inserções, foi igualmente encontrada a -312/-243 do gene D e a -329/-260 do gene F. Uma região equivalente mas com menor semelhança foi igualmente encontrada no gene E, mas não no gene H. A análise das regiões a jusante dos genes poliubiquitina (Fig.III.2B) não permitiu encontrar nenhum motivo significativo comum.

Tabela III.1. Motivos conservados na região não codificante 5' dos gene poliubiquitina dos Grupos II e III.

Gene	início	Sequência	fim
D	-207	ATG TCTA CTT TcTA TTT TCTA cTc TCat	-180
E	-204	ATG TCTA tTT cCgA TTT TCTA TTt TCTc	-177
F	-223	ATG TCTA CTT TCaA TTg TCTA TTc TCTt	-196
H	-216	ATG TCTA CTT TcTA TTa TCTA Tct TCTc	-189
Consensus 1		ATG TCTA CTT TCNA TTN TCTA TTY TCTY	
D	-155	ACTCAaAATcTATCgTATCTGAT	-133
E	-151	cCTCAaAATTTATaATATCTaAT	-129
F	-158	ACTCAaAATTCATCaTATa-GAT	-137
H	-156	ACTCAaAATTTATCgTATCTGAc	-134
G	-128	ACTCAaAATTTATCaaATCTGAT	-106
Consensus 2		ACTCAWAATTTATCRTATCTGAT	

2.3. Utilização de codões pelos genes poliubiquitina.

Tal como muitos organismos unicelulares, a *Tetrahymena* apresenta uma escolha preferencial típica entre os tripletos sinónimos que codificam para um determinado aminoácido (Barahona *et al.*, 1988; Martindale, 1989). Este padrão preferencial na utilização de codões é mais acentuado para alguns genes do que para outros. Genes muito expressos utilizam usualmente os nucleótidos T ou C na terceira posição dos codões, enquanto outros genes são mais ricos em AT, possivelmente devido ao baixo conteúdo em GC do genoma de *Tetrahymena* (Gorovsky, 1980). De modo a analisar a utilização de

codões para os diferentes grupos de genes poliubiquitina de *Tetrahymena*, a ocorrência de cada base em cada posição do codão foi calculada (Tabela III.2).

Tabela III.2. Frequência nucleotídica por posição de codão em cada grupo de genes poliubiquitina de *Tetrahymena*.

	1ª posição do codão			2ª posição do codão			3ª posição do codão		
	GpI	GpII	GpIII	GpI	GpII	GpIII	GpI	GpII	GpIII
%A	36.5	37.9	43.1	36.8	38.2	36.4	24.2	34.0	36.9
%C	16.4	12.8	12.0	20.1	18.5	19.1	34.2	14.6	14.2
%G	31.6	30.9	22.7	13.0	12.8	9.3	13.0	12.7	12.4
%T	15.5	18.4	22.2	30.0	30.4	35.1	28.6	38.8	36.4
%G+C	48.0	43.7	34.7	33.2	31.4	28.4	47.2	27.3	26.7

A acentuada preferência pelos nucleótidos TC na terceira posição do codão, onde representam 62.8% do total, é claramente visível nos genes do Grupo I. Mesmo os 37.2% observados para os nucleótidos AG são um reflexo da existência de codões que usam sempre A ou G na terceira posição, e que representam 28.5% do total. Ainda assim, este grupo de codões exhibe igualmente uma escolha preferencial, como é o caso dos codões codificantes para os aminoácidos Gln (sempre TAA ou CAA), Glu (sempre GAA), Arg (sempre AGA) e Lys (95.2% codificados por AAG). O conteúdo em GC na terceira posição dos codões dos genes do grupo I (47.2%) é o dobro do conteúdo global em GC do genoma de *Tetrahymena*, mas é semelhante ao encontrado em várias regiões codificantes (Martindale, 1989).

O conteúdo em GC na terceira posição dos genes do Grupo II (27.3%) ou do Grupo III (26.7%) é bastante mais próximo do encontrado no genoma de *Tetrahymena* (25%), o que sugere uma utilização de codões menos favorável. Este decréscimo é sobretudo devido a mutações de C em A ou T. De facto, o conteúdo em C de 34.2% observado no Grupo I baixa para 14.6% no Grupo II ou 14.2% no Grupo III. Contudo, esta perda de pressão selectiva só é observável na terceira posição do codão. O conteúdo em GC na segunda posição é aproximadamente igual nos três grupos, e ligeiramente superior ao encontrado no genoma de *Tetrahymena*. Os conteúdos em GC da primeira posição dos codões do Grupo II (43.7%) e do Grupo III (34.7%) são ainda comparáveis ao do Grupo I (48.0%), e claramente acima do conteúdo global em *Tetrahymena*. Estes valores são devidos sobretudo à elevada percentagem de G observada na primeira posição dos três grupos: a mutação de G em A ou T não ocorre nesta posição, diferentemente da mutação de C em A ou T observada para a terceira posição. Esta diferença não é devida a uma taxa de mutação global de G inferior à de C, uma vez que na terceira posição do codão ocorrem igualmente mutações de G em A: 95.2% das Lys do Grupo I são codificadas por AAG, mas o mesmo codão codifica apenas para 46.7% das Lys do Grupo II e 38.1% do Grupo III. Estes dados sugerem que as duas primeiras posições dos codões dos Grupos II e III estão

ainda sujeitas a uma pressão selectiva elevada, mas que esta pressão se atenuou para a terceira posição.

Tabela III.3. Índice de Adaptação de Codões (CAI) dos genes de *Tetrahymena*.

Nível	Genes ou famílias de genes	CAI
Alto	Calmodulinas ^a	0.81- 0.84
	Factores de elongação ^b	0.81
	Ubiquitina - gene A	0.81
	Mutases ^c	0.79
	Proteínas do citoesqueleto ^d	0.77 - 0.85
	Histonas ^e	0.73 - 0.77
	Proteínas ribossomais ^f	0.72 - 0.84
Médio	Proteínas do tipo calmodulina ^a	0.64 - 0.74
	Outras proteínas ^g	0.51 - 0.71
	Histona (variante) ^e	0.50
Baixo	Ubiquitina - gene D	0.37
	Ubiquitina - gene H	0.32
	Ubiquitina - gene E	0.31
	Ubiquitina - gene F	0.31
	Outras proteínas ^g	0.27 - 0.31
	Proteínas específicas da conjugação ^h	0.24 - 0.31
	Ubiquitina - gene G	0.23
	Antígenos de superfície ⁱ	0.20

^a Calmodulinas, 4 genes: calmodulina (X52242), TCBP-23 (J05227) e TCBP-25 (J05109) de *T.thermophila*, e calmodulina (D10521) de *T.pyriformis*.

^b Factores de elongação, 1 gene: EF-1 α (D11083) de *T.thermophila*.

^c Mutase, 1 gene: fosfoenolpiruvato mutase (M85236) de *T.pyriformis*.

^d Proteínas do citoesqueleto, 7 genes: actina (M13939) e duas β -tubulinas (L01415; L01416) de *T.thermophila*, e actina (X01195), uma α - (X12767) e duas β -tubulinas (X12768; X12769) de *T.pyriformis*.

^e Histonas, 7 genes: H1 (M14854), H2B-1 (X05543), H2B-2 (X05544), H1v (X15548), H3-1 (M87304), H4-1 (X00417) e H4-2 (X04755) de *T.thermophila*.

^f Proteínas ribossomais, 5 genes: S25 (Nielsen *et al.*, 1986), L21 (M37892), L29 (M76718) e L37 (M59428) de *T.thermophila*, e TUF53 (X56693) de *T.pyriformis*.

^g Outras proteínas: Médio, 7 genes: hemoglobina (D13920) de *T.pyriformis*, profilina (D00813), enzima tipo citrato sintetase (D90117), proteína HMG-B (M63425), proteína tipo antígeno F (M59429), fosfoglicerato cinase (X63528) e Ile-tRNA sintetase (M30942) de *T.thermophila*; Baixo, 2 genes: hemoglobina (D13919) e cisteína protease (L03212) de *T.thermophila*.

^h Proteínas específicas da conjugação, 2 genes: CnjB (X06462) e CnjC (M29447) de *T.thermophila*.

ⁱ Antígenos de superfície, 1 gene: SerH3 (M60425) de *T.thermophila*.

A diferente utilização de codões em *Tetrahymena* é igualmente bem caracterizada pelo uso do Índice de Adaptação de Codões (CAI - "Codon Adaptation Index"; Sharp & Li, 1987b). A Tabela III.3 apresenta uma tabela de CAI construída usando como genes de

referência genes ribossomais, de histonas, tubulinas, actinas e de factores de elongação, já usados como referência na construção de tabelas equivalentes para levedura, *E.coli* e ciliados (Sharp & Li, 1987b; Martindale, 1989). Como esperado, o gene A (TU20) do Grupo I tem um valor semelhante ao obtido para diversos genes muito expressos. Os outros genes incluídos neste grupo, embora incompletos, possuem valores que estão em acordo com o obtido para o gene A. Já os genes dos Grupos II e III apresentam valores que reflectem claramente a sua diferente utilização de codões, e são mais próximos dos obtidos para, por exemplo, os genes de antígenos de superfície ou para os genes específicos da conjugação de *Tetrahymena*.

3. Análise comparativa entre ubiquitinas de diferentes espécies.

As sequências ubiquitina de *Tetrahymena* e as de diversos organismos foram comparadas entre si, sendo o resultado apresentado na Fig.III.4 sob a forma de um dendrograma de homologia. Este dendrograma foi construído através de um emparelhamento progressivo de sequências e grupos de sequências, e permite visualizar melhor as relações de similaridade entre os diferentes grupos de ubiquitinas. De acordo com a figura, pode-se observar que as ubiquitinas do Grupo I (representadas por *Tpyr*) formam uma ramificação diferenciada que inclui outras 17 ubiquitinas de organismos tão distintos como a levedura ou o homem. Dentro desta ramificação as ubiquitinas apresentam uma elevada similaridade, com não mais de quatro substituições de aminoácidos entre si (ver Fig.II.4), o que sugere que a ubiquitina do Grupo I é funcionalmente equivalente às dos

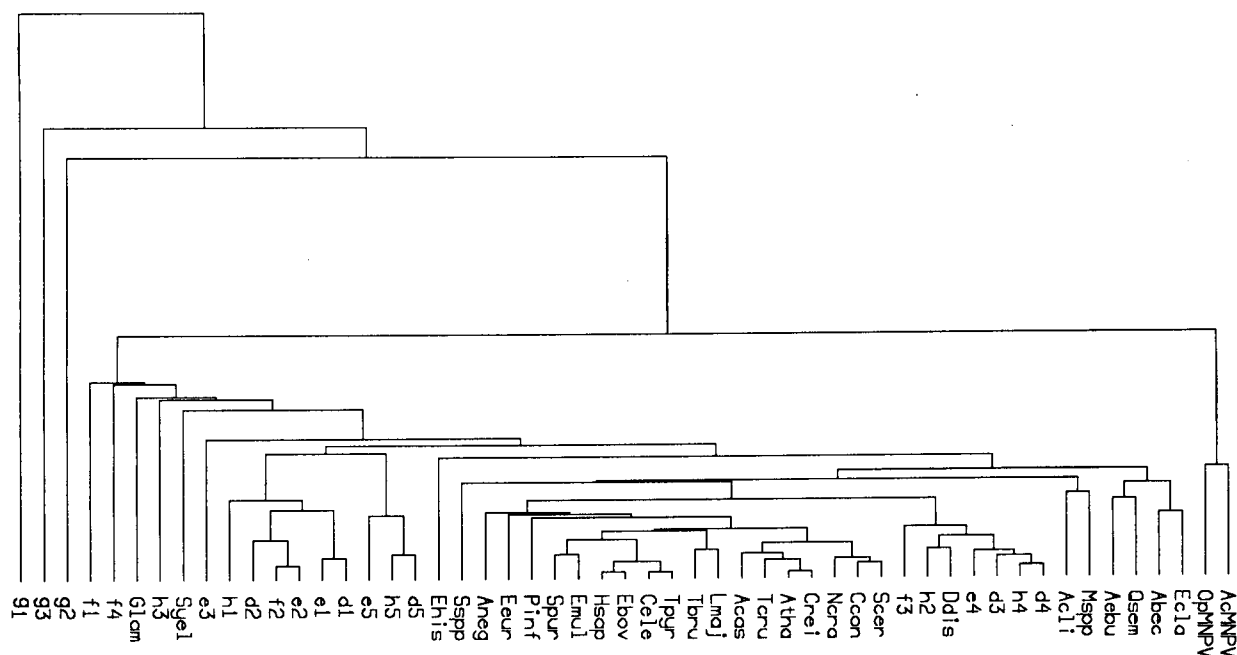


Fig.III.4. Dendrograma representando as relações de similaridade entre as diversas ubiquitinas. A distância ao longo do eixo vertical é proporcional às diferenças entre as sequências; a distância ao longo do eixo horizontal é arbitrária. As espécies representadas são as mesmas presentes na Fig.II.4, sendo aqui designadas pela primeira letra do gênero somada às primeiras três letras da espécie. OpMNPV e AcMNPV representam as ubiquitinas dos Baculovirus (ver Fig.II.6). *Tpyr* representa as ubiquitinas do Grupo I de *T.pyriformis*; as unidades dos Grupos II e III encontram-se representadas individualmente.

outros organismos. Curiosamente, a ubiquitina de *C.elegans* ou mesmo a humana são mais semelhantes à ubiquitina do Grupo I que a do único ciliado conhecido, *Euplotes eurystomus*. Contudo, isto não é reflexo de um maior parentesco da *Tetrahymena* com o homem, mas somente uma consequência do baixo grau de liberdade mutacional na sequência de ubiquitina, que impede o seu uso na construção de árvores filogenéticas.

As ubiquitinas do Grupo II apresentam um nível de semelhança mais baixo em relação às do Grupo I ou da maioria dos outros organismos. São ubiquitinas que se encontram extensivamente mutadas, mas cujas mutações se encontram na sua maioria localizadas em regiões delimitadas na molécula de ubiquitina (áreas sombreadas na Fig.III.3). Curiosamente, nestas regiões, que representam 27.6% da molécula, encontram-se não só a maioria das substituições do Grupo II (85.1%) como também a maioria das mutações das ubiquitinas eucariotas (81.0%; ver Fig.II.4) ou das dos Baculovirus (73.0%; ver Fig.II.6), o que sugere estarmos na presença de regiões que permitem um certo grau de liberdade mutacional.

As substituições observadas são predominantemente conservadas, estando entre as mais frequentes as substituições Ile3→Val, Val26→Ile, Pro38→Ser e Leu56→Val. São igualmente assinaláveis as alterações sofridas entre os resíduos 18 e 22, onde a sequência EAS é alterada frequentemente a DSS, NST ou AQT, e o resíduo Thr22 a Asn. Contudo, o resíduo Asp21, que se encontra envolvido numa ligação de hidrogénio na estrutura da ubiquitina (Vijay-Kumar *et al.*, 1987a), apenas é mutado uma vez apesar da elevada frequência de alterações da sua vizinhança. Algumas alterações são típicas da posição da unidade de ubiquitina no gene. Por exemplo, todas as primeira unidades possuem as substituições Ile13→Val, Val26→Ile, Ala28→Gln e Leu67→Val, enquanto a mutação Gln62→Ser ocorre somente nas últimas unidades. Outras substituições curiosas são a inexistência de mutações Ile3→Val nas primeiras ou últimas unidades, e o facto de somente os genes E e F possuírem substituições Pro38. Esta distribuição é facilmente observável no dendrograma da Fig.III.4, onde unidades com o mesmo número aparecem localizadas na mesma ramificação. No dendrograma, as unidades do Grupo II aparecem distribuídas por duas ramificações principais: uma que inclui a maioria das unidades nº 3 e 4, e que é a mais próxima da ramificação que inclui a ubiquitina do Grupo I, e uma segunda que inclui as restantes unidades. É curioso notar, contudo, que apesar de o Grupo II formar ramificações diferenciadas das dos outros organismos, existem três ubiquitinas eucariotas que se incluem nos seus ramos: a de *D.discoideum* na primeira ramificação, e as de *G.lambli*a e *S.yellowstonensis* na segunda ramificação.

Além das várias mutações conservadas observadas no Grupo II, é possível também notar a presença de algumas mutações não conservadas, que parecem ocorrer aleatoriamente. As unidades H3, F1 e F4 são as mais afectadas, o que se reflecte na sua posição no dendrograma, mas outras unidades mais "consensuais", como a unidade E5, apresentam igualmente este tipo de mutações. Nenhuma das últimas unidades possui a Gly terminal, um resíduo essencial para a activação da ubiquitina. A substituição deste resíduo por Ala inactiva a actividade da molécula, o mesmo sucedendo com versões encurtadas da ubiquitina por deleção do resíduo Leu73 (Ecker *et al.*, 1987), o que sugere que as últimas unidades ubiquitina do Grupo II são incapazes de conjugar.

As ubiquitinas do Grupo III não possuem, para além da última Gly, dois outros tripletos, tendo as restantes mutações espalhadas por todas as unidades e sendo por isso as ubiquitinas mais divergentes no dendrograma da Fig.III.4. Nenhuma das unidades possui a sequência terminal Gly-Gly-X, o que impede o seu processamento em monómeros de ubiquitina (Lopez-Otin *et al.*, 1989).

4. Análise da estrutura da molécula de ubiquitina de *Tetrahymena pyriformis*.

4.1. Perfis de hidropatia.

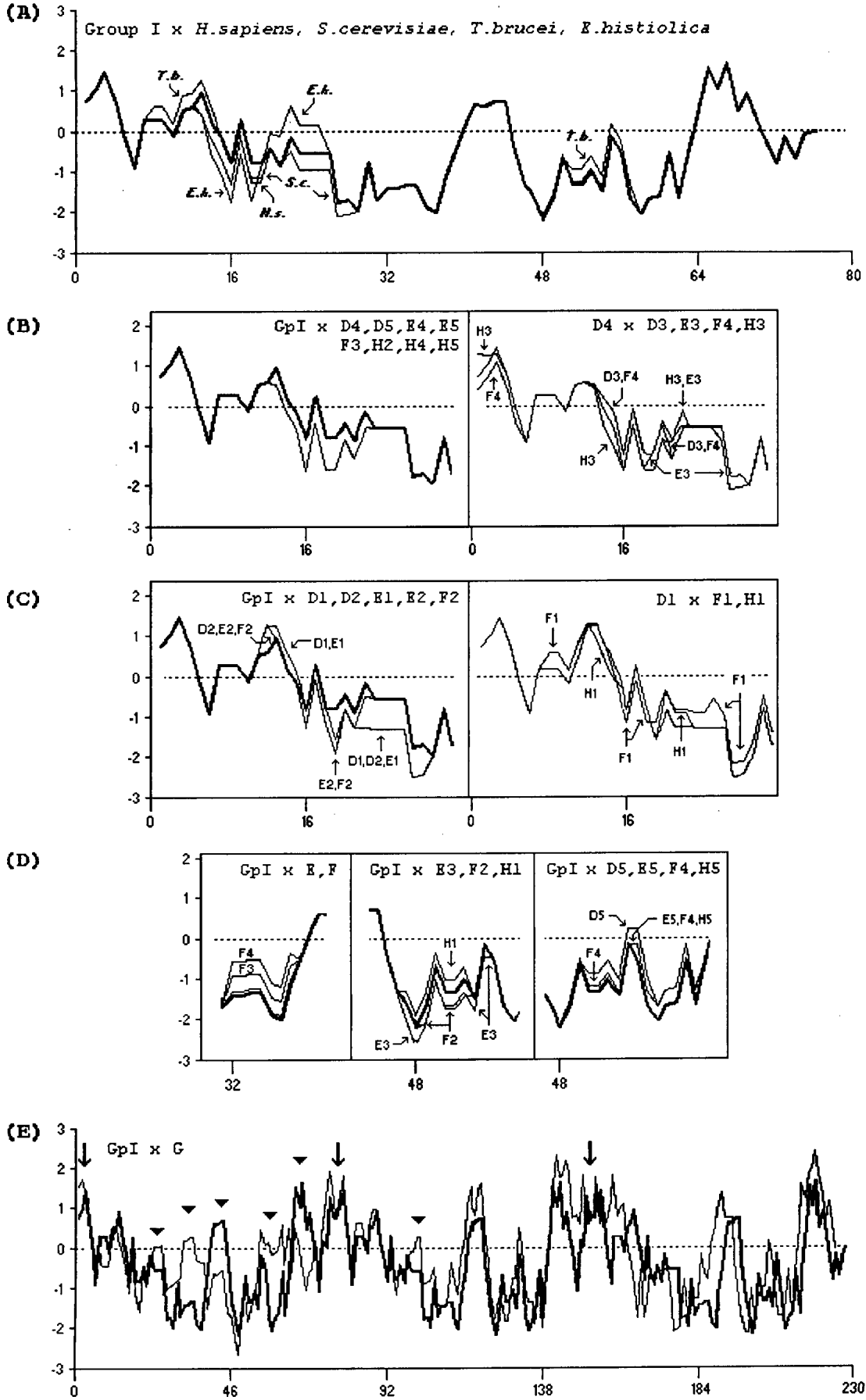
De modo a determinar possíveis alterações à estrutura da ubiquitina causadas pelas substituições observadas nas ubiquitinas de *T. pyriformis*, foram calculados e analisados os diferentes perfis de hidropatia. A comparação do perfil de hidropatia da ubiquitina do Grupo I com os perfis de alguns outros organismos (Fig.III.5A) não apresenta alterações marcantes entre os diversos perfis, o que sugere a inexistência de alterações estruturais significativas entre a ubiquitina do Grupo I e as dos outros organismos.

Uma vez que as ubiquitinas do Grupo II são consideravelmente mais divergentes, seriam de esperar diferenças mais acentuadas nos seus perfis. Contudo, a comparação destes com o perfil do Grupo I não apresenta alterações muito marcantes. De acordo com o padrão observado na região que compreende os resíduos 16 a 28, o Grupo II apresenta dois subgrupos com perfis semelhantes. O primeiro subgrupo, representado na Fig.III.5B, é constituído por 12 unidades, oito das quais possuem um perfil idêntico entre si nesta região (Fig.III.5B, quadro esquerdo), que, curiosamente, é muito semelhante ao perfil obtido para a levedura ou para o homem. As outras quatro unidades apresentam um perfil semelhante ao anterior, mas não idêntico (Fig.III.5B, quadro direito). As alterações são causadas por algumas substituições não conservadas, de que são exemplos os três primeiros resíduos de H3, ou a mutação Asn25→Lys em F4 e H3. O segundo subgrupo, representado na Fig.III.5C, contém 7 unidades, das quais só duas, E2 e F2, apresentam um perfil idêntico entre si. As outras cinco apresentam variações que parecem ser igualmente causadas por algumas mutações não conservadas observáveis nestas unidades.

Para além da região mencionada atrás, os perfis de hidropatia das ubiquitinas do Grupo II são idênticos ao do Grupo I com a excepção de alguns casos onde ocorrem algumas substituições não conservadas (Fig.III.5D). Nos genes E e F, a substituição Pro38→Ser não altera substancialmente o perfil, ao contrário das mutações Pro38→Ala (F3) ou Pro38→Ser (F4) (Fig.III.5D, quadro esquerdo). Na região que compreende os resíduos 52 a 57, apenas sete unidades apresentam alterações nos seus perfis de hidropatia (Fig.III.5D, quadros central e direito) que podem ser causadas por certas mutações como, por exemplo, His68→Asn em F4 ou Gly53→Glu em F2. Estes resultados sugerem que a maioria das substituições observadas nas ubiquitinas do Grupo II não é susceptível de alterar a estrutura da ubiquitina, embora esta possa ser afectada por algumas mutações não conservadas que parecem ocorrer aleatoriamente em algumas unidades.

O perfil de hidropatia do Grupo III (gene G) foi igualmente calculado. Contudo, uma vez que o produto do gene G não é susceptível de processamento em monómeros, o seu perfil foi comparado como um todo com uma hipotética triubiquitina do Grupo I (Fig.III.5E). Os perfis são claramente diferentes, mas apenas a primeira unidade apresenta transições acentuadas entre regiões hidropáticas e hidrofóbicas.

Fig.III.5. (Página ao lado) Perfis de hidropatia das unidades ubiquitina. Os perfis de hidropatia foram determinados de acordo com o método de Kyte e Doolittle (1982) usando-se uma janela de sete resíduos. (A) Desvios observados entre a ubiquitina do Grupo I de *T. pyriformis* e a de diversos organismos. (B) Perfis do primeiro subgrupo do Grupo II na região compreendendo os resíduos 16 a 28: à esquerda, desvio observado com a ubiquitina do Grupo I, e à direita desvios observados dentro do subgrupo. (C) Idem para o segundo subgrupo do Grupo II. (D) Desvios observados entre a ubiquitina do Grupo I e as do Grupo II nas regiões contendo o resíduo Pro38 (esquerda) e os resíduos 52 a 57 (centro e direita). (E) Desvios observados entre uma hipotética triubiquitina do Grupo I e o produto do gene G. As setas indicam a localização da metionina inicial de cada unidade, e as pontas de setas as regiões onde se observa uma inversão dos valores do perfil.



4.2. Estrutura terciária.

A ubiquitina é uma molécula compacta e globular com um núcleo hidrofóbico e uma extensão C-terminal protuberante que participa na conjugação ubiquitina-proteína (Vijay-Kumar *et al.*, 1987a). Esta estrutura é muito estável a diversos valores de pH e/ou de temperatura (Low *et al.*, 1979), e é aparentemente única para ubiquitinas de levedura, plantas ou de mamíferos (Vijay-Kumar *et al.*, 1987b). Mesmo as substituições observadas na ubiquitina divergente de *E.histolytica* encontram-se localizadas na mesma região da estrutura terciária, estando a maioria exposta à superfície da molécula (Wöstmann *et al.*, 1992).

De modo a determinar o modo como as mutações existentes nas ubiquitinas do Grupo II poderão afectar a estrutura típica da molécula de ubiquitina, as principais substituições foram introduzidas numa imagem estereo da estrutura terciária da ubiquitina produzida pelo programa Turbo-FRODO (Jones, 1978). A Fig.III.6A representa um esqueleto de carbono α onde foram introduzidas as mutações encontradas na unidade E2, por esta unidade possuir as oito mutações mais frequentes de todas as unidades do Grupo II. Uma vez que todas as primeiras unidades de cada gene do Grupo II possuem várias mutações únicas (Ile13→Val, Ala28→Gln e Leu67→Val), a Fig.III.6B representa uma imagem estereo equivalente para a unidade E1. Sete das 12 posições observadas são externas e, tal como para a *E.histolytica*, localizam-se na mesma região da superfície da molécula. O facto mais interessante é observado para as substituições internas: quatro das cinco substituições observadas - Val3, Ile26, Val56 e Val67 - localizam-se na mesma bolsa hidrofóbica interna. Contudo, não se encontram suficientemente próximas para interagirem entre si, e três modificam-se para um resíduo aminoácido menos volumoso. A quinta substituição - Val13 - é igualmente para um resíduo menos volumoso e localiza-se muito próximo dos resíduos da folha β 40-45, cuja posição poderá ser afectada pela substituição Pro38→Ser localizada na sua "reverse turn".

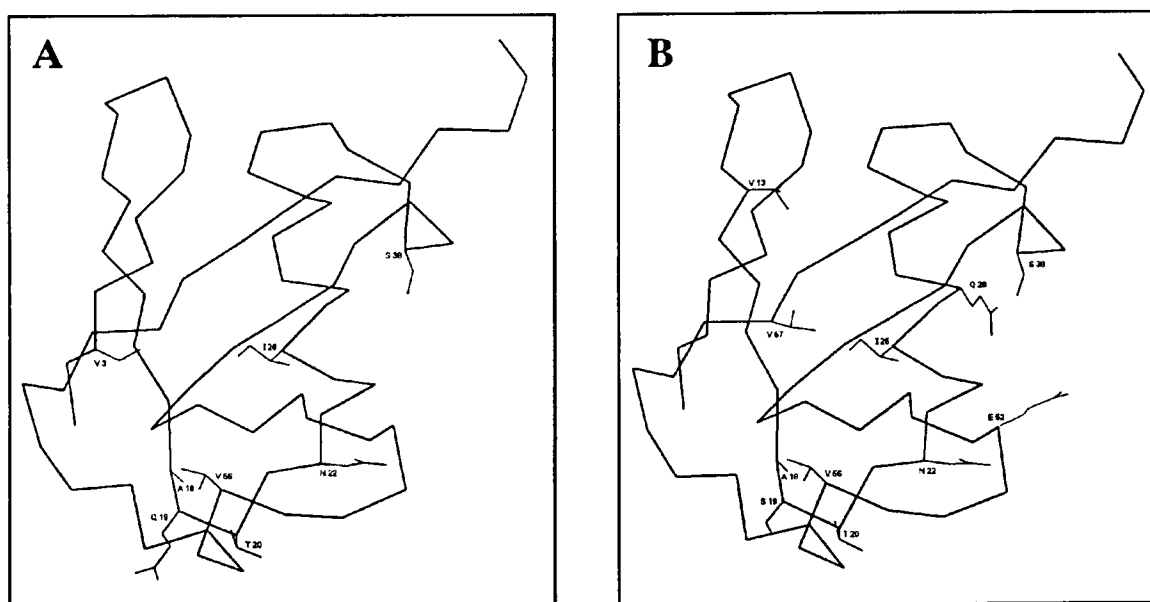


Fig.III.6. Imagem estereo das diferenças entre a estrutura terciária do esqueleto de carbono α da ubiquitina do Grupo I e a (A) ubiquitina E2; (B) a ubiquitina E1. De notar que as substituições internas Val3, Ile26, Val56 e Val67 se encontram localizadas numa mesma bolsa hidrofóbica e que as substituições externas Ala18, Ser/Gln19, Thr20, Asn22, Gln28, Ser38 e Glu52 se encontram expostas na mesma superfície externa.

5. Expressão dos genes poliubiquitina.

O padrão de expressão dos genes ubiquitina em células de *T.pyriformis* em condições normais de crescimento ou sujeitas a um "stress" hipertérmico encontra-se já descrito (Neves *et al.*, 1988; ver Fig.IV.4A). As análises por Northern de RNAs preparados a partir de células em crescimento exponencial apresentam três bandas, com um tamanho aparente de 0.75, 1.4 e 1.8 kb, sendo a de 1.4 kb dificilmente detectável. Quando as células são sujeitas a um "stress" hipertérmico, observa-se um aumento da intensidade da banda de 1.8 kb e o aparecimento de uma nova banda com cerca de 1.6 kb.

O uso de sondas específicas permitiu identificar os transcritos de três dos genes poliubiquitina pertencentes ao Grupo I. Assim, o gene B transcreve um mRNA de 1.6 kb em células sujeitas a "stress", enquanto ambos os genes A e C transcrevem mRNAs de 1.8 kb induzíveis por "stress" (Neves *et al.*, 1991a; 1991c). Cada banda observável por Northern pode então corresponder a transcritos de vários genes poliubiquitina.

Para analisar a expressão dos genes poliubiquitina dos Grupos II e III, usaram-se sondas específicas para cada um dos genes D a H, quer sob a forma de oligômeros sintéticos, quer sob a forma de fragmentos de regiões não codificantes. A especificidade destas sondas foi testada por análise em Southern (# 6), tendo sido determinadas igualmente as condições ideais de hibridação. Contudo, não foi obtido nenhum sinal quando estas sondas foram usadas para análise de Northern "blots" contendo mRNA-poli(A)⁺ de células em crescimento exponencial ou sujeitas a "stress" hipertérmico, o que sugere que estes genes não serão expressos nestas condições de crescimento. O mesmo resultado foi obtido em ensaios de "run-on" (Fig.III.7), uma técnica que permite detectar a transcrição de RNAs no núcleo, mesmo que estes não venham a ser processados e transportados para o citoplasma. O uso de fragmentos contendo apenas regiões não codificantes não é aparentemente muito eficaz neste tipo de análise, provavelmente por as regiões "leader" dos mensageiros de *Tetrahymena* serem habitualmente muito pequenas e ricas em nucleótidos A ou U. A "leader" do transcrito do gene A possui 108/114 nucleótidos (Neves *et al.*, 1991a), e deveria sobrepor-se em 81/87 nucleótidos ao fragmento não codificante usado na análise, mas a ausência de sinal sugere que esta sobreposição não é suficiente para se obter hibridação. Contudo, quando a região codificante deste gene foi usada obtiveram-se sinais fortes, observando-se uma indução da transcrição nas células previamente sujeitas a um "stress" hipertérmico. Apesar disso, não foram detectados nenhuns sinais quando as regiões codificantes dos genes F, G e H foram usadas na análise.

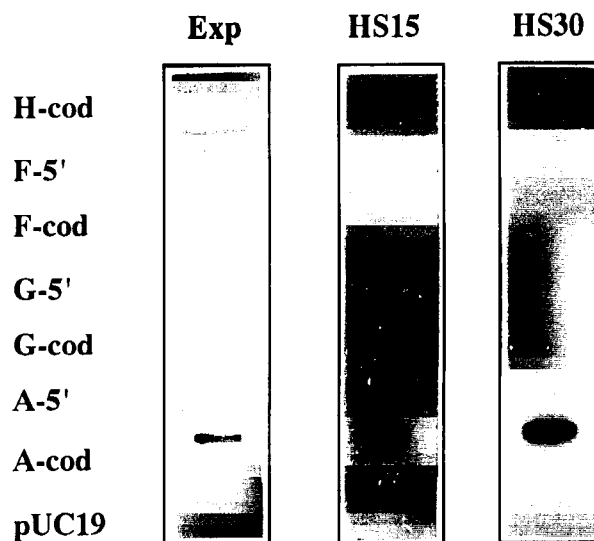


Fig.III.7. Ensaios de transcrição aparente ("run-on"). mRNAs isolados de núcleos incubados com uma mistura de transcrição contendo [α -³²P]UTP foram hibridados com DNA previamente ligado a nitrocelulose. A-5', F-5' e G-5' correspondem a fragmentos de DNA das regiões 5' não codificantes enquanto que A-cod, F-cod, G-cod e H-cod incluem região codificante (ver Tabela M2). O pUC19 linearizado foi usado como controlo. Exp, núcleos de células em crescimento exponencial; HS15 e HS30, núcleos de células sujeitas a um "stress" hipertérmico de 15 e 30 min.

6. Organização genómica dos genes poliubiquitina de *Tetrahymena pyriformis*.

A disposição dos genes poliubiquitina nos fragmentos de DNA sequenciados sugere uma organização em que todos os genes estão arrançados em série e separados por regiões intergénicas de 1.5 a 2.5 kb. De modo a caracterizar melhor a organização genómica destes genes em *T. pyriformis*, o DNA genómico deste organismo foi digerido com a enzima de restrição *ClaI*, separado em geis de agarose e transferido para filtros, sendo os fragmentos obtidos analisados por hibridação com o fragmento codificante *HindIII* de 0.23 kb do clone TU20 (Fig.III.8 - primeira pista). A enzima *ClaI* foi escolhida por não ser possível existir nenhuma sequência de reconhecimento na região codificante dos genes do Grupo I, e porque reconhece locais contidos nas regiões intergénicas dos clones TU1, TU3 e TU19H. As sete bandas observadas indicam a provável existência de pelo menos sete genes ubiquitina. Contudo, uma vez que não existem locais reconhecidos pela enzima *ClaI* entre os genes B e C ou entre os genes F e G (Fig.III.1), deverão existir no mínimo nove genes ubiquitina em *T. pyriformis*.

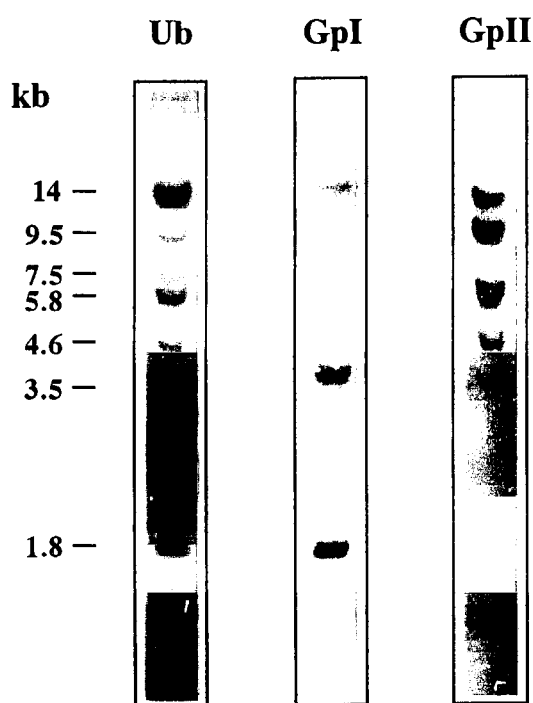


Fig.III.8. Detecção de sequências ubiquitina em DNA de *T. pyriformis*. Aliquotas de DNA macronuclear foram digeridas com *ClaI*, analisadas em geis de agarose a 1% e transferidas para filtros de nylon. A hibridação foi efectuada com sondas para ubiquitina (Ub), e em condições restritivas com sondas específicas para ubiquitinas dos Grupos I (GpI) e II (GpII).

14 kb contém três genes (B, C e D), o fragmento de 9.5 kb contém dois (E e F), e o fragmento de 3.5 kb contém o gene Y e a região a montante do gene X. A banda de 1.8 kb, que foi clonada como TU20 (Fig.III.1), contém o gene A, e a banda de sinal pouco intenso de 7.5 kb tinha já sido identificada anteriormente como contendo o gene de fusão ubiquitina-proteína de extensão (Neves *et al.*, 1991b). Como o fragmento de 4.6 kb contém o gene H, só o de 5.8 kb não contém nenhum gene conhecido.

De modo a caracterizar a distribuição dos genes poliubiquitina dos Grupos I e II, digestões de DNA genómico com a enzima *ClaI* foram analisadas por hibridação em condições mais restritivas com uma sonda codificante do Grupo I correspondente ao fragmento *HindIII* de 0.23 kb do clone TU20, e com uma sonda codificante do Grupo II correspondente ao fragmento *EcoRI* de 0.23 kb do clone TU1 (Fig.III.8). Os resultados mostram que a sonda do Grupo I hibrida preferencialmente com as bandas *ClaI* de 1.8, 3.5, 7.5 e 14 kb, enquanto a sonda do Grupo II hibrida preferencialmente com as bandas *ClaI* de 4.6, 5.8 e 14 kb. Contudo, é possível que a banda de 5.8 kb possua igualmente genes do Grupo I.

A identificação dos fragmentos *ClaI* que contêm os diferentes genes ubiquitina clonados foi realizada por hibridação de digestões de DNA genómico com sondas específicas para cada gene poliubiquitina. De acordo com os resultados compilados na Tabela III.4, o fragmento *ClaI* de

Tabela III.4. Tamanho dos fragmentos de DNA que contêm genes poliubiquitina. DNA genómico de *T.pyriformis* foi digerido com as enzimas de restrição indicadas e hibridado com fragmentos de DNA específicos dos genes A a Y. Os dados representam o tamanho aparente das bandas de hibridação observadas.

Enzima	Genes poliubiquitina								
	A	B	C	D	E	F	H	X-5'	Y
<i>Cla</i> I	1.8	14	14	14	9.5	9.5	4.6	3.5	3.5
<i>Cla</i> I/ <i>Bam</i> HI	1.8	4.1	10	14	2.0	5.2	4.6	3.5	3.5
<i>Eco</i> RI	n.d.	8.4	n.d.	n.d.	n.d.	n.d.	n.d.	8.4	n.d.

n.d.: Não determinado.

Os dados obtidos com as digestões *Cla*I permitem já um primeiro arranjo esquemático da organização dos genes poliubiquitina em *T.pyriformis*. Por exemplo, uma vez que os genes D e E do fragmento TUλ19H hibridam com os fragmentos *Cla*I de 14 e 9.5 kb e os genes F e H do fragmento TU1 com os de 9.5 e 4.6 kb, estes três fragmentos podem então ser ordenados 14-9.5-4.6 (Fig.III.9A). De modo a clarificar a organização, realizaram-se digestões do DNA genómico com diversas enzimas ou combinações de enzimas, que foram igualmente hibridadas com sondas específicas para os genes poliubiquitina. A Tabela III.4 contém os resultados mais significativos dessas hibridações. De acordo com estes, os genes B e X reconhecem um mesmo fragmento *Eco*RI de 8.4 kb, o que indica que o gene X pertence igualmente ao fragmento *Cla*I de 14 kb, e que o fragmento de 3.5 kb lhe está associado (Fig.III.9B). O resultado mais surpreendente é dado contido pela digestão dupla *Bam*HI/*Cla*I do DNA genómico. Uma vez que a enzima *Bam*HI reconhece um local entre os genes B e C (Fig.III.1), serão de esperar duas bandas diferentes para cada gene, e, de facto, o gene B reconhece um fragmento *Bam*HI/*Cla*I de 4.1 kb enquanto o gene C reconhece um fragmento de 10 kb. Contudo, embora o gene D devesse estar incluído num destes fragmentos, o resultado da hibridação apresenta novamente uma banda de 14 kb, o que sugere que o DNA genómico de *T.pyriformis* contém não um mas dois fragmentos *Cla*I de 14 kb com genes poliubiquitina. Estes dois fragmentos não são resolvidos por geis de electroforese de campo eléctrico alternado ("Pulse Field Gel Electrophoresis"), pelo que a diferença de tamanho entre eles deverá ser muito reduzida (resultados não apresentados).

Estes resultados evidenciam uma separação física entre os genes poliubiquitina do Grupo I e os dos Grupos II e III no genoma de *T.pyriformis*, onde existirão pelo menos duas grandes regiões contendo genes poliubiquitina (Fig.III.9). A região que compreende os fragmentos *Cla*I 3.5-14 kb (Fig.III.9B) contém quatro dos genes do Grupo I (B, C, X e Y), que deverão ser os únicos existentes, já que as diferenças de intensidade nos sinais de hibridação obtidos com a digestão dupla *Bam*HI/*Cla*I indicam que não deverão existir mais genes a jusante do gene C (resultados não apresentados). Muito possivelmente, o fragmento TU4 terá continuidade no fragmento TU3, por ser a única possibilidade de se incluir um gene Y de 3-4 unidades no fragmento *Cla*I de 3.5 kb. O quinto gene do Grupo I, o gene A, deverá encontrar-se numa região independente do genoma, como sugerido por algumas análises em Southern (resultados não apresentados). A segunda grande região compreende os fragmentos *Cla*I 14-9.5-4.6 kb e contém todos os genes poliubiquitina dos Grupos II e III conhecidos (D a H; Fig.III.9A). Contudo, de acordo com a análise efectuada ao fago

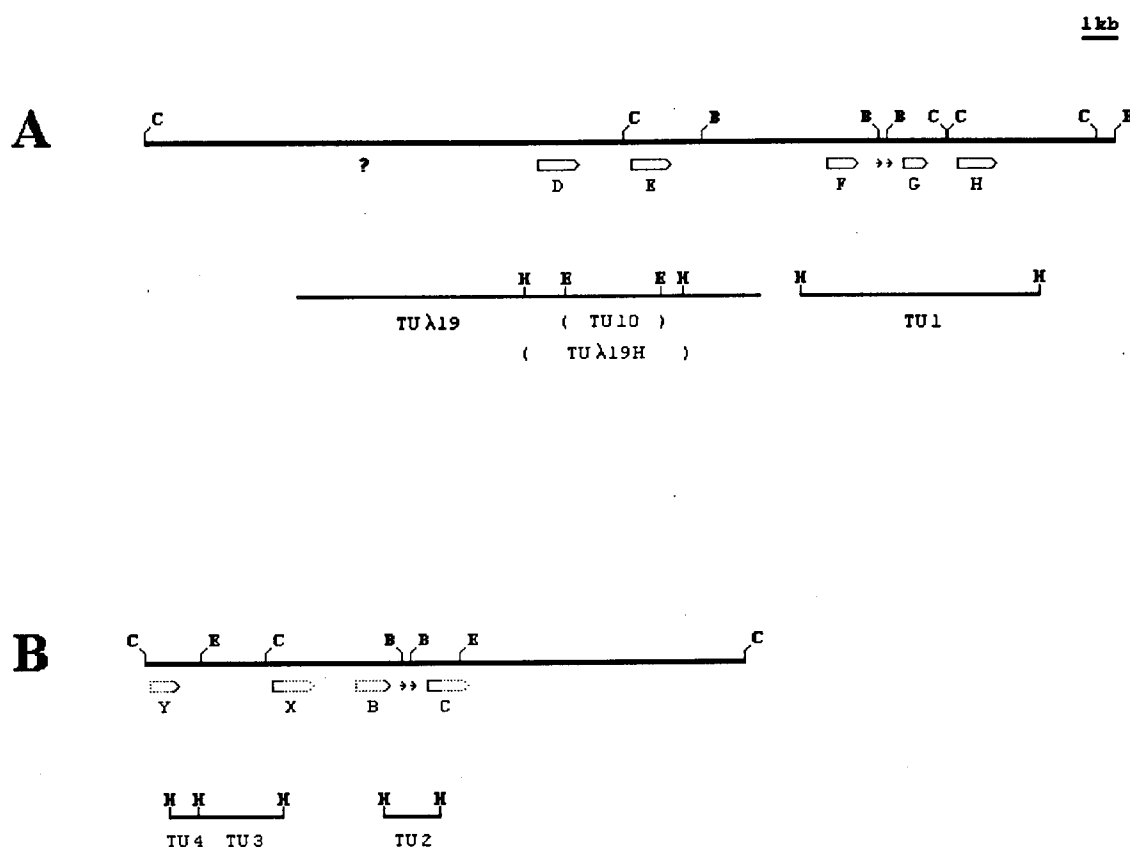


Fig.III.9. Diagrama esquemático da organização dos genes poliubiquitina em *T.pyriformis*. (A) Região contendo os genes poliubiquitina dos Grupos II e III. Esta região compreende os fragmentos *Cla*I de 14 (II), 9.5 e 4.6 kb, contém os genes D a H e engloba os fragmentos clonados TU1 e TUλ19 (este último contendo os fragmentos TU10 e TUλ19H). O ponto de interrogação corresponde a um ou mais genes detectados por hibridação em TUλ19 mas não sequenciados. (B) Região contendo os genes poliubiquitina do Grupo I. Esta região compreende os fragmentos de 3.5 e 14 (I) kb, contém os genes Y, X, B e C e engloba os fragmentos clonados TU2, TU3 e TU4. O gene A deverá encontrar-se numa região distinta do genoma (ver texto), e a localização do fragmento *Cla*I de 5.8 kb não é conhecida. As setas abertas indicam os genes poliubiquitina, e as setas indicam genes de 5S-rRNAs. B-*Bam*HI; C-*Cla*I; E-*Eco*RI; H-*Hind*III.

TUλ19, é de esperar pelo menos mais um gene poliubiquitina a montante do gene D (resultados não apresentados). Juntamente com pelo menos um gene localizado no fragmento *Cla*I de 5.8 kb, ao qual não é possível atribuir uma localização, o Grupo II terá um mínimo de seis genes poliubiquitina.

Em conclusão, o genoma de *T.pyriformis* contém um mínimo de 11 genes poliubiquitina, sendo cinco do Grupo I, seis (a nove) do Grupo II e um do Grupo III. Quatro dos genes do Grupo I encontram-se organizados em série numa região comum do genoma, enquanto o quinto deverá ter uma localização distinta. Por seu lado, todos os genes dos Grupos II e III (com a eventual excepção dos localizados no fragmento *Cla*I de 5.8 kb) encontram-se igualmente organizados em série numa região comum, e distinta da do Grupo I.

7. Discussão.

A presença de genes poliubiquitina em *T.pyriformis* foi já previamente demonstrada (Neves, 1991). Os resultados obtidos evidenciam, contudo, que os genes poliubiquitina deste organismo se organizam numa família alargada que pode ser dividida em três grupos de sequências relacionadas entre si.

O primeiro grupo, Grupo I, inclui os genes poliubiquitina mais conservados, codificando todas as suas unidades para proteínas idênticas. Estas proteínas são muito semelhantes às encontradas noutros organismos, não apresentando mais de quatro substituições em relação a um conjunto de ubiquitinas de organismos evolucionariamente tão distantes como fungos, plantas e animais. Existe uma preferência clara no uso de alguns codões, indicando que estes genes são muito expressos. De facto, os genes deste grupo originam pelo menos um mensageiro de 1.6 kb e dois de 1.8 kb (Neves, 1991). A sua transcrição é regulada por agentes de stress, como demonstrado pela presença de elementos de choque térmico nas suas regiões promotoras e pela indução dos mensageiros por "stress" hipertérmico. É pois plausível que este grupo de genes poliubiquitina tenha um papel equivalente ao descrito para as ubiquitinas de outros organismos.

O segundo grupo de genes, Grupo II, apresenta várias particularidade pouco usuais, como sejam o número pouco habitual de substituições, a divergência entre as repetições de cada gene e a ausência do resíduo Gly76 na última repetição de cada gene. Contudo, todas as análises efectuadas indicam que estes genes não podem ser pseudogenes com origem em genes do Grupo I. A ser assim, a distribuição das unidades nestes genes (com as primeiras unidades muito semelhantes entre si, as segundas igualmente, etc; Fig.III.4) sugere a existência de um pseudogene inicial que acumulou as mutações observadas, e uma duplicação muito recente deste nos genes actuais. Ora, a comparação entre as regiões 5' e 3' não codificantes destes genes não apresenta homologias significativas, o que invalida uma duplicação recente. Também não são encontradas homologias entre estas mesmas regiões e as regiões equivalentes dos genes do Grupo I, pelo que não deverá igualmente ter ocorrido uma separação recente destes grupos. Contudo, uma separação antiga com imediata conversão dos genes do Grupo II a pseudogenes não é compatível com o elevado grau de homologia ainda observável entre as unidades de ambos os grupos (mais de 80% de identidade nas sequências de aminoácidos). Na verdade, esta é uma das inúmeras linhas de evidência que sugerem que os genes do Grupo II têm vindo a ser sujeitos a uma pressão selectiva durante um longo período. Os pseudogenes possuem elevadas taxas de substituições de nucleótidos, mas sem nenhuma selecção da posição nos codões (Li *et al.*, 1981), o que se traduz num elevado número de substituições de aminoácidos. No Grupo II, observa-se de facto uma taxa de substituições de nucleótidos muito elevada, mas que afecta predominantemente a terceira posição dos codões, mesmo em regiões muito conservadas no Grupo I. Assim, isto traduz-se numa utilização de codões pouco favorável e num número reduzido de substituições de aminoácidos. Mais ainda, as mutações de aminoácidos não se encontram aleatoriamente distribuídas, como esperado em pseudogenes, mas localizadas nas mesmas regiões da sequência primária da ubiquitina onde se observam a maioria das variações entre as ubiquitinas dos diversos eucariotas estudados. Estas substituições não afectam o perfil hidrofóbico da molécula (Fig.III.5), e a sua localização na estrutura terciária da ubiquitina revela que se situam na mesma superfície exterior da molécula, ou na mesma bolsa hidrofóbica interior (Fig.III.6).

Apesar destes dados, uma análise mais detalhada sugere igualmente que os genes poliubiquitina do Grupo II de *T.pyriformis* podem não ser funcionais. É possível observar,

para além de um padrão de substituições conservativas e não aleatórias, várias substituições não conservativas cuja distribuição aparenta ser aleatória. Estas substituições, que se acentuam mais numas unidades do que noutras mas que ocorrem em todos os genes deste grupo, podem afectar resíduos envolvidos na estrutura da ubiquitina, como sugerido por alguns dos perfis de hidropatia obtidos (Fig.III.5). Sobretudo, não foi detectada transcrição associada a estes genes (Fig.III.7). Embora não possuam codões de terminação internos ou percam a grelha de leitura, este tipo de mutações pode ser suficiente para justificar a sua não funcionalidade.

Em conjunto, estes dados sugerem que os genes do Grupo II serão muito possivelmente pseudogenes, embora não relacionados com os genes do Grupo I. A sua origem é ainda uma incógnita, mas poderá estar relacionada com as singularidades próprias dos ciliados e em particular da espécie *T.pyriformis*, um ciliado que ao contrário da maioria das espécies do género não possui micronúcleo (ver Capítulo I2). A perda do micronúcleo, apesar de constituir um fenómeno natural, não é reprodutível em laboratório sem que ocorra em simultâneo uma captura de DNA micronuclear pelo macronúcleo (Karrer *et al.*, 1984). É pois possível que os genes poliubiquitina do Grupo II de *T.pyriformis* sejam um resultado dos acontecimentos associados à perda inicial do micronúcleo. Sendo o macronúcleo poliplóide, os genes resultantes poderiam ter escapado a uma selecção e ter acumulado mutações que seriam letais em cópias únicas. Este fenómeno ocorre, por exemplo, com um gene de histona H3 de *Xenopus*, que não possui igualmente codões de terminação internos ou perda da grelha de leitura, mas que apresenta cinco substituições de aminoácidos únicas. Crê-se que este gene é na realidade um pseudogene que foi mantido devido à tetraploidização parcial do genoma de *Xenopus* (Wells *et al.*, 1986). É possível que um fenómeno semelhante tenha ocorrido no genoma de *T.pyriformis*, o que é apoiado por alguns dados referentes à organização dos genes no genoma deste organismo. Quando o genoma de *T.pyriformis* é separado por geis de electroforese em campo eléctrico ortogonal alterno, a hibridação com uma sonda de ubiquitina apresenta uma única banda, ao contrário do observado com outras espécies de *Tetrahymena* (Conover & Brunk, 1986). Como aparentemente os genes do Grupo I e do Grupo II estão fisicamente separados no genoma (# 6), ou ambos os grupos estão contidos em regiões distintas no mesmo fragmento macronuclear, ou, alternativamente, existem dois fragmentos "alélicos" com o mesmo tamanho relativo, um contendo os genes do Grupo I e o outro os genes do Grupo II. A presença de bandas *Cla*I com o mesmo tamanho relativo (14 kb, Tabela III.4) mas contendo genes distintos suporta esta última hipótese. Este fenómeno pode mesmo ser frequente em *T.pyriformis*, pois os dois genes β -tubulina encontrados neste organismo parecem localizar-se igualmente em dois fragmentos *Bam*HI distintos mas de tamanho relativo idêntico (L.Cyrne, comunicação pessoal). Assim, os fragmentos "alélicos" teriam origem nos rearranjos causados pela perda do micronúcleo, e as eventuais mutações sofridas num dos fragmentos mantidas pela poliploidização. Contudo, esta hipótese não explica a diferente utilização de codões observada entre os genes do Grupo II e os do Grupo I, mesmo em regiões onde não ocorrem substituições de aminoácidos.

É igualmente possível que os gene poliubiquitina do Grupo II constituam um ramo antigo de ubiquitinas divergentes, retido ainda em alguns eucariotas inferiores. Nos últimos dois anos, têm sido encontradas ubiquitinas muito divergentes em alguns organismos, como *G.lamblia* ou em protozoários sarcodinos, que em alguns casos estão mais próximas de algumas unidades do Grupo II do que das ubiquitinas canónicas (Fig.III.4). Estas ubiquitinas, que em alguns casos são as únicas ubiquitinas existentes no organismo (Krebber *et al.*, 1994), podem possuir funções desconhecidas que se foram perdendo nos eucariotas

superiores. Em *Tetrahymena*, estas funções podem estar relacionadas com o micronúcleo. Uma vez que a *T.pyriformis* perdeu o seu micronúcleo, as poliubiquitinas do Grupo II podem ter igualmente perdido a sua função e iniciado a acumulação das mutações aleatórias e não conservativas observadas. Esta hipótese justificaria porque, ao contrário dos genes do Grupo I, os genes do Grupo II têm uma utilização de codões menos preferencial, mais próxima da encontrada, por exemplo, em genes específicos do processo de conjugação (Martindale, 1990; Martindale & Taylor, 1988). Uma das funções propostas para as ubiquitinas divergentes encontradas em Baculovirus, que são proteínas expressa no período tardio do ciclo viral, foi um possível papel na inibição da transcrição e tradução dos genes do hospedeiro que se observa na fase tardia do processo de infecção (Guarino, 1990). As poliubiquitinas do Grupo II podem eventualmente possuir um papel equivalente na inibição da transcrição do micronúcleo durante a fase vegetativa do ciclo das células de *Tetrahymena*, ou na inibição da transcrição do macronúcleo durante a fase germinativa.

O Grupo III apresenta as ubiquitinas mais divergentes encontradas até agora. O número elevado de mutações observado em conjunto com a perda do resíduo Gly76 terminal sugerem que o gene G pode ser um pseudogene muito antigo do Grupo II. Não foi detectado nenhum sinal de transcrição para este gene, sugerindo igualmente que não é funcional. Contudo, apesar de extensamente mutado, o gene não apresenta codões de terminação internos ou perda da grelha de leitura, como acontece frequentemente em pseudogenes. O facto de as unidades de ubiquitina deste gene não possuírem a sequência *consensus* necessária ao processamento da poliubiquitina levanta a hipótese de este gene poder codificar para uma proteína de 225 aminoácidos expressa sob condições fisiológicas muito específicas. Este tipo de proteínas relacionadas com a ubiquitina têm sido frequentemente encontradas nos mais diversos organismos (ver Capítulo I1 - # 4.2). Por exemplo, a UCPR é uma proteína de 15kDa com duas regiões similares à ubiquitina que é induzida em células humanas apenas na presença de interferão β (Haas *et al.*, 1987). Mas contrariamente a esta proteína (Loeb & Haas, 1992), o produto hipotético do gene G não pode conjugar com outras proteínas usando o sistema ubiquitina, uma vez que não possui uma região C-terminal canónica.

Capítulo IV

RESULTADOS 2

Genes poliubiquitina em *T.thermophila*



1. Análise de sequências ubiquitina no DNA genómico de *T.thermophila*.

A presença de genes poliubiquitina divergentes em *T.pyriformis* levantou questões sobre a sua origem e/ou função cujo esclarecimento poderia passar pela análise das sequências ubiquitina noutros ciliados do género *Tetrahymena*. Sendo *T.pyriformis* uma espécie amiconucleada, uma eventual exclusividade dos genes poliubiquitina dos Grupos II e III neste organismo sugeriria uma sua aquisição posterior à perda do micronúcleo, podendo mesmo ter sido esta perda a causa do aparecimento destes genes. No caso destes genes existirem igualmente em espécies micronucleadas, é possível então que possam possuir funções necessárias a estes organismos, eventualmente relacionadas com o micronúcleo. Dentro das espécies micronucleadas, a espécie *T.thermophila* é a candidata óbvia a modelo biológico, já que se inclui no chamado complexo *T.pyriformis* (Nanney & McCoy, 1976), e é o ciliado *Tetrahymena* onde existe um maior conhecimento acumulado, incluindo áreas tão importantes como organização genómica, conjugação celular, rearranjos nucleares e manipulação genética.

Como primeira abordagem à presença de genes ubiquitina em *T.thermophila*, o DNA total deste organismo foi digerido com várias enzimas de restrição, e, após separação por electroforese em gel de agarose e transferência para filtro de nitrocelulose, analisado por hibridação com o fragmento codificante *Hind*III de 0.23 kb do clone TU20 (Neves *et al.*, 1991a). O DNA total deste organismo é constituído por 92 a 98% de DNA macronuclear, pelo que os resultados obtidos reflectem com mais propriedade a composição do macronúcleo. Como esperado, o padrão de hibridação observado é substancialmente diferente do obtido para o ciliado *T.pyriformis* (Fig.IV.1A). Três das enzimas, *Bam*HI, *Cla*I e *Stu*I, apresentam respectivamente 2, 2 e 7 bandas para *T.thermophila*, e 5, 7 e 1 bandas para *T.pyriformis*, enquanto as outras duas, *Eco*RI e *Hind*III, apresentam um número superior de bandas cujo padrão é igualmente distinto para ambos os organismos.

Apesar destas diferenças, é possível deduzir uma organização genómica dos genes ubiquitina em *T.thermophila*, que aparenta ser semelhante à encontrada em *T.pyriformis*. O número de genes deverá ser aproximadamente igual em ambos os organismos, e é seguramente superior a seis em *T.thermophila*. Para se chegar a este número, é preciso tomar em atenção que o número máximo de unidades ubiquitina por gene em *T.thermophila* deverá ser cinco: o maior mRNA observado neste organismo é equivalente ao de *T.pyriformis*, onde corresponde a um gene de cinco unidades (# 3). Sendo assim, o maior gene terá cerca de 1.2 kb, e o maior fragmento codificante possível de se obter com uma enzima que reconheça dois locais no gene (um por unidade) terá no máximo cerca de 1 kb: fragmentos superiores a este valor conterão sempre região não codificante, sejam regiões a montante e/ou a jusante do gene, sejam regiões intergénicas. Assim, como se observam pelo menos 12 bandas *Eco*RI superiores a 1 kb, estas corresponderão de 6 (cada fragmento com a extremidade de um único gene, seis montantes e seis jusantes) a 12 genes (cada fragmento com um gene completo). Por outro lado, estes genes deverão, à semelhança do que ocorre em *T.pyriformis*, estar organizados em série. Esta constatação provém da observação de que o elevado número de bandas obtidas com as enzimas *Hind*III e *Eco*RI se reduz a sete com a enzima *Stu*I e a duas com as enzimas *Cla*I e *Bam*HI. Se a enzima *Stu*I não reconhecer nenhum local interno dos genes, as sete bandas observadas corresponderão a sete genes, um número próximo do mínimo calculado com a enzima *Eco*RI. Contudo, a enzima *Stu*I possui um local de restrição, AGGCCT, pouco frequente em *Tetrahymena*, o que é facilmente comprovado pela acumulação de fragmentos de elevada massa molecular observada em geis corados com brometo de etídio (resultado não mostrado). De facto, em

T.pyriformis, observa-se apenas uma banda de elevada massa molecular em hibridações de DNA digerido com *StuI*, pelo que as sete bandas observadas em *T.thermophila*, entre 1.5 a 11 kb, serão mais facilmente explicadas se a enzima reconhecer um local interno dos genes. Assim, sendo todas as bandas superiores a 1.2 kb, o número mínimo de genes possível sem que cada fragmento contivesse mais de um gene seria de quatro, correspondendo a seis fragmentos com as extremidades de três genes e um com um gene completo. Como este valor é inferior ao mínimo calculado com a enzima *EcoRI*, alguns fragmentos *StuI* terão necessariamente de conter mais de um gene, originando o seu agrupamento em série. Seguindo o mesmo raciocínio com as duas bandas observadas com as enzimas *ClaI* e *BamHI*, facilmente se deduz que pelo menos uma corresponderá a um fragmento contendo vários genes.

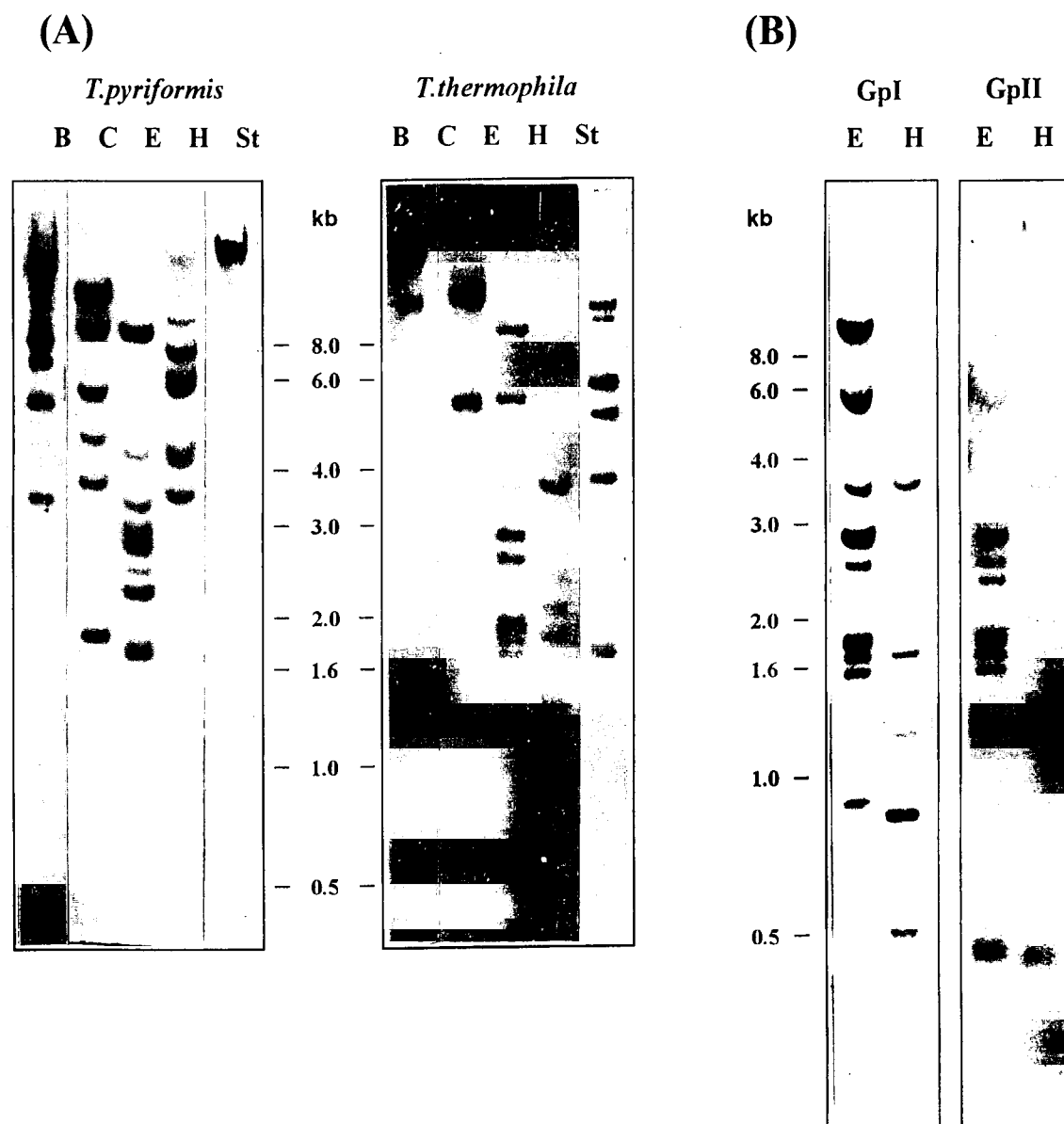


Fig.IV.1. Detecção de seqüências ubiquitina em DNA de *T.thermophila*. (A) Aliquotas de DNA macronuclear de *T.pyriformis* (A1) e de DNA total de *T.thermophila* (A2) foram digeridas com as enzimas indicadas, analisadas em geis de agarose a 1% e transferidas para filtros de nylon. A hibridação foi efectuada com sonda para ubiquitina (Ub). Enzimas de restrição: B-*BamHI*; C-*ClaI*; E-*EcoRI*; H-*HindIII*; St-*StuI*. (B) Filtros contendo DNA total de *T.thermophila* digerido com as enzimas *EcoRI* e *HindIII* foram hibridados em condições ou restritivas (65°C) com sondas específicas para ubiquitinas dos Grupos I (GpI) e II (GpII).

De modo a determinar a eventual presença de genes poliubiquitina do Grupo II em *T.thermophila*, digestões genômicas com as enzimas *EcoRI* e *HindIII* foram analisadas por hibridação em condições restritivas com uma sonda codificante deste grupo, correspondente ao fragmento *EcoRI* de 0.23 kb do clone TU1 (Fig.IV.1B). Como controlo, digestões equivalentes foram hibridadas nas mesmas condições com uma sonda codificante do Grupo I, correspondente ao fragmento *HindIII* de 0.23 kb do clone TU20. Os resultados obtidos não são tão conclusivos como os conseguidos com *T.pyriformis*, já que a maioria das hibridações obtidas com a sonda do Grupo II poderão ter sido consequência de hibridação cruzada com genes do Grupo I. Contudo é possível observarem-se várias bandas cujo sinal é claramente superior com a sonda do Grupo II do que com a sonda do Grupo I, como as bandas *EcoRI* de 2.2 e 0.45 kb, e a *HindIII* de 1.4 kb. Também é curioso notar que, embora certas bandas intensas presentes com a sonda do Grupo I praticamente desapareçam com a sonda do Grupo II (p.ex., *EcoRI* de 3.5 kb, *HindIII* de 1.6 kb), outras há que se mantêm com ambas as sondas, apesar de menos intensas. Estes dados, embora não permitam concluir com clareza da existência de genes poliubiquitina do Grupo II em *T.thermophila*, também não permitem excluir a sua presença neste organismo.

2. Isolamento e caracterização dos genes poliubiquitina em *T.thermophila*.

2.1. Isolamento de genes ubiquitina.

A estratégia seguida para o isolamento de genes ubiquitina em *T.thermophila* consistiu na construção em λ DashII de um banco genómico de DNA total parcialmente digerido com *Sau3AI* (ver Material e Métodos). O banco foi pesquisada por hibridação *in situ* com o fragmento *HindIII* de 0.23 kb do clone pTU20 (Neves *et al.*, 1991a) em condições pouco restritivas, de modo a permitir isolar clones contendo genes poliubiquitina dos Grupos I ou II. Obtiveram-se quatro sinais positivos, mas destes apenas foi possível amplificar e isolar um único, que foi analisado de forma mais aprofundada.

O clone fágico foi caracterizado por digestão com diversas enzimas de restrição, electroforese em gel de agarose e transferência para um filtro de nylon. O filtro foi então hibridado com a sonda atrás descrita, e o resultado da hibridação (não apresentado) permitiu concluir pela presença de um único gene ubiquitina, contido num fragmento *XbaI* de 3.3 kb. Este fragmento foi subclonado em pUC19 e caracterizado por restrição enzimática e sequenciação, revelando a presença de um gene poliubiquitina de cinco unidades numa das suas extremidades (Fig.IV.2). Curiosamente, a outra extremidade revelou conter dois exões e um intrão de um gene codificante para a região C-terminal da subunidade 2 da RNA polimerase II (ver apêndice 2).

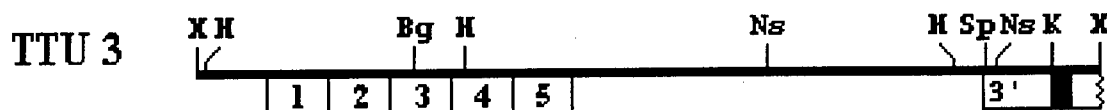


Fig.IV.2. Mapa de restrição do fragmento TTU3X. O mapa de restrição foi determinado por digestão com as seguintes enzimas de restrição: Bg-BglII; H-HindIII; K-KpnI; Ns-NsiI; Sp-SpeI; X-XbaI. As regiões codificantes foram determinadas por hibridação e confirmadas por sequenciação. As caixas numeradas representam unidades ubiquitina. No extremo do fragmento encontram-se dois exões (caixas abertas) e um intrão (caixa fechada) correspondentes à região terminal de um gene codificante para a subunidade 2 da RNA polimerase II.

2.2. Análise do gene poliubiquitina TTU3.

A sequência nucleotídica do gene poliubiquitina encontrado no fragmento TTU3X encontra-se representada na Fig.IV.3. Este gene, a que se denominou TTU3, codifica para cinco unidades ubiquitina iguais entre si e com as ubiquitinas do Grupo I de *T.pyrifomis*, pertencendo por isso a este grupo. Tal como os outros genes do Grupo I, o gene TTU3 apresenta uma utilização de códons otimizada, reflectida pelo seu valor de CAI, 0.737, que, embora inferior ao encontrado para o gene A de *T.pyrifomis*, é semelhante ao encontrado para genes de proteínas do citoesqueleto, histonas ou proteínas ribossomais (ver Tabela III.3). Contudo, o codão extra encontrado após a sua última unidade é diferente dos encontrados em *T.pyrifomis*, embora codifique para um aminoácido, Phe, que é o mais frequentemente encontrado nesta posição em genes poliubiquitina de vários organismos, incluindo o do ciliado *Euplotes eurytostomus* (ver Tabela II.1).

A análise das sequências nucleotídicas codificantes para cada uma das cinco unidades revela que estas são muito semelhantes entre si, com uma média de 13.8 substituições/par, um valor semelhante ao encontrado no gene A de *T.pyrifomis* (14.2 substituições/par). Contudo, neste gene as unidades mais semelhantes entre si, 2 e 4, divergem em seis nucleótidos, enquanto no gene TTU3 as unidades 2 e 5 são absolutamente idênticas. Estes dados reflectem a acção de uma elevada pressão evolutiva sobre estes genes, no sentido de uma evolução concertada entre as suas subunidades, um fenómeno que se observa em vários organismos (como, por exemplo, no homem) mas não em todos (Sharp & Li, 1987b). A homologia intragénica das unidades dos genes poliubiquitina é então mais elevada do que a homologia intergénica ou interespecie, reflectindo-se num aumento da média de substituições entre as unidades do gene TTU3 e as do gene A (18.64 substituições/par), C (25.4 substituições/par) ou X (25.2 substituições/par) de *T.pyrifomis*.

1	TCTaGAAGctatcaattaaatcagtttttttcttatctatttcattcaataaaaactcttt	60
61	tcattcaaattaattaattaaaaaaagataaaagctttt TTcatAGAGA aaaccattca	120
121	gacttgaacagatcgactctcttttctttcctataaaattgcgttttgctaccttcaaaaaa	180
181	aaaagaaaaacaattcaaaaagaattaaattcttattaaacaaaaacctccaaaaaaaagc	240
241	aaaaaactaaaaATGCAAATCTTCGTTAAGACCCTTACTGGTAAAACCATTACTTTAGAC	300
81	M Q I F V K T L T G K T I T L D	100
301	GTTGAAGCCTCCGACACTATTGAAAACGTCAAGGCTAAGATTCAAGACAAGGAAGGTATC	360
101	V E A S D T I E N V K A K I Q D K E G I	120
361	CCTCCTGATCAATAAAGACTTATCTTCGCTGGTAAGCAATTAGAAGACGGCAGAACCCTT	420
121	P P D Q Q R L I F A G K Q L E D G R T L	140
421	TCTGACTACAACATTCAAAGGAATCAACCCTCCACTTGGTTTTAAGATTAGAGGTGGT	480
141	S D Y N I Q K E S T L H L V L R L R G G	160
481	ATGCAAATTTTCGTCAAGACCTTGACTGGTAAGCATCACTCTTGACGTTGAAGCCTCT	540
161	M Q I F V K T L T G K T I T L D V E A S	180
541	GACACTATTGAAAACGTCAAGGCTAAGATCCAAGACAAGGAAGGTATCCCCCTGATCAA	600
181	D T I E N V K A K I Q D K E G I P P D Q	200
601	TAAAGACTTATCTTCGCTGGTAAGCAATTAGAAGACGGCAGAACCCTTTCTGACTACAAC	660
201	Q R L I F A G K Q L E D G R T L S D Y N	220
661	ATCCAAAAGGAATCTACCCTCCACTTAGTCCTCAGATTAAGAGGTGGTATGTAAATCTTC	720
221	I Q K E S T L H L V L R L R G G M Q I F	240
721	GTCAAGACCTTAAGTGGCAAGACCATCACTCTTGATGTTGAAGCCTCCGACACCATTGAA	780
241	V K T L T G K T I T L D V E A S D T I E	260
781	AATGTCAAGGCTAAGATCTAAGACAAGGAAGGTATCCCCCTGATCAATAAAGACTTATC	840
261	N V K A K I Q D K E G I P P D Q Q R L I	280
841	TTGCTGGTAAGCAATTAGAAGACGGCAGAACCCTTTCTGACTACAACATCCAAAAGGAA	900
281	F A G K Q L E D G R T L S D Y N I Q K E	300
901	TCTACTCTCCACTTAGTCCTTAGATTGAGAGGTGGTATGTAAATCTTCGTCAGACCTTA	960
301	S T L H L V L R L R G G M Q I F V K T L	320
961	ACTGGCAAGACCATCACTCTTGATGTGCAAGCTTCTGACACCATTGAAAACGTCAAGGCT	1020
321	T G K T I T L D V E A S D T I E N V K A	340
1021	AAGATCCAAGACAAGGAAGGTATCCCCCTGATCAATAAAGACTTATCTTCGCTGGTAAG	1080
341	K I Q D K E G I P P D Q Q R L I F A G K	360
1081	CAATTAGAAGACGGCAGAACCCTTTCTGACTACAACATTCAAAGGAATCAACCCTCCAC	1140
361	Q L E D G R T L S D Y N I Q K E S T L H	380
1141	TTGGTTTTAAGATTAAAGAGGTGGTATGCAATTTTCGTCAAGACCTTGACTGGTAAGACC	1200
381	L V L R L R L R G G M Q I F V K T L T G K T	400
1201	ATCACTCTTGACGTTGAAGCCTCTGACACTATTGAAAACGTCAAGGCTAAGATCCAAGAC	1260
401	I T L D V E A S D T I E N V K A K I Q D	420
1261	AAGGAAGGTATCCCCCTGATCAATAAAGACTTATCTTCGCTGGTAAGCAATTAGAAGAC	1320
421	K E G I P P D Q Q R L I F A G K Q L E D	440
1321	GGCAGAACCCTTTCTGACTACAACATCCAAAAGGAATCTACCCTCCACTTAGTCCTCAGA	1380
441	G R T L S D Y N I Q K E S T L H L V L R	460
1381	TTAAGAGGTGGTTTCTGAattgaatgaaaatcgcatttaaaacaacattaaaaataaaac	1440
461	L R G G F *	480
1441	atcacttaaaaaatgaataaaaattgatttgttactctctatatatatacataaaatacttgt	1500
1501	tgttttgaatttaacatctctaaattcaattaaataaaatttccttcctttcaaattaatt	1560
1561	aaagaaaaaaataaaaaattaactaatttttcaatcaaacaataaaaatttttcttttctt	1620
1621	aaatgtgatctaagaaatgattgattaatattcattaaaaatttcataaaaaatttgttta	1680
1681	tgaattttttttgaaattcatcatcataacctatttgatgaaatatgtaaatataaaatt	1740
1741	ctgataattatttgttttttagtcagaaaaaaagaaattattttttctatttttaaaagt	1800
1801	atgaaaatagaaatttttaagtgagaatggtttttatttataaattaccgtttgttttaaaa	1860
1861	aatgttttgattgaattaattttcatatttttaagttaattcaaattaaaaatttatttgca	1920
1921	tttagtaggtctgcctctattaattggaaaatctatatgaattttttctactatcttat	1980

Fig.IV.3. Sequência nucleotídica do gene poliubiquitina TTU3. A região codificante encontra-se apresentada em letras maiúsculas, e as respectivas regiões não codificantes em letras minúsculas. A sequência de aminoácidos predita encontra-se indicada sob a respectiva sequência nucleotídica, sendo o codão de terminação indicado por um asterisco. O motivo AGAGA encontra-se em negrito, e um possível elemento de choque térmico truncado encontra-se sublinhado.

Para além da região codificante, a análise da sequência revela a presença de um elemento AGAGA situado a -142 do codão de iniciação, uma posição praticamente idêntica à do elemento encontrado no gene A de *T.pyriformis* (-143). Já a presença de um elemento de choque térmico, HSE, não pode ser confirmada, embora possa ser deduzida. É possível de facto encontrar-se uma sequência que aparenta ser parte de um elemento HSE e que coincide com o extremo do fragmento (a sublinhado na Fig.IV.3). Os oito nucleótidos encontrados são idênticos aos do elemento HSE do gene A de *T.pyriformis*, e localizam-se a uma distância semelhante do codão de iniciação (-245 para o gene TTU3, contra -276 do gene A). O elemento HSE do gene A contém igualmente uma sequência reconhecível pela enzima de restrição *Xba*I, pelo que a sequência parcial encontrada no gene TTU3 pode efectivamente corresponder a um verdadeiro elemento HSE, embora só a clonagem e sequenciação de um fragmento contendo a promotora deste gene o possa confirmar em absoluto.

A comparação das regiões não codificantes do gene TTU3 com as regiões equivalentes dos diversos genes poliubiquitina encontrados em *T.pyriformis* revela apenas uma única semelhança apreciável, de cerca de 73%, entre os cerca de 100 nucleótidos imediatamente a montante dos genes TTU3 de *T.thermophila* e A de *T.pyriformis*, que correspondem aproximadamente à região "leader" do RNA mensageiro deste último gene (Neves *et al.*, 1991a). A semelhança decai acentuadamente a montante desta região entre estes dois genes, e atinge apenas 55-60% entre a região 5' do gene TTU3 e as dos outros genes de *T.pyriformis* que não o gene A, ou entre a região 3' do gene TTU3 e as de qualquer outro gene de *T.pyriformis*, incluindo a do gene A. A comparação entre as regiões não codificantes de diversos genes ortólogos sequenciados em ambos os ciliados *T.pyriformis* e *T.thermophila* (actina, tubulinas, histonas H3/H4, calmodulina e Ran/TC4) revelam homologies de 70 a 90%, que são contudo observáveis quer nas regiões a montante, quer nas regiões a jusante. Este facto não permite afirmar com segurança que o gene TTU3 é, em *T.thermophila*, o gene ortólogo do gene A de *T.pyriformis*, já que não é observada uma homologia entre as suas regiões 3'. Contudo, a elevada semelhança observada entre os dois genes, a nível das sequências codificantes, "leader", elementos HSE e localização dos elementos promotores, não observável com nenhum outro gene conhecido de *T.pyriformis*, não permite igualmente excluir essa hipótese.

3. Expressão dos genes ubiquitina em *T.thermophila*.

A análise comparativa da expressão dos genes ubiquitina de *T.thermophila* com a expressão dos genes ubiquitina de *T.pyriformis* (Fig.IV.4A) revela que fundamentalmente ambos os organismos transcrevem os mesmos tipos de mRNAs ubiquitina: três populações de mensageiros de 1.8, 1.6 e 1.4 kb correspondentes a genes poliubiquitinas, e uma população de mensageiros de 0.75 kb correspondentes ao(s) gene(s) de fusão. Contudo, é possível observarem-se diferenças significativas nos seus padrões de expressão, particularmente no que diz respeito aos mensageiros de 1.4 e 1.6 kb. Enquanto em *T.pyriformis* os mRNAs de 1.6 kb são expressos em situações de "stress" hipertérmico, não sendo detectáveis durante o crescimento exponencial, em *T.thermophila* assiste-se a uma clara indução dos níveis dos mRNAs de 1.4 kb nas mesmas condições, embora neste caso se observe uma expressão vestigial em células em crescimento exponencial. Este cruzamento parece suceder igualmente com os mRNAs de 1.4 kb de *T.pyriformis*, um

mensageiro dificilmente visível e levemente induzido em situações de "stress", e com os mensageiros de 1.6 kb em *T. thermophila*, cujo padrão de expressão é semelhante. Este facto sugere uma eventual recombinação dos genes codificantes para os mensageiros de 1.6/1.4 kb e 1.4/1.6 kb de *T. pyriformis*/*T. thermophila*, que terão perdido/ganho uma unidade ubiquitina durante a sua evolução. Já os mRNAs de 1.8 e 0.75 kb são os preferencialmente expressos durante o crescimento exponencial em ambos os organismos, sendo os de 1.8 kb igualmente induzidos durante o "stress" hipertérmico, jejum, fase estacionária e reciliação (Fig. IV.4B; Soares *et al.*, 1993). É curioso notar que, destas quatro situações de "stress", apenas o "stress" hipertérmico e a reciliação induzem os mensageiros de 1.4 kb (1.6 kb em *T. pyriformis*), o que poderá significar que o(s) gene(s) codificante(s) para estes mensageiros poderá(ão) ser induzido(s) apenas por alterações rápidas e drásticas do ambiente, e não por situações de "stress" continuado.

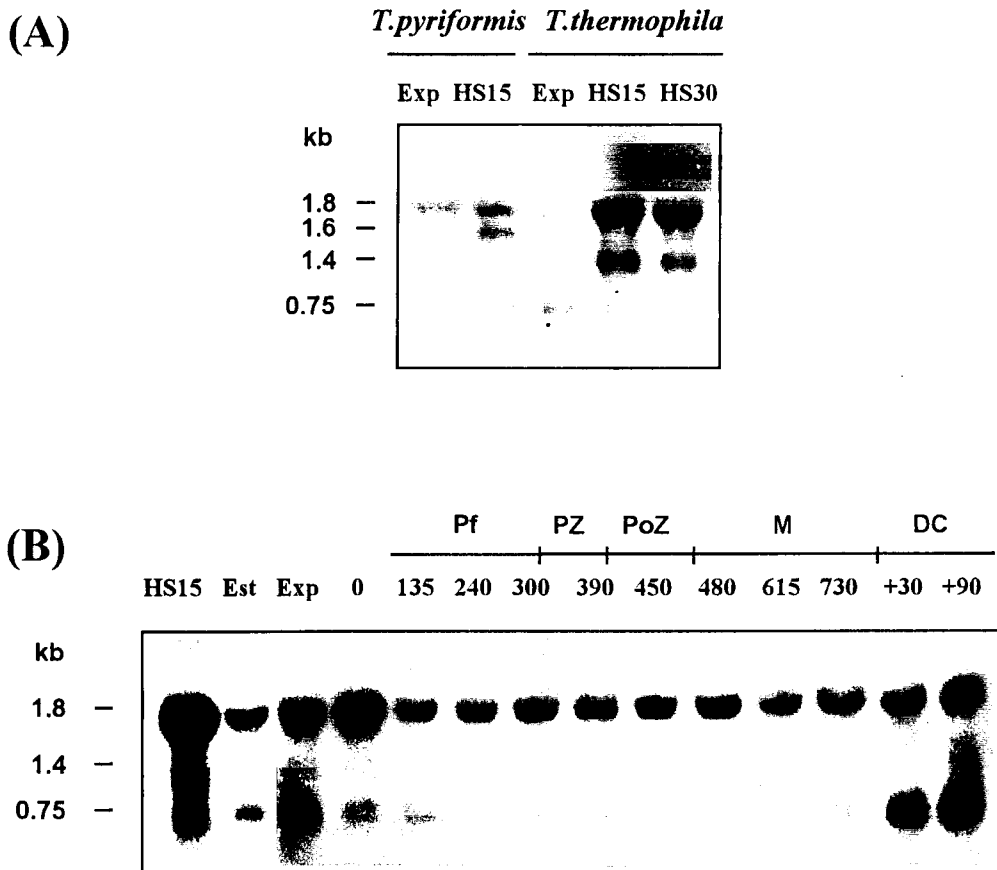


Fig. IV.4. Expressão dos genes ubiquitina de *T. thermophila*. Cerca de 20µg de RNA total extraído de células de *Tetrahymena* foram separados em géis de agarose a 2% com formaldeído, transferidos para nitrocelulose e hibridados com sonda ubiquitina Ub. (A) RNA extraído de células de *T. pyriformis* ou *T. thermophila* em crescimento exponencial (Exp) ou sujeitas a um "stress" hipertérmico de 15 (HS15) ou 30 min (HS30) (B) RNA extraído de células de *T. thermophila* cultivadas em PPY diluído 1:3 em crescimento exponencial (Exp), em fase estacionária ($> 10^6$ células/ml) (Est) ou submetidas a um "stress" hipertérmico de 15 min (HS15). Para a conjugação, células *T. thermophila* B4 e SB1918 em jejum foram postas em contacto (0) e subdivididas em aliquotas contendo cerca de 6×10^6 células cada, extraíndo-se o RNA de uma aliquota ao fim de cada tempo indicado. Após 780 min de contacto, as células das aliquotas restantes foram realimentadas por adição de meio volume de PPY, sendo extraído RNA de aliquotas 30 (+30) e 90 min (+90) após a realimentação. Os estados citológicos a cada tempo foram analisados como indicado nos Materiais e Métodos, representando a barra o período temporal em que pelo menos 50% das células se encontram nos estados referidos: Pf, prófase meiótica; PZ, divisão pré-zigótica; PoZ, divisão pós-zigótica; M, desenvolvimento macronuclear; DC, primeira divisão celular.

A expressão dos genes ubiquitina foi igualmente estudada durante a conjugação de *T.thermophila*. Este é o único estado fisiológico conhecido em que existe actividade transcricional no micronúcleo (ver Capítulo I2), e que não ocorre em *T.pyriformis*. Células de *T.thermophila* das estirpes B4 e SB1918 submetidas a jejum durante 24 horas foram postas em contacto, subdivididas em aliquotas com cerca de 6×10^6 células cada e deixadas em repouso a 30°C. A subdivisão em pequenos volumes é necessária para uma melhor oxigenação: uma má oxigenação inibe a conjugação. Ao fim de vários tempos (Fig.IV.4B), uma das aliquotas foi retirada para análise da percentagem de conjugantes, estados citológicos e extracção do RNA total. Após 135 min de contacto já cerca de 40% das células se encontravam emparelhadas, atingindo-se o valor máximo de conjugantes (80-90%) ao fim de 240 min de contacto. O desemparelhamento iniciou-se com cerca de 615 min de contacto, com 60-65% de conjugantes, decrescendo este valor para 40-45% após 730 min de contacto e 20% após 780 min, altura em que as células foram realimentadas por adição de meio volume de PPY. A análise dos estados citológicos presentes em cada ponto temporal, representada pela barra na Fig.IV.4B, mostrou uma boa correlação com o descrito na literatura (Martindale *et al.*, 1982), apresentando contudo uma atraso de cerca de 30 min. Estes desfasamentos entre diversas experiências são comuns e dependem de diversos factores, como por exemplo o tempo de jejum da células antes da sua junção (Martindale *et al.*, 1982).

A análise da expressão dos genes ubiquitina durante a conjugação de *T.thermophila* (Fig.IV.4B) revela a ausência de uma indução específica de qualquer um dos mRNAs codificantes para poliubiquitinas. Apenas no início da conjugação (ponto 0) se observa uma indução dos mRNAs de 1.8 kb, mas este ponto corresponde mais propriamente a um jejum de 24 horas, um estado fisiológico que se sabe ser indutor de "stress" em *Tetrahymena* (Soares *et al.*, 1993). Pelo contrário, é possível observar-se uma redução acentuada dos níveis dos mRNAs codificados pelo(s) gene(s) de fusão durante toda a conjugação. A síntese proteica aumenta durante a fase inicial da conjugação, correspondente ao início da prófase mitótica, mas decresce imediatamente após esta fase (Martindale *et al.*, 1985), o que pode justificar a redução dos níveis de um mensageiro codificante para uma proteína ribossomal. Só quando as células são realimentadas se observa uma indução do mRNA da proteína de fusão a níveis superiores ao do crescimento exponencial. A realimentação induz a divisão celular, iniciada com a divisão dos exconjugantes a células carionides, pelo que é de esperar um rápido aumento da síntese proteica e da correspondente necessidade de proteínas ribossomais.

4. Discussão.

Tal como no ciliado *T.pyriformis*, a ubiquitina em *T.thermophila* é codificada por uma família multigénica. A sua análise por Southern genómicos sugere que o número de genes constituintes da família é igualmente elevado, de entre seis a uma dezena, que se encontram preferencialmente organizados em série. Contudo, não é possível determinar com clareza em *T.thermophila* se os genes poliubiquitina se distribuem igualmente por grupos distintos aos encontrados em *T.pyriformis*. Se a presença de genes do Grupo I é clara, e foi confirmada com a posterior clonagem e sequenciação de um gene do grupo, a presença de genes do Grupo II é ambígua, e se não pode ser confirmada até agora, também não deve ser eliminada. O uso em condições restritivas de uma sonda específica deste grupo

em Southern genômicos revela efectivamente algumas bandas aparentemente mais específicas do que em Southern equivalentes onde foi usada uma sonda específica para o Grupo I. Todavia, os sinais obtidos não são muito intensos, e podem ser consequência de uma hibridação cruzada entre a sonda do Grupo II e genes do Grupo I. Há, contudo, dois dados a reter, e que suportam a presença de genes do Grupo II. O primeiro prende-se com a utilização de codões, que é semelhante mas não idêntica entre as duas *Tetrahymena*, e com a eventual diferença entre os genes do Grupo II dos dois organismos. Estes genes, a existirem em *T. thermophila*, poderão divergir significativamente dos genes encontrados em *T. pyriformis*, quer porque estes últimos possuem uma utilização de codões muito desfavorável (que poderá não ser tão acentuada em *T. thermophila*), quer porque as análises neles efectuadas sugerem que acumularam várias mutações aleatórias para além das observadas pelo padrão característico do grupo. A diferente utilização de codões é evidenciada pela enzima *StuI*: aparentemente esta enzima reconhece em *T. thermophila* sequências internas às regiões codificantes dos genes. Ora, a única sequência susceptível de reconhecimento por esta enzima nas ubiquitinas do Grupo I é 18-EAS-20, que coincide com a região mais variável nos genes poliubiquitina do Grupo II, e apenas desde que seja codificada por GAGGCCTCN. Contudo, este facto implica que o aminoácido Glu seja necessariamente codificado pelo codão GAG, um codão extremamente raro em *Tetrahymena* que constitui apenas 2% dos codões codificantes para Glu conhecidos. Assim, é possível que a baixa intensidade dos sinais obtidos com a sonda específica dos genes do Grupo II de *T. pyriformis* em Southern de DNA genómico de *T. thermophila* seja devida apenas a diferenças significativas nas respectivas sequências de cada organismo.

O segundo dado prende-se com as restrições *HindIII* e *EcoRI*. É curioso notar (ver Fig.III.1) que, nos genes conhecidos de *T. pyriformis*, a enzima de restrição *HindIII* reconhece exclusivamente locais em genes poliubiquitina do Grupo I, enquanto a enzima *EcoRI* o faz exclusivamente em genes do Grupo II. Sendo assim, os fragmentos *EcoRI* inferiores a 0.8 kb (correspondente a um gene de três unidades, eventualmente codificante para um mensageiro de 1.4 kb) serão resultantes de genes com locais de restrição *EcoRI* internos, e por isso poderão conter sequências de genes poliubiquitina do Grupo II. Existem três bandas condizentes em *T. thermophila*, de 0.7, 0.45 e 0.23 kb, que correspondem aproximadamente a três, duas e uma unidade ubiquitina. Dessas três, apenas a de 0.45 kb é claramente mais intensa com a sonda do Grupo II nas condições restritivas. É de notar ainda que, caso a enzima *EcoRI* reconheça efectivamente apenas locais em genes do Grupo II, várias das bandas observadas com mais de 1 kb conterão genes completos do Grupo I, o que sugere igualmente que o número total de genes em *T. thermophila* será superior a seis, e muito mais próximo do valor de doze atrás calculado.

De modo a procurar um maior esclarecimento sobre a organização dos genes ubiquitina em *T. thermophila*, e a determinar com maior rigor da presença de genes poliubiquitina do Grupo II, pesquisou-se um banco genómico total deste organismo com sondas ubiquitina em condições pouco restritivas. Dos quatro clones positivos obtidos, apenas foi possível isolar-se um, o qual foi analisado e caracterizado com maior profundidade. Este clone revelou conter um único gene pertencente ao Grupo I, cujas principais características são a de codificar para uma ubiquitina idêntica à da *T. pyriformis*, possuir um codão extra distinto dos encontrados nos genes deste organismo, e possuir igualmente na região promotora um elemento AGAGA e, aparentemente, um elemento HSE. Outros dados de interesse são a presença de um local de restrição *HindIII* interno, a ausência de locais *EcoRI* e *StuI*, e a ausência de outros genes em série, tendo-se encontrado apenas um gene codificante para a segunda subunidade da RNA polimerase II. Com a

excepção da presença e ausência de locais *HindIII* e *EcoRI*, que estão de acordo com as hipóteses apresentadas atrás, a ausência de locais *StuI* e de outros genes ubiquitina em série parecem contradizer o modelo de organização exposto. Contudo, é preciso notar que dificilmente as sete bandas *StuI* observadas conterão apenas extremidades de genes. Se assim fosse, o número máximo de genes ubiquitina seria seis, mas devido às análises efectuadas com a enzima *EcoRI* parece mais provável que o número seja superior a este. Por outro lado, a ausência de outros genes poliubiquitina em série poderá ser explicada pelo facto de, possivelmente, o gene TTU3 ser o gene ortólogo do gene A de *T.pyriformis*. Neste organismo, o gene A não possui aparentemente outros genes em série, e parece localizar-se numa região distinta do genoma relativamente aos outros genes, pelo que o gene TTU3 poderá seguir o mesmo comportamento em *T.thermophila*.

A análise da expressão dos genes ubiquitina em *T.thermophila* revela que este organismo transcreve mensageiros cuja massa molecular relativa é idêntica à dos encontrados em *T.pyriformis*, mas que enquanto os mRNAs de 1.8 e 0.75 kb parecem possuir um comportamento semelhante em ambos os organismos, as populações de 1.6 e 1.4 kb possuem comportamentos cruzados. Em *T.pyriformis*, os mRNAs de 1.6 kb são induzidos quando submetidos a um "stress" hipertérmico, enquanto em *T.thermophila* o mesmo "stress" induz mRNAs de 1.4 kb, apesar de estes serem também residualmente expressos durante a fase de crescimento exponencial da célula. Os mensageiros de 1.4 kb, em *T.pyriformis*, e de 1.6 kb, em *T.thermophila*, apresentam apenas um sinal residual em ambos os organismos. A sua função é desconhecida, mas é curioso notar que, em *T.pyriformis*, o mRNA de 1.4 kb é fortemente induzido na presença das drogas despolimerizantes dos microtúbulos, como a colchicina ou o nocadazole, sugerindo um possível envolvimento da ubiquitina com os microtúbulos (H.Soares, comunicação pessoal).

A expressão dos genes ubiquitina durante a conjugação em *T.thermophila* não revela alterações significativas, com excepção do(s) gene(s) de fusão, cujos mensageiros tem um decréscimo acentuado durante todo o processo, recuperando apenas após a realimentação das células. A não indução das espécies de mRNAs conhecidas, ou o aparecimento de uma nova espécie durante o início da meiose profásica (que coincide com o pico de actividade transcricional do micronúcleo; Martindale *et al.*, 1985) ou durante a maturação do novo macronúcleo sugere que a ubiquitina poderá não desempenhar um papel importante durante as fases principais da conjugação. Contudo, a ubiquitina é expressa ao longo de todo o processo de conjugação, o que não sucede com outras proteínas. Estudos de transcrição aparente revelam que, por exemplo, a tubulina não é transcrita durante a conjugação (Stargell *et al.*, 1990), e é possível observar-se uma redução muito acentuada dos níveis do mRNA de tubulina durante toda a conjugação, à semelhança do que sucede com os mRNAs do(s) gene(s) de fusão (H.Soares, comunicação pessoal). A manutenção dos níveis do(s) mensageiro(s) de 1.8 kb de ubiquitina pode então significar uma necessidade desta proteína durante o ciclo sexual de *T.thermophila*.

CONCLUSÕES E PERSPECTIVAS

CONCLUSÕES E PERSPECTIVAS

Os resultados apresentados neste trabalho permitiram ampliar e aprofundar os conhecimentos anteriores sobre a organização genómica e expressão dos genes ubiquitina em *Tetrahymena*. Os estudos realizados em *T.pyriformis* permitiram concluir que os genes ubiquitina constituem neste organismo uma família multigénica com um elevado número de genes, que se calcula ser superior a onze, codificando pelo menos um para uma proteína de fusão ubiquitina-proteína ribossomal de 53 aminoácidos, e os restantes para poliubiquitinas. Fundamentalmente, foi possível identificar dois grupos evolutivamente distintos de genes poliubiquitina neste organismo, um facto inédito relativamente aos conhecimentos actuais sobre genes poliubiquitina noutros organismos. O primeiro grupo é constituído por genes que codificam proteínas idênticas entre si, e semelhantes às encontradas noutros eucariotas, pelo que deverão possuir uma função equivalente. O segundo grupo é constituído por genes cujas unidades codificam para possíveis ubiquitinas distintas entre si e com as dos outros genes. As características próprias da sua estrutura permitem classificá-los como um grupo independente do grupo anterior, eliminando a hipótese de serem pseudogenes deste grupo. Os genes de ambos os grupos encontram-se organizados de uma maneira semelhante, com os seus genes agrupados e dispostos em série, à excepção do gene A do Grupo I, que parece ter uma localização distinta. Contudo, os resultados obtidos indicam igualmente que ambos os grupos se encontram localizados em regiões distintas do genoma, cuja proximidade não é conhecida. Foi igualmente detectada a presença de um gene aparentemente distinto dos constituintes de cada um destes dois grupos, tendo sido por isso classificado num terceiro. Possui todavia algumas características que não o permitem excluir totalmente do segundo grupo, como a falta da Gly terminal e a sua localização entre genes deste grupo, o que, aliado à sua extensa mutação, pode indicá-lo como sendo um pseudogene do Grupo II. A ser o caso, será o primeiro pseudogene a ser detectado num ciliado.

A presença de um grupo raro de genes poliubiquitina em *T.pyriformis* levantou várias questões, relativas às suas possíveis funções, e à razão da sua aparente exclusividade neste organismo. Contudo, os trabalhos realizados não permitiram detectar nenhum sinal de actividade destes genes, quer determinada pela presença de mRNAs específicos, quer através da detecção de transcrição aparente. Mais ainda, a análise dos genes evidência a acumulação recente e gradual de mutações não específicas, o que sugere uma não funcionalidade efectiva para estes genes. Sendo *T.pyriformis* uma espécie particular dentro das *Tetrahymena*, devido à ausência de micronúcleo, considerou-se necessário realizar estudos da organização genómica de uma espécie micronucleada, de modo a determinar da ubiquidade do segundo grupo de poliubiquitinas dentro do *genus*. Os trabalhos efectuados em *T.thermophila* permitiram concluir que este organismo possui uma organização dos seus genes ubiquitina aparentemente semelhante à encontrada em *T.pyriformis*, com um número de genes elevado (entre 6 a 12) que deverão estar organizados em série. A presença de genes do Grupo I ficou inequivocamente demonstrada pela clonagem e caracterização de um gene do grupo, cuja estrutura é equivalente à do gene A de *T.pyriformis* e que codifica para ubiquitinas idênticas às deste organismo. Já a presença de genes do Grupo II é ambígua, devido às hibridações cruzadas que ocorrem entre os genes de ambos os grupos, e às eventuais diferenças que os genes do Grupo II de *T.thermophila*, a existirem, possam apresentar relativamente aos genes do mesmo grupo de *T.pyriformis*. A análise da expressão dos genes ubiquitina em *T.thermophila* revelou que este organismo transcreve mensageiros de massas moleculares aparentes semelhantes, embora os padrões de expressão observados em algumas situações fisiológicas sejam equivalentes para dois desses mensageiros, e

aparentemente alternados para os outros dois. Durante a conjugação, uma parte do ciclo celular inexistente em *T.pyriformis* e o único onde é detectada actividade do micronúcleo, não foi detectada nenhuma indução ou repressão específica dos genes poliubiquitina, mas observou-se que a presença de ubiquitina parece ser requerida ao longo de todo o processo, o que não sucede com outras proteínas como a tubulina ou a proteína de fusão.

Longe de se apresentar concluído, o trabalho aqui apresentado levanta várias questões e aponta diversas perspectivas para uma investigação futura. A questão mais importante prende-se com a presença ou ausência de genes poliubiquitina do Grupo II em espécies *Tetrahymena* micronucleadas, que não ficou inequivocamente demonstrada. Poder-se-á, para tal, proceder a uma nova pesquisa no banco genómico construído em *T.thermophila* de modo a obterem-se novos clones. Estes permitirão aprofundar os conhecimentos sobre a organização dos genes, como a distribuição em série e o seu número real, mas poderá ser um processo moroso e dispendioso de demonstrar a presença de genes do Grupo II. A sua ausência poderia, até, levar à necessidade de clonar e sequenciar todos os genes conhecidos para o demonstrar. Parece ser preferível, pois, usar-se um outro tipo de abordagem para este propósito. O mais eficaz parece relacionar-se com a diferente "afinidade" das enzimas de restrição *EcoRI* e *HindIII* pelos dois grupos: em *T.pyriformis*, fragmentos *HindIII* inferiores a 0.8 kb são oriundos de genes do Grupo I, enquanto fragmentos *EcoRI* inferiores a 0.8 kb são oriundos de genes do Grupo II. Existindo fragmentos *EcoRI* de 0.23, 0.45 e 0.7 kb em *T.thermophila*, parece razoável supor que estes conterão genes do Grupo II, e será por isso mais eficaz construir bancos genómicos parciais *EcoRI* que incluam esses fragmentos. A não obtenção de genes do Grupo II nestes fragmentos não excluirá definitivamente a sua presença em *T.thermophila*, mas torna-la-á muito improvável.

Inevitavelmente, a questão seguinte por-se-á com a organização dos genes ubiquitina no micronúcleo. Existirá uma multiplicação dos genes durante a maturação do macronúcleo, ou o número de genes é equivalente em ambos os núcleos? Ocorrerão reorganizações que possam justificar o aparecimento dos genes do Grupo II em *T.pyriformis*, ou caso estes existam igualmente em *T.thermophila*, serão preexistentes no micronúcleo? Estas e outras questões só poderão ser respondidas com o estudo do genoma micronuclear, através do uso de sondas específicas em Southern genómicos.

A eventual função dos genes do Grupo II permanece ainda sem resposta. A sua ausência em *T.thermophila* poderá indicar que estes genes não são nem foram funcionais, mas se-lo-ão as putativas ubiquitinas por eles codificadas? Uma resposta poderá ser obtida através da sua integração noutros organismos, como *S.cerevisiae*. Esta levedura possui um único gene poliubiquitina, não essencial excepto em situações de stress ambiental, e a sua substituição por um gene do Grupo II poderá fornecer indicações importantes sobre a interconvertibilidade das proteínas, ou da sua toxicidade.

Finalmente, há uma questão que permanece sem resposta e que se prende com o gene de fusão ubiquitina-proteína ribossomal de 76-81 aminoácidos (Ub76). De facto, este gene nunca foi detectado não só em *Tetrahymena*, como em nenhum outro protozoário. A ausência da proteína ribossomal de 76-81 aminoácidos nestes organismos levantaria questões curiosas sobre a composição do ribossoma ao longo da escala filogenética, pelo que seria interessante avaliar a sua presença. É possível que exista isolado da ubiquitina, levantando questões não menos interessantes sobre a utilidade dos genes de fusão, ou que exista fundido a uma sequência homóloga à ubiquitina (como em *C.elegans*), sendo interessante verificar então a sua semelhança aos grupos II e III. A melhor metodologia a este problema passará pelo uso do PCR, já que o uso de sondas heterólogas para o gene de fusão Ub76 apenas permitiram, em *T.pyriformis*, clonar e identificar o gene de fusão Ub53.

BIBLIOGRAFIA

- Abrahamsen, M.S., Clark, T.G., Mascolo, P., Speer, C.A. & White, M.W. (1993). Developmental gene expression in *Eimeria bovis*. *Mol. Biochem. Parasitol.* **57**, 1-14.
- Acker, J., Wintzerith, M., Vigneron, M. & Kédinger, C. (1992). Primary structure of the second largest subunit of human RNA polymerase II (or B). *J. Mol. Biol.* **226**, 1295-1299.
- Agarwal, M.L. & Cullis, C.A. (1991). The ubiquitin-encoding multigene family of flax, *Linum usitatissimum*. *Gene* **99**, 69-75.
- Ahn, K.-S. & Henney Jr., H.R. (1994). An *Acanthamoeba* ubiquitin-fusion protein; cDNA and deduced protein sequence. *Biochem. Biophys. Acta* **1218**, 109-111.
- Ajioka, J. & Swindle, J. (1993). The calmodulin-ubiquitin associated genes of *Trypanosoma cruzi*: their identification and transcription. *Mol. Biochem. Parasitol.* **57**, 127-136.
- Allen, S. & Gibson, I. (1973). Genetics of *Tetrahymena*. Em: Elliott, A.M. (ed). *Biology of Tetrahymena*. Dowden, Hutchinson & Ross, Inc., Pennsylvania, pp. 307-373.
- Allen, S.L. (1967). Cytogenetics of genomic exclusion in *Tetrahymena*. *Genetics* **55**, 797-822.
- Allis, C.D. & Dennison, D.K. (1982). Identification and purification of young macronuclear anlagen from conjugating cells of *Tetrahymena thermophila*. *Dev. Biol.* **93**, 519-533.
- Allis, C.D. & Wiggins, J.C. (1984). Histone rearrangements accompany nuclear differentiation and dedifferentiation in *Tetrahymena*. *Dev. Biol.* **101**, 282-294.
- Allis, C.D., Allen, R.L., Wiggins, J.C., Chicoine, L.G. & Richman, R. (1984). Proteolytic processing of H1-like histones in chromatin: a physiologically and developmentally regulated event in *Tetrahymena* micronuclei. *J. Cell Biol.* **99**, 1669-1677.
- Allis, C.D., Richman, R., Gorovsky, M.A., Ziegler, Y.S., Touchstone, B., Bradley, W.A. & Cook, R.G. (1986). hv1 is an evolutionary conserved H2A variant that is preferentially associated with active genes. *J. Biol. Chem.* **261**, 1941-1948.
- Allito, B.A. & Karrer, K.M. (1986). A family of DNA sequences is reproducibly rearranged in the somatic nucleus of *Tetrahymena*. *Nucleic Acids Res.* **14**, 8007-8025.
- Altschuler, M.I. & Yao, M.-C. (1985). Macronuclear DNA of *Tetrahymena thermophila* exists as defined subchromosomal-sized molecules. *Nucleic Acids Res.* **13**, 5817-5831.
- Amar, L. (1994). Chromosome end formation and internal sequence elimination as alternative genomic rearrangements in the ciliate *Paramecium*. *J. Mol. Biol.* **236**, 421-426.
- Ammerman, D. (1971). Morphology and development of the macronuclei of the ciliates *Stylonychia mytilus* and *Euplotes aediculatus*. *Chromosoma* **33**, 209-238.
- Arfin, S.M. & Bradshaw, R.A. (1988). Cotranslational processing and protein turnover in eukaryotic cells. *Biochem.* **27**, 7979-7984.

- Arribas, C., Sampedro, J. & Izquierdo, M. (1986). The ubiquitin genes in *D. melanogaster*: transcription and polymorphism. *Biochem. Biophys. Acta* **868**, 119-127.
- Austerberry, C.F. & Yao, M.-C. (1987). Nucleotide sequence structure and consistency of a developmentally regulated DNA deletion in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **7**, 435-443.
- Austerberry, C.F., Allis, C.D. & Yao, M.-C. (1984). Specific DNA rearrangements in synchronously developing nuclei of *Tetrahymena*. *Proc. Natl. Acad. Sci. USA* **81**, 7383-7387.
- Austerberry, C.F., Snyder, R.O. & Yao, M.-C. (1989). Sequence microheterogeneity is generated at junctions of programmed DNA deletions in *Tetrahymena thermophila*. *Nucleic Acids Res.* **17**, 7263-7272.
- Baba, M.L., Dargy, L.L., Goodman, M. & Czelusniak, J. (1981). Evolution of cytochrome *c* investigated by the maximum parsimony method. *J. Mol. Evol.* **17**, 197-213.
- Bachmair, A. & Varshavsky, A. (1989). The degradation signal in a short-lived protein. *Cell* **56**, 1019-1032.
- Bachmair, A., Becker, F., Masterson, R.V. & Schell, J. (1990). Perturbation of the ubiquitin system causes leaf curling, vascular tissue alterations and necrotic lesions in a higher plant. *EMBO J.* **9**, 4543-4549.
- Bachmair, A., Finley, D. & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its N-terminal residue. *Science* **234**, 179-186.
- Baird, S.E., Fino, G.M., Tausta, S.L. & Klobutcher, L.A. (1989). Micronuclear genome organization in *Euplotes crassus*: a transposonlike element is removed during macronuclear development. *Mol. Cell. Biol.* **9**, 3793-3807.
- Baird, S.E. & Klobutcher, L.A. (1989). Characterization of chromosome fragmentation in two protozoans and identification of a candidate fragmentation sequence in *Euplotes crassus*. *Genes Dev.* **3**, 585-597.
- Baker, R.T. & Board, P.G. (1987). The human ubiquitin gene family: structure of a gene and pseudogenes from the UbB subfamily. *Nucleic Acids Res.* **15**, 443-463.
- Baker, R.T. & Board, P.G. (1989). Unequal crossover generates variation in ubiquitin coding unit number at the human *UbC* polyubiquitin locus. *Am. J. Hum. Genet.* **44**, 534-542.
- Baker, R.T. & Board, P.G. (1991). The human ubiquitin-52 amino acid fusion protein gene shares several structural features with mammalian ribosomal protein genes. *Nucleic Acids Res.* **19**, 1035-1040.
- Baker, R.T. & Board, P.G. (1992). The human ubiquitin/52-residue ribosomal protein fusion gene subfamily (UbA₅₂) is composed primarily of processed pseudogenes. *Genomics* **14**, 520-522.
- Baker, R.T., Tobias, J.W. & Varshavsky, A. (1992). Ubiquitin-specific proteases of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 23364-23375.
-

- Ball, E., Karlik, C.C., Beall, C.J., Saville, D.L., Sparrow, J.C., Bullard, B. & Fyrberg, E.A. (1987). Arthrin, a myofibrillar protein of insect flight muscle, is an actin-ubiquitin conjugate. *Cell* **51**, 221-228.
- Banerjee, A., Gregori, L., Xu, Y. & Chau, V. (1993). The bacterially expressed yeast *CDC34* gene product can undergo autoubiquitination to form a multiubiquitin chain-linked protein. *J. Biol. Chem.* **268**, 5668-5675.
- Banerji, J., Sands, J., Strominger, J.L. & Spies, T. (1990). A gene pair from the human major histocompatibility complex encodes large proline-rich proteins with multiple repeated motifs and a single ubiquitin-like domain. *Proc. Natl. Acad. Sci. USA* **87**, 2374-2378.
- Barahona, I., Soares, H., Cyrne, L., Penque, D., Denoulet, P. & Rodrigues-Pousada, C. (1988). Sequence of one α - and two β -tubulin genes of *Tetrahymena pyriformis*. *J. Mol. Biol.* **202**, 365-382.
- Barahona, I., Galego, L., Jacquet, M. & Rodrigues-Pousada, C. (1985). Cloning of two β -tubulin related genes in *Tetrahymena pyriformis*. *FEBS Lett.* **188**, 227-232.
- Baroin-Tourancheau, A., Delgado, P., Perasso, R. & Adoutte, A. (1992). A broad molecular phylogeny of ciliates: identification of major evolutionary trends and radiations within the phylum. *Proc. Natl. Acad. Sci. USA* **89**, 9764-9768.
- Barrell, B.G., Bankier, A.T. & Drouin, J. (1979). A different genetic code in human mitochondria. *Nature* **282**, 189-194.
- Barrio, R., del Arco, A., Cabrera, H. & Arribas, C. (1994). Structure and expression of the *Drosophila* ubiquitin-80-amino-acid fusion-protein gene. *Biochem. J.* **302**, 237-244.
- Bartel, B., Wüning, I. & Varshavsky, A. (1990). The recognition component of the N-end rule pathway. *EMBO J.* **9**, 3179-3189.
- Bartling, D., Rehling, P. & Weiler, E.W. (1993). Functional expression and molecular characterization of AtUBC2-1, a novel ubiquitin-conjugating enzyme (E2) from *Arabidopsis thaliana*. *Plant Mol. Biol.* **23**, 387-396.
- Berger, J.D. (1973). Selective inhibition of DNA synthesis in macronuclear fragments in *Paramecium aurelia* exconjugants and its reversal during macronuclear regeneration. *Chromosoma* **44**, 33-48.
- Binet, M.-N., Weil, J.-H. & Tessier, L.-H. (1991). Structure and expression of sunflower ubiquitin genes. *Plant Mol. Biol.* **17**, 395-407.
- Birnboim, H.C. & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**, 1513-1523.
- Bishoff, S.T. & Schwartz, L.M. (1990). Characterization of a ubiquitin-fusion gene from the tobacco hawkmoth, *Manduca sexta*. *Nucleic Acids Res.* **20**, 6039-6043.
- Blackburn, E.H. (1991). Structure and function of telomeres. *Nature* **350**, 569-573.
- Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, J.H. & Falkow, S. (1977). Construction and characterization of new cloning vectors. II. A multipurpose cloning system. *Gene* **2**, 95-113.

- Bond, U. & Schlesinger, M.J. (1985). Ubiquitin is a heat-shock protein in chicken embryo fibroblasts. *Mol. Cell. Biol.* **5**, 949-956.
- Bond, U. & Schlesinger, M.J. (1986). The chicken ubiquitin gene contains a heat-shock promoter and expresses an unstable mRNA in heat-shocked cells. *Mol. Cell. Biol.* **6**, 4602-4610.
- Bonifacino, J.S. & Lippincott-Schwartz, J. (1991). Degradation of proteins within the endoplasmatic reticulum. *Curr. Biol.* **3**, 592-600.
- Boyer, H.W. & Roulland-Dussoix, D. (1969).. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* **41**, 459-472.
- Brunk, C.F. & Sadler, L.A. (1990). Characterization of the promoter region of *Tetrahymena* genes. *Nucleic Acids Res.* **18**, 323-329.
- Brunk, C.F. & Conover, R.K. (1985). Elimination of micronuclear specific DNA sequences early in anlagen development. *Mol. Cell. Biol.* **5**, 93-98.
- Bruns, P.J. & Brussard, T.E.B. (1981). Nullisomic *Tetrahymena*: eliminating germinal chromosomes. *Science* **213**, 549-551.
- Bruns, P.J. (1986). Genetic organization of *Tetrahymena*. Em: Gall, J.G. (ed). *The molecular biology of ciliated protozoa*, Academic Press, London, pp. 27-44.
- Bruns, P.J., Brussard, T.B. & Merriam, E.V. (1983). Nullisomic *Tetrahymena*. II. A set of nullisomics define the germinal chromosomes. *Genetics* **104**, 257-270.
- Bullock, R.H., Fernandez, J.M. & Short, J.M. (1987). XL1-Blue: high efficiency plasmid transforming *recA Escherichia coli* strain with β -galactosidase section. *Biotechniques* **5**, 376-379.
- Burke, T.J., Callis, J. & Vierstra, R.D. (1988). Characterization of a polyubiquitin gene from *Arabidopsis thaliana*. *Mol. Gen. Genet.* **213**, 435-443.
- Butt, T.R., Jonnalagadda, S., Monia, B.P., Sternberg, E.J., Marsh, J.A., Stadel, J.M., Ecker D.J. & Crooke, S.T. (1989). Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86**, 2540-2544.
- Cabrera, H.L., Barrio, R. & Arribas, C. (1992). Structure and expression of the *Drosophila* ubiquitin-52-amino acid fusion protein gene. *Biochem. J.* **286**, 281-288.
- Callahan, R.C., Shalke, G. & Gorovsky, M.A. (1984). Developmental rearrangements associated with a single type of expressed α -tubulin gene in *Tetrahymena*. *Cell* **36**, 441-445.
- Callis, J., Raasch, J.A. & Vierstra, R.D. (1990). Ubiquitin extension proteins of *Arabidopsis thaliana*. *J. Biol. Chem.* **265**, 12486-12493.
- Cameron, I.L. (1973). Growth characteristics of *Tetrahymena*. Em: Elliott, A.M. (ed). *Biology of Tetrahymena*. Dowden, Hutchinson & Ross, Inc., Pennsylvania, pp. 199-226.
- Caron, F. & Meyer, E. (1985). Does *Paramecium primaurelia* use a different genetic code in its macronucleus? *Nature* **314**, 185-188.
- Caron, F. & Meyer, E. (1989). Molecular basis of surface antigen variation in *Paramecium*. *Annu. Rev. Microbiol.* **43**, 23-42.
-

-
- Caron, F. (1992). A high degree of macronuclear chromosome polymorphism is generated by variable DNA rearrangements in *Paramecium primaurelia* during macronuclear differentiation. *J. Mol. Biol.* **225**, 661-678.
- Casci, R.J. & Hufnagel, L.A. (1988). Cell pairing during mating in *Tetrahymena*: I. does phagocytosis or a cell surface receptor participate in Con A block? *J. Protozool.* **35**, 424-430.
- Cech, T.R. & Brehm, S.L. (1981). Replication of the extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila*. *Nucleic Acids Res.* **9**, 3532-3543.
- Cech, T.R. & Rio, D.C. (1979). Localization of transcribed regions on the extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila* by R-loop mapping. *Proc. Natl. Acad. Sci. USA* **76**, 5051-5055.
- Cech, T.R. (1987). The chemistry of self-splicing RNA and RNA enzymes. *Science* **236**, 1532-1539.
- Cech, T.R. (1988). Conserved sequences and structures of group I introns: building an active site for RNA catalysis - a review. *Gene* **73**, 259-271.
- Cech, T.R. (1992). Catalytic RNA: structure and mechanism. *Biochem. Soc. Trans.* **21**, 229-234.
- Cech, T.R., Zaug, A.J. & Grabowski, P.J. (1981). *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* **27**, 487-496.
- Chau, V., Tobias, J.W., Bachmair, A., Marriot, D., Ecker, D.J., Gonda, D.K. & Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a target short-lived protein. *Science* **243**, 1576-1583.
- Chen, K. & Rubenstein, I. (1991). Characterization of the structure and transcription of an ubiquitin fusion gene from maize. *Gene* **107**, 205-212.
- Chen, P., Johnson, P., Sommer, T., Jentsch, S. & Hochstrasser, M. (1993). Multiple ubiquitin-conjugating enzymes participate in the *in vivo* degradation of the yeast MAT α 2 repressor. *Cell* **74**, 357-369.
- Cherry, J.M. & Blackburn, E.H. (1985). The internally located telomeric sequences in the germ-line chromosomes of *Tetrahymena* are at the ends of transposon-like elements. *Cell* **43**, 747-758.
- Christensen, A.H. & Quail, P.H. (1989). Sequence analysis and transcriptional regulation by heat shock of polyubiquitin transcripts from maize. *Plant Mol. Biol.* **12**, 619-632.
- Christensen, A.H., Sharrock, R.A. & Quail, P.H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* **18**, 675-689.
- Chung, S.-H. & Swindle, J. (1990). Linkage of the calmodulin and ubiquitin *loci* in *Trypanosoma cruzi*. *Nucleic Acids Res.* **18**, 4561-4569.
- Ciechanover, A., Elias, S., Heller, H. & Hershko, A. (1982). "Covalent affinity" purification of ubiquitin-activating enzyme. *J. Biol. Chem.* **257**, 2537-2542.
-

- Ciechanover, A., Ferber, S., Ganoth, D., Elias, S., Hershko, A. & Arfin, S. (1988). Purification and characterization of arginyl-tRNA-protein transferase from rabbit reticulocytes. *J. Biol. Chem.* **263**, 11155-11167.
- Ciechanover, A., Finley, D. & Varshavsky, A. (1984). Ubiquitin-dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell* **37**, 57-66.
- Ciechanover, A., Hod, Y. & Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochem. Biophys. Res. Commun.* **81**, 1100-1105.
- Cleffman, G. (1980). Chromatin elimination and the genetic organization of the macronucleus in *Tetrahymena thermophila*. *Chromosoma* **78**, 313-325.
- Cohen, S.N. & Chang, A.C.Y. (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by the R-factor DNA. *Proc. Natl. Acad. Sci.* **69**, 2110-2114.
- Conover, R.K. & Brunk, C.F. (1986). Macronuclear DNA molecules of *Tetrahymena thermophila*. *Mol. Cell. Biol.* **6**, 900-905.
- Cook, J.C. & Chock, P.B. (1991a). Immunocytochemical localization of ubiquitin-activating enzyme in the cell nucleus. *Biochem. Biophys. Res. Commun.* **174**, 564-571.
- Cook, J.C. & Chock, P.B. (1991b). Association of ubiquitin-activating enzyme with HeLa cell chromosomes during mitosis. *Proc. Natl. Acad. Sci. USA* **88**, 11388-11392.
- Cook, J.C. & Chock, P.B. (1992). Isoforms of mammalian ubiquitin-activating enzyme. *J. Biol. Chem.* **267**, 24315-24321.
- Corliss, J.O. (1973). History, taxonomy, ecology and evolution of species of *Tetrahymena*. Em: Elliott, A.M. (ed). *Biology of Tetrahymena*. Dowden, Hutchinson & Ross, Inc., Pennsylvania, pp. 1-55.
- Corliss, J.O. (1979). Em: *The ciliated protozoa: characterization, classification and guide to the literature*. Pergamon, Oxford, pp. 1-455.
- Cowland, J.B., Wiborg, O. & Vuust, J. (1988). Human ubiquitin genes: one member of the UbB gene subfamily is a tetrameric non-processed pseudogene. *FEBS Lett.* **231**, 187-191.
- Csank, C., Taylor, F.M. & Martindale, D.W. (1990). Nuclear pre-mRNA introns: analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes. *Nucleic Acids Res.* **18**, 5133-5141.
- Dahlmann, B., Kopp, F., Kuehn, L., Nidel, B., Pfeifer, G., Hegerl, R. & Baumeister, W. (1989). The multicatalytic proteinase (prosome) is ubiquitous from eukaryotes to archeobacteria. *FEBS Lett.* **251**, 125-131.
- Dale, R.M.K., McClure, B.A. & Houchins, J.P. (1985). A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* **13**, 31-40.
-

- Davie, J.R. & Murphy, L.C. (1990). Level of ubiquitinated histone H2B in chromatin is coupled to ongoing transcription. *Biochem.* **29**, 4752-4757.
- Davies, K.M. & King, G.A. (1993). Isolation and characterization of *Asparagus officinalis* L. cDNA clones encoding two forms of ubiquitin mRNA. *New Zealand J. Crop Hort. Sci.* **21**, 153-159.
- Davis, M.C., Ward, J.G., Herrick, G. & Allis, C.D. (1992). Programmed nuclear death: apoptotic-like degradation of specific nuclei in conjugating *Tetrahymena*. *Dev. Biol.* **154**, 419-432.
- Dice, J.F. (1990). Peptide sequences that target cytosolic proteins for lysosomal proteolysis. *Trends Biochem. Sci.* **15**, 305-309.
- Djaballah, H., Rowe, A.J., Harding, S.E. & Rivett, A.J. (1993). The multicatalytic proteinase complex (proteasome): structure and conformational changes associated with changes in proteolytic activity. *Biochem. J.* **292**, 857-862.
- Doak, T.G., Doerder, F.P., Jahn, C.L. & Herrick, G. (1994). A proposed superfamily of transposable genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proc. Natl. Acad. Sci. USA* **91**, 942-946.
- Doerder, F.P. (1979). Regulation of macronuclear DNA content in *Tetrahymena thermophila*. *J. Protozool.* **26**, 28-35.
- Dohmen, J., Madura, K., Bartel, B. & Varshavsky, A. (1991). The N-end rule is mediated by the UBC2 (RAD6) ubiquitin-conjugating enzyme. *Proc. Natl. Acad. Sci. USA* **356**, 215-221.
- Douglass, J., Civelli, O. & Herbert, E. (1984). Polyprotein gene expression: generation of diversity of neuroendocrine peptides. *Ann. Rev. Biochem.* **53**, 665-715.
- Driscoll, J., Frydman, J. & Goldberg, A.L. (1992). An ATP-stabilized inhibitor of the proteasome is a component of the 1500-kDa ubiquitin conjugate-degrading complex. *Proc. Natl. Acad. Sci.* **89**, 4986-4990.
- Driscoll, J. & Goldberg, A.L. (1990). The proteasome (multicatalytic protease) is a component of the 1500 kDa proteolytic complex which degrades ubiquitin-conjugated proteins. *J. Biol. Chem.* **265**, 4789-4792.
- Dunigan, D.D., Dietzgen, R.G., Schoelz, J.E. & Zaitlin, M. (1988). Tobacco mosaic virus particles contain ubiquitinated coat protein subunits. *Virology* **165**, 310-312.
- Dworkin-Rastl, E., Shrutkowski, A. & Dworkin, M.B. (1984). Multiple ubiquitin mRNAs during *Xenopus laevis* development contain tandem repeats of the 76 amino acid coding sequence. *Cell* **39**, 321-325.
- Ecker, D.J., Butt, T.R., Marsh, J., Sternberg, E.J., Margolis, N., Monia, B.P., Jonnalagadda, S., Khan, M.I., Weber, P.L., Mueller, L. & Crooke, S.T. (1987). Gene synthesis, expression, structures, and functional activities of site-specific mutants of ubiquitin. *J. Biol. Chem.* **262**, 14213-14221.
- Ecker, D.J., Stadel, J.M., Butt, T.R., Marsh, J.A., Monia, B.P., Powers, D.A., Gorman, J.A., Clark, P.E., Warren, F., Shatzman, A. & Crooke, S.T. (1989). Increasing gene expression in yeast by fusion to ubiquitin. *J. Biol. Chem.* **264**, 7715-7719.

- Ellison, K.S., Gwozd, T., Prendergast, J.A., Paterson, M.C. & Ellison, M.J. (1991). A site-directed approach for constructing temperature-sensitive ubiquitin-conjugating enzymes reveals a cell cycle function and growth function for RAD6. *J. Biol. Chem.* **266**, 24116-24120.
- Emori, Y., Tsukahara, T., Kawasaki, H., Ishiura, S., Sugita, H. & Suzuki, K. (1991). Molecular cloning and functional analysis of three subunits of yeast proteasome. *Mol. Cell. Biol.* **11**, 344-353.
- Engberg, J., Anderson, P. & Leick, V. (1976). Free ribosomal DNA molecules from *Tetrahymena pyriformis* GL are giant palindromes. *J. Mol. Biol.* **104**, 455-470.
- Engberg, J., Christiansen, G. & Leick, V. (1974). Autonomous rDNA molecules containing single copies of ribosomal RNA genes in the macronucleus of *Tetrahymena pyriformis*. *Biochem. Biophys. Res. Commun.* **59**, 1356-1365.
- Epstein, L.M. & Forney, J.D. (1984). Mendelian and non-Mendelian mutations affecting surface antigen expression in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **4**, 1583-1590.
- Eytan, E., Armon, T., Heller, H., Beck, S. & Hershko, A. (1993). Ubiquitin C-terminal hydrolase activity associated with the 26S protease complex. *J. Biol. Chem.* **268**, 4668-4674.
- Eytan, E., Ganoth, D., Armon, T. & Hershko, A. (1989). ATP-dependent incorporation of 20S protease into the 26S complex that degrades proteins conjugated to ubiquitin. *Proc. Natl. Acad. Sci.* **86**, 7751-7755.
- Falkenburg, D., Dworniczak, B., Faust, D.M. & Bautz, E.K.F. (1987). RNA polymerase II of *Drosophila*. Relation of its 140,000 M_r subunit to the β subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **195**, 929-937.
- Feilotter, H., Lingner, C., Rowley, R. & Young, P.G. (1992). Regulation of the G₂-mitosis transition. *Biochem. Cell Biol.* **70**, 954-971.
- Ferber, S. & Ciechanover, A. (1987). Role of arginine-tRNA in protein degradation by the ubiquitin pathway. *Nature* **326**, 808-811.
- Finch, J.S., St. John, T., Krieg, P., Bonham, K., Smith, H.T., Fried, V.A. & Bowden, G.T. (1992). Overexpression of 3 ubiquitin genes in mouse epidermal tumors is associated with enhanced cellular proliferation and stress. *Cell Growth Differ.* **3**, 269-278.
- Finley, D. & Chau, V. (1991). Ubiquitination. *Ann. Rev. Cell Biol.* **7**, 25-70.
- Finley, D., Bartel, B. & Varshavsky, A. (1989). The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* **338**, 394-401.
- Finley, D., Ciechanover, A. & Varshavsky, A. (1984). Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell* **37**, 43-55.
- Fleischmann, J. & Campbell, D.A. (1994). Expression of the *Leishmania tarentolae* ubiquitin-encoding and mini-exon genes. *Gene* **144**, 45-51.
- Foster, L.M., Loftus, M.G. & Ross, I.K. (1990). A novel form of ubiquitin found in the basidiomycete fungus, *Coprinus congregatus*. *Nucleic Acids Res.* **18**, 6449.
-

- Garbarino, J.E., Rockhold, D.R. & Belknap, W.R. (1992). Expression of stress-responsive ubiquitin genes in potato tubers. *Plant Mol. Biol.* **20**, 235-244.
- Gausung, K. & Barkardottir, R. (1986). Structure and expression of ubiquitin genes in higher plants. *Eur. J. Biochem.* **158**, 57-62.
- Gausung, K. & Jensen, C.B. (1990). Two ubiquitin-long-tail fusion genes arranged as closely spaced direct repeats in barley. *Gene* **94**, 165-171.
- Genschik, P., Parmentier, Y., Durr, A., Marbach, J., Criqui, M.-C., Janet, E. & Fleck, J. (1992). Ubiquitin genes are differentially regulated in protoplast-derived cultures of *Nicotiana sylvestris* and in response to various stresses. *Plant Mol. Biol.* **20**, 897-910.
- Gilmour, D.S., Thomas, G.H. & Elgin, S.C.R. (1989). *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science* **245**, 1487-1490.
- Giordia, R. & Ennis, H.L. (1987). Structure of two developmentally regulated *Dictyostelium discoideum* ubiquitin genes. *Mol. Cell. Biol.* **6**, 2097-2103.
- Glotzer, M., Murray, A.W. & Krishner, M.W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132-138.
- Godiska, R. & Yao, M.-C. (1990). A programmed site-specific DNA rearrangement in *Tetrahymena thermophila* requires flanking polypurine tracts. *Cell* **61**, 1237-1246.
- Godiska, R., James, C. & Yao, M.-C. (1993). A distant 10-bp sequence specifies the boundaries of a programmed DNA deletion in *Tetrahymena*. *Genes Dev.* **7**, 2357-2365.
- Goebel, M.G., Yochem, J., Jentsch, S., McGrath, J.P., Varshavsky, A. & Byers, B. (1988). The yeast cell cycle gene *CDC34* encodes a ubiquitin-conjugating enzyme. *Science* **241**, 1331-1335.
- Goldberg, A.L. & Rock, K.L. (1992). Proteolysis, proteasomes, and antigen presentation. *Nature* **357**, 375-379.
- Goldknopf, I.L. & Busch, H. (1977). Isopeptide linkage between non-histone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proc. Natl. Acad. Sci. USA* **74**, 864-868.
- Goldstein, G., Scheid, M., Hammerling, U., Boyse, E.A., Schlesinger, D.H. & Niall, H.D. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci. USA* **72**, 11-15.
- Gonda, D.K., Bachmair, A., Wüning, I., Tobias, J.W., Lane, W.S. & Varshavsky, A. (1989). Universality and structure of the N-end rule. *J. Biol. Chem.* **264**, 16700-16712.
- Gonen, H., Schwartz, A.L. & Ciechanover, A. (1991). Purification and characterization of a novel protein that is required for the degradation of N- α -acetylated proteins by the ubiquitin system. *J. Biol. Chem.* **266**, 19221-19231.
- Gorovsky, M.A. (1973). Macro- and micronuclei of *Tetrahymena pyriformis*: a model system for studying the structure and function of eukaryotic nuclei. *J. Protozool.* **20**, 19-25.

- Gorovsky, M.A. (1980). Genome organization and reorganization in *Tetrahymena*. *Annu. Rev. Genet.* **14**, 203-239.
- Gorovsky, M.A., Glover, C., Johmann, C.A., Keevert, J.B., Mathis, D.J. & Samuelson, M. (1977). Histones and chromatin structure in *Tetrahymena* macro- and micronuclei. *Cold Spring Harbor Symp. Quant. Biol.* **42**, 493-503.
- Gorovsky, M.A., Yao, M.-C., Keevert, J.B. & Pleger, G.L. (1975). Isolation of micro- and macronuclei of *Tetrahymena pyriformis*. *Methods Cell Biol.* **9**, 311-327.
- Gottschling, D.E. & Zakian, V.A. (1986). Telomere proteins specific recognition and protection of the natural termini of *Oxytricha* macronuclear DNA. *Cell* **47**, 195-205.
- Grabowski, P.J., Zaug, A.J. & Cech, T.R. (1981). The intervening sequence of the ribosomal precursor is converted to a circular RNA in isolated nuclei in *Tetrahymena*. *Cell* **23**, 467-476.
- Graham, R.W., Jones, D. & Candido, E.P.M. (1989). *UbiA*, the major polyubiquitin locus in *Caenorhabditis elegans*, has unusual structural features and is constitutively expressed. *Mol. Cell. Biol.* **9**, 268-277.
- Gray, J.T., Celander, D.W., Price, C.M. & Cech, T.R. (1991). Cloning and expression of genes for the *Oxytricha* telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell* **67**, 807-814.
- Greider, C.W. & Blackburn, E.H. (1987). The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* **51**, 887-898.
- Greider, C.W. & Blackburn, E.H. (1989). A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* **337**, 331-337.
- Grenfell, S.J., Trausch-Azar, J.S., Handley-Gearhart, P.M., Ciechanover, A. & Schwartz, A.L. (1994). Nuclear localization of the ubiquitin-activating enzyme, E1, is cell-cycle-dependent. *Biochem. J.* **300**, 701-708.
- Greslin, A.F., Prescott, D.M., Oka, Y., Loukin, S.H. & Chappell, J.C. (1989). Reordering of nine exons is necessary to form a functional actin gene in *Oxytricha nova*. *Proc. Natl. Acad. Sci. USA* **86**, 6264-6268.
- Grivell, L.A. (1986). Deciphering divergent codes. *Nature* **324**, 109-110.
- Gronenborg, A.M. & Clore, G.M. (1991). Similarity of protein G and ubiquitin. *Science* **254**, 581-582.
- Groudine, M., Peretz, M. & Weintraub, H. (1981). Transcriptional regulation of haemoglobin switching in chicken embryos. *Mol. Cell. Biol.* **1**, 281-288.
- Grziwa, A., Baumeister, W., Dahlmann, B. & Kopp, F. (1991). Localization of subunits in proteasomes from *Thermoplasma acidophilum* by immunoelectron microscopy. *FEBS Lett.* **290**, 186-190.
- Guarino, L.A. (1990). Identification of a viral gene encoding a ubiquitin-like protein. *Proc. Natl. Acad. Sci. USA* **87**, 409-413.
-

- Guerreiro, P., Neves, A. & Rodrigues-Pousada, C. (1993). Clusters of 5S rRNAs in the intergenic region of ubiquitin genes in *Tetrahymena pyriformis*. *Biochem. Biophys. Acta* **1216**, 137-139.
- Haas, A.L. & Rose, I.A. (1982). The mechanism of ubiquitin activating enzyme. *J. Biol. Chem.* **257**, 10329-10337.
- Haas, A.L., Ahrens, P., Bright, P.M. & Ankel, H. (1987). Interferon induces a 15-kilodalton protein exhibiting marked homology to ubiquitin. *J. Biol. Chem.* **262**, 11315-11323.
- Haas, A.L., Reback, P.M., Pratt, G. & Rechsteiner, M. (1990). Ubiquitin-mediated degradation of histone H3 does not require the substrate-binding ubiquitin protein ligase, E3, or attachment of polyubiquitin chains. *J. Biol. Chem.* **265**, 21664-21669.
- Hadari, T., Warms, J.B.V., Rose, I.A. & Hershko, A. (1992). A ubiquitin C-terminal isopeptidase that acts on polyubiquitin chains - role in protein degradation. *J. Biol. Chem.* **267**, 719-727.
- Hamilton, E.P., Suhr-Jessen, P.B. & Orias, E. (1988). Pronuclear fusion failure: an alternative conjugation pathway in *Tetrahymena thermophila*, induced by vinblastine. *Genetics* **118**, 627-636.
- Hanyu, N., Kuchino, Y. & Nishimura, S. (1986). Dramatic events in ciliate evolution: alteration of UAA and UAG termination codons to glutamine codons due to anticodon mutations in two *Tetrahymena* tRNAs^{Gln}. *EMBO J.* **5**, 1307-1311.
- Harper, D.S. & Jahn, C.L. (1989). Differential use of termination codons in ciliated protozoa. *Proc. Natl. Acad. Sci. USA* **86**, 3253-3256.
- Harrington, L. & Greider, C. (1991). Telomerase primer specificity and chromosome healing. *Nature* **353**, 451-453.
- Harumoto, T. (1986). Induced change in a non-Mendelian determinant by transplantation of macronucleoplasm in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **6**, 3498-3501.
- Hatfield, P.M. & Vierstra, R.D. (1989). Ubiquitin-dependent proteolytic pathway in wheat germ: isolation of multiple forms of ubiquitin-activating enzyme, E1. *Biochem.* **28**, 735-742.
- Hatfield, P.M. & Vierstra, R.D. (1992). Multiple forms of ubiquitin-activating enzyme E1 from wheat. *J. Biol. Chem.* **267**, 14799-14803.
- Hauser, L.J., Roberson, A.E. & Olins, D.E. (1991). Structure of the macronuclear polyubiquitin gene in *Euplotes*. *Chromosoma* **100**, 386-394.
- Hayashi, T., Noga, M. & Matsuda, M. (1994). Nucleotide sequence and expression of the rat polyubiquitin mRNA. *Biochem. Biophys. Acta* **1218**, 232-234.
- Heinemeyer, W., Gruhler, A., Möhrle, Y., Mahé, Y. & Wolf, D.H. (1993). PRE2, highly homologous to the human major histocompatibility complex-linked *RING10* gene, codes for a yeast proteasome subunit necessary for chymotryptic activity and degradation of ubiquitinated proteins. *J. Biol. Chem.* **268**, 5115-5120.
- Heinemeyer, W., Kleinschmidt, J.A., Saidowsky, J., Escher, C. & Wolf, D.H. (1991). Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase:

- mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival. *EMBO J.* **10**, 555-562.
- Helftenbein, E. & Müller, E. (1988). Both α -tubulin genes are transcriptionally active in *Stylonychia lemnae*. *Curr. Genet.* **13**, 425-432.
- Helftenbein, E. (1985). Nucleotide sequence of a macronuclear DNA molecule coding for α -tubulin from the ciliate *Stylonychia lemnae*. Special codon usage: TAA is not a translation termination codon. *Nucleic Acids Res.* **13**, 415-433.
- Heller, H. & Hershko, A. (1990). A ubiquitin-protein ligase specific for type III protein substrates. *J. Biol. Chem.* **265**, 6532-6535.
- Herrick, G., Cartinhour, S.W., Williams, K.R. & Kotter, K.P. (1987). Multiple sequence versions of the *Oxytricha fallax* 81-MAC alternate processing family. *J. Protozool.* **34**, 429-434.
- Herrick, G., Cartinhour, S., Dawson, D., Ang, D., Sheets, R., Lee, A. & Williams, K. (1985). Mobile genetic elements bounded by C4A4 telomeric repeats in *Oxytricha fallax*. *Cell* **43**, 759-768.
- Hershko, A. & Ciechanover, A. (1992). The ubiquitin system for protein degradation. *Ann. Rev. Biochem.* **61**, 761-807.
- Hershko, A. (1991). The ubiquitin pathway for protein degradation. *Trends Biochem. Sci.* **16**, 265-268.
- Hershko, A., Ganoth, D., Pehrson, J., Palazzo, R.E. & Cohen, L.H. (1991). Methylated ubiquitin inhibits cyclin degradation in clam embryo extracts. *J. Cell Biol.* **266**, 16376-16379.
- Herschlag, D. & Cech, T.R. (1990). DNA cleavage catalysed by the ribozyme from *Tetrahymena*. *Nature* **344**, 405-409.
- Herskowitz, I., Rine, J. & Strathern, J. (1992). Mating-type determination and mating-type interconversion in *Saccharomyces cerevisiae*. In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, vol.2, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 583-656.
- Hilt, W., Enenkel, C., Gruhler, A., Singer, T. & Wolf, D.H. (1993). The *PRE4* gene codes for a subunit of the yeast proteasome necessary for a peptidylglutamyl-peptide-hydrolyzing activity: mutations link the proteasome to stress- and ubiquitin-dependent proteolysis. *J. Biol. Chem.* **268**, 3479-3486.
- Hingamp, P.M., Arnold, J.E., Mayer, R.J. & Dixon, L.K. (1992). A ubiquitin-conjugating enzyme encoded by African Swine Fever Virus. *EMBO J.* **11**, 361-366.
- Hirel, P.H., Schmitter, J.M., Dessen, P., Fayat, G. & Blanquet, S. (1989). Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. USA* **86**, 8247-8251.
- Hochstrasser, M. & Varshavsky, A. (1990). *In vivo* degradation of a transcriptional regulator: the yeast $\alpha 2$ repressor. *Cell* **61**, 697-708.
- Hochstrasser, M. (1992). Ubiquitin and intracellular protein degradation. *Curr. Opinion Cell. Biol.* **4**, 1024-1031.

- Hochstrasser, M., Ellison, M.J., Chau, V. & Varshavsky, A. (1991). The short-lived MAT α 2 transcriptional regulator is ubiquitinated *in vivo*. *Proc. Natl. Acad. Sci. USA* **88**, 4606-4610.
- Hodgins, R.R.W., Ellison, K.S. & Ellison, M.J. (1992). Expression of a ubiquitin derivate that conjugates to protein irreversibly produces phenotypes consistent with a ubiquitin deficiency. *J. Biol. Chem.* **267**, 8807-8812.
- Hoffman, N.E., Ko, K., Milkowski, D. & Pichersky, E. (1991). Isolation and characterization of tomato cDNA and genomic clones encoding the ubiquitin gene *ubi3*. *Plant Mol. Biol.* **17**, 1189-1201.
- Hoffmann, W., Bach, T.C., Seliger, H. & Kreil, G. (1983). Biosynthesis of caerulein in the skin of *Xenopus laevis* partial sequences of precursors as deduced from cDNA clones. *EMBO J.* **2**, 111-114.
- Holloway, S.L., Glotzer, M., King, R.W. & Murray, A.W. (1993). Anaphase is initiated by proteolysis rather than by the inactivating of maturation-promoting factor. *Cell* **73**, 1393-1402.
- Holmes, D.S. & Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**, 193-197.
- Horowitz, S. & Gorovsky, M.A. (1985). An unusual genetic code in nuclear genes of *Tetrahymena*. *Proc. Natl. Acad. Sci. USA* **82**, 2452-2455.
- Hough, R., Pratt, G. & Rechsteiner, M. (1987). Purification of two high molecular weight proteases from rabbit reticulocyte lisate. *J. Biol. Chem.* **262**, 8303-8313.
- Huibregtse, J.M., Scheffner, M. & Howley, P. (1991). A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J.* **10**, 4129-4135.
- Hunt, T. (1991). Destruction's our delight. *Nature* **349**, 100-101.
- Imai, N., Kaneda, S., Nagai, Y., Seno, T., Ayusawa, D., Hanaoka, F. & Yamao, F. (1992). Cloning and sequence of a functionally active cDNA encoding the mouse ubiquitin-activating enzyme E1. *Gene* **118**, 279-282.
- Jabben, M., Shanklin, J. & Vierstra, R.D. (1989). Ubiquitin-phytochrome conjugates. Pool dynamics during *in vivo* phytochrome degradation. *J. Biol. Chem.* **264**, 4998-5005.
- Jacinto, A. (1992). *Contribuição para o estudo dos genes ubiquitina em Lupinus albus*. Tese de licenciatura, Faculdade de Ciências de Lisboa, Lisboa.
- Jacinto, A., Neves, A.M., Vassilevskaia, T.D., Ricardo, C.P. & Rodrigues-Pousada, C. (1994). Cloning and characterization of two ubiquitin::79-amino-acid extension protein-encoding fusion genes from *Lupinus albus*. *Gene* **139**, 201-205.
- Jahn, C.L., Krikau, M.F. & Shyman, S. (1989). Developmentally coordinated en masse excision of a highly repetitive element in *E. crassus*. *Cell* **59**, 1009-1018.
- Jamet, E., Durr, A., Parmentier, Y., Criqui, M.-C. & Fleck, J. (1990). Is ubiquitin involved in the dedifferentiation of higher plant cells? *Cell Differentiation Dev.* **29**, 37-46.

- Jaraczewski, J.W. & Jahn, C.L. (1993). Elimination of Tec elements involves a novel excision process. *Genes Dev.* **7**, 95-105.
- Jennissen, H.P., Botzet, G., Majetschak, M., Laub, M., Ziegenhagen, R. & Demiroglou, A. (1992). Ca^{2+} -dependent ubiquitination of calmodulin in yeast. *FEBS Lett.* **296**, 51-56.
- Jenson, J., Goldstein, G. & Breslow, E. (1980). Physical-chemical properties of ubiquitin. *Biochem. Biophys. Acta* **624**, 378-385.
- Jentsch, S. (1992). The ubiquitin-conjugation system. *Ann. Rev. Genet.* **26**, 179-207.
- Jentsch, S., McGrath, J.P. & Varshavsky, A. (1987). The DNA repair gene RAD6 encodes a ubiquitin-conjugating enzyme. *Nature* **329**, 131-134.
- Jentsch, S., Seufert, W. & Hauser, H.-P. (1991). Genetic analysis of the ubiquitin system. *Biochem. Biophys. Acta* **1089**, 127-139.
- Jentsch, S., Seufert, W., Sommer, T. & Reins, H.-A. (1990). Ubiquitin-conjugating enzymes: novel regulators of eukaryotic cells. *Trends Biochem. Sci.* **15**, 195-198.
- Jofuku, K.D. & Goldberg, R.B. (1988). Analysis of plant genome structure. In *Plant Molecular Biology: a practical approach*. Shaw, C.H. (eds.), IRL Press, Oxford, pp. 37-66.
- Johnson, E.S., Gonda, D.K. & Varshavsky, A. (1990). *Cis-trans* recognition and subunit-specific degradation of short-lived proteins. *Nature* **346**, 287-291.
- Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268-272.
- Jones, D. & Candido, E.P.M. (1993). Novel ubiquitin-like ribosomal protein fusion genes from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *J. Biol. Chem.* **268**, 19545-19551.
- Jungmann, J., Reins, H.-A., Schobert, C. & Jentsch, S. (1993). Resistance to cadmium mediated by ubiquitin-dependent proteolysis. *Nature* **361**, 369-371.
- Kalderon, D., Roberts, B.L., Richardson, W.D. & Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499-509.
- Kaney, A.R. & Speare, V.J. (1983). An amiconucleate mutant of *Tetrahymena thermophila*. *Exp. Cell Res.* **143**, 461-467.
- Kapler, G.M. & Blackburn, E.H. (1994). A weak germ-line excision mutation blocks developmentally controlled amplification of the rDNA minichromosome of *Tetrahymena thermophila*. *Genes Dev.* **8**, 84-95.
- Karrer, K.M. & Gall, J.G. (1976). The macronuclear ribosomal DNA of *Tetrahymena pyriformis* is a palindrome. *J. Mol. Biol.* **104**, 421-453.
- Karrer, K.M. (1983). Germ line specific DNA sequences are present on all five micronuclear chromosomes of *Tetrahymena thermophila*. *Mol. Cell. Biol.* **3**, 1909-1919.
- Karrer, K.M. (1986). The nuclear DNAs of holotrichous ciliates. Em: Gall, J.G. (ed). *The molecular biology of ciliated protozoa*, Academic Press, London, pp. 85-110.
-

- Karrer, K., Stein-Gavens, S. & Allito, B.A. (1984). Micronucleus-specific DNA sequences in an amiconucleate mutant of *Tetrahymena*. *Dev. Biol.* **105**, 121-129.
- Kas, K., Michiels, L. & Merregaert, J. (1992). Genomic structure and expression of the human *fau* gene: encoding the ribosomal protein fused to a ubiquitin-like protein. *Biochem. Biophys. Res. Comm.* **187**, 927-933.
- Katoh, M., Hirono, M., Takemasa, T., Kimura, M. & Watanabe, Y. (1993). A micronucleus-specific sequence exists in the 5'-upstream region of calmodulin gene in *Tetrahymena thermophila*. *Nucleic Acids Res.* **21**, 2409-2414.
- Kaufmann, J., Florian, V. & Klein, A. (1992). TGA cysteine codons and intron sequences in conserved and non-conserved positions are found in macronuclear RNA polymerase genes of *Euplotes octocarinatus*. *Nucleic Acids Res.* **20**, 5985-5989.
- Kawagishi, M., Yamagishi, M. & Ishihama, A. (1993). Cloning and sequence determination of the *Schizosaccharomyces pombe rpb2* gene encoding the subunit 2 of RNA polymerase II. *Nucleic Acids Res.* **21**, 469-473.
- Kawaguchi, Y., Honda, H., Taniguchi-Morimura, J. & Iwasaki, S. (1989). The codon CUG is read as serine in an asporogenic yeast *Candida cylindracea*. *Nature* **341**, 164-166.
- Kawalleck, P., Somssich, I.E., Feldbrügge, M., Hahlbrock, K. & Weisshaar, B. (1993). Polyubiquitin gene expression and structural properties of the *ubi4-2* gene in *Petroselinum crispum*. *Plant Mol. Biol.* **21**, 673-684.
- Kemp, L.M. & Latchman, D.S. (1988). The herpes simplex virus type I immediate early protein ICP4 specifically induces increased transcription of the human ubiquitin B gene without affecting the ubiquitin A and C genes. *Virology* **166**, 258-261.
- Kile, J.P., Love Jr, H.D., Hubach, C.A. & Bannon, G.A. (1988). Reproducible and variable rearrangements of a *Tetrahymena thermophila* surface protein gene family occur during macronuclear development. *Mol. Cell. Biol.* **8**, 5043-5046.
- Kimmel, A.R. & Gorovsky, M.A. (1976). Numbers of 5S and tRNA genes in macro- and micronuclei of *Tetrahymena*. *Chromosoma* **54**, 327-337.
- Kirchhoff, L.V., Kim, K.S., Engman, D.M. & Donelson, J.E. (1988). Ubiquitin genes in *Trypanosomatidae*. *J. Biol. Chem.* **263**, 12698-12704.
- Klausner, R.D. & Sitia, R. (1990). Protein degradation in the endoplasmic reticulum. *Cell* **62**, 611-614.
- Klemperer, N.S., Berleth, E.S. & Pickart, C.M. (1989). A novel, arsenite-sensitive E2 of the ubiquitin pathway: purification and properties. *Biochem.* **28**, 6035-6041.
- Klobutcher, L.A. & Prescott, D.M. (1986). The special case of the hypotrichs. Em: Gall, J.G. (ed). *The molecular biology of ciliated protozoa*, Academic Press, London, pp. 111-154.
- Klobutcher, L.A., Swanton, M.T., Donini, P. & Prescott, D.M. (1981). All gene-sized molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc. Natl. Acad. Sci. USA* **78**, 3015-3019.
- Klobutcher, L.A., Turner, L.R. & LaPlante, J. (1993). Circular forms of developmentally excised DNA in *Euplotes crassus* have a heteroduplex junction. *Genes Dev.* **7**, 84-94.

- Kloetzel, J.A. (1970). Compartmentalization of the developing macronucleus following conjugation in *Stylonychia* and *Euplotes*. *J. Cell Biol.* **47**, 395-407.
- Klug, A. & Rhodes, D. (1987). Zinc finger's: a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12**, 464-469.
- Knight Jr., E., Fahey, D., Cordova, B., Hillman, M., Kutny, R., Reich, N. & Blomstrom, D. (1988). A 15-kDa interferon-induced protein is derived by COOH-terminal processing of a 17-kDa precursor. *J. Biol. Chem.* **263**, 4520-4522.
- Kolman, C.J., Toth, J. & Gonda, D.K. (1992). Identification of a portable determinant of cell cycle function within the carboxyl-terminal domain of the yeast CDC34 (UBC3) ubiquitin conjugating (E2) enzyme. *EMBO J.* **11**, 3081-3090.
- Kraulis, P.J. (1991). Similarity of protein G and ubiquitin. *Science* **254**, 581.
- Krebber, H., Wöstmann, C., Bakker-Grunwald, T. (1994). Evidence for the existence of a single ubiquitin gene in *Giardia lamblia*. *FEBS Lett.* **343**, 234-236.
- Krikau, M.F. & Jahn, C.L. (1991). Tec2, a second transposon-like element demonstrating developmentally programmed excision in *Euplotes crassus*. *Mol. Cell. Biol.* **11**, 4751-4759.
- Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. & Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**, 147-157.
- Kuchino, Y., Hanyu, N., Tashiro, F. & Nishimura, S. (1985). *Tetrahymena thermophila* glutamine tRNA and its gene that corresponds to UAA termination codon. *Proc. Natl. Acad. Sci. USA* **82**, 4758-4762.
- Kuehn, L., Dahlmann, B. & Reinauer, H. (1992). Evidence indicating that the multicatalytic proteinase of rabbit reticulocytes is not incorporated as a core enzyme into a 26S proteinase complex. *Arch. Biochem. Biophys.* **295**, 55-60.
- Kuhse, M.G., Strickland, R. & Palmer, J.D. (1990). An ancient group I intron shared by eubacteria and chloroplasts. *Science* **250**, 1570-1573.
- Kulka, R.G., Raboy, B., Schuster, R., Parag, H.A., Diamond, G., Ciechanover, A. & Marcus, M. (1988). A Chinese hamster cell cycle mutant arrested at G2 phase has a temperature-sensitive ubiquitin activating enzyme, E1. *J. Biol. Chem.* **263**, 15726-15731.
- Kumar, S., Yoshida, Y. & Noda, M. (1993). Cloning of a cDNA which encodes a novel ubiquitin-like protein. *Biochem. Biophys. Res. Comm.* **195**, 393-399.
- Kurjan, J. & Herskowitz, I. (1982). Structure of a yeast pheromone gene (*MF α*): a putative α -factor precursor contains four tandem copies of mature α -factor. *Cell* **30**, 933-943.
- Kyte, J. & Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.
- Lambowitz, A.M. & Perlman, P.S. (1990). Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. *Trends Biochem. Sci.* **15**, 440-444.
-

- Landick, R., Stewart, J. & Lee, D.N. (1990). Amino acid changes in conserved regions of the β -subunit of *Escherichia coli* RNA polymerase alter transcription pausing and termination. *Gene Dev.* **4**, 1623-1636.
- Laub, M. & Jennissen, H.P. (1991). Ubiquitination of endogenous calmodulin in rabbit tissue extracts. *FEBS Lett.* **294**, 229-233.
- Lauth, M.R., Spear, B.B., Heumann, J. & Prescott, D.M. (1976). DNA of ciliated protozoa: DNA sequence diminution during macronuclear development of *Oxytricha*. *Cell* **7**, 67-74.
- Lee, H., Simon, J.A. & Lis, J.T. (1988). Structure and expression of ubiquitin genes in *Drosophila melanogaster*. *Mol. Cell. Biol.* **8**, 4727-4735.
- Levinger, L. & Varshavsky, A. (1982). Selective arrangement of ubiquitinated and D1 protein-containing nucleosomes within the *Drosophila* genome. *Cell* **28**, 375-385.
- Leyser, H.M.O., Lincoln, C.A., Timpfe, C., Lammer, D., Turner, J. & Estelle, M. (1993). *Arabidopsis* auxin-resistance gene *AXR1* encodes a protein related to ubiquitin-activating enzyme E1. *Nature* **364**, 161-164.
- Li, W.-H., Gojobori, T. & Nei, M. (1981). Pseudogenes as a paradigm of neutral evolution. *Nature* **292**, 237-239.
- Liang, A. & Heckmann, K. (1993a). *Blepharisma* uses UAA as a termination codon. *Naturwissenschaften* **80**, 225-226.
- Liang, A. & Heckmann, K. (1993b). The macronuclear γ -tubulin-encoding gene of *Euplotes octocarinatus* contains two introns and an in-frame TGA. *Gene* **136**, 319-322.
- Linnen, J.M., Bailey, C.P. & Weeks, D.L. (1993). Two related localised mRNAs from *Xenopus laevis* encode ubiquitin-like fusion proteins. *Gene* **128**, 181-188.
- Loeb, K.R. & Haas, A.L. (1992). The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. *J. Biol. Chem.* **267**, 7806-7813.
- Lopez-Otin, C., Simon-Mateo, C., Martinez, L. & Viñuela, E. (1989). Gly-Gly-X, a novel consensus sequence for the proteolytic processing of viral and cellular proteins. *J. Biol. Chem.* **264**, 9107-9110.
- Low, T.L.K., Thurman, G.B., McAdoo, M., McClure, J., Rossio, J.L., Naylor, P.H. & Goldstein, A.L. (1979). The chemistry and biology of thymosin. *J. Biol. Chem.* **254**, 981-986.
- Lund, P.K., Moats-Staats, B.M., Simmons, J.G., Hoy, E., D'Ercole, A.J., Martin, F. & van Wyk, J.J. (1985). Nucleotide sequence analysis of a cDNA encoding human ubiquitin reveals that ubiquitin is synthesized as a precursor. *J. Biol. Chem.* **260**, 7609-7613.
- Luporini, P. & Miceli, C. (1986). Mating pheromones. Em: Gall, J.G. (ed). *The molecular biology of ciliated protozoa*, Academic Press, London, pp. 263-299.
- Lynn, D.H. & Corliss, J.O. (1991). *Microscopic Anatomy of Invertebrates: Protozoa*, Wiley-Liss, New York, Vol.1, pp 35-179.

- Lynn, D.H. & Sogin, M.L. (1988). Assessment of phylogenetic relationships among ciliated protists using partial ribosomal RNA sequences derived from reverse transcripts. *BioSystems* **21**, 249-254.
- Majetschak, M., Laub, M. & Jennissen, H.P. (1993). A ubiquitin-calmodulin synthetase that effectively recognizes the Ca^{2+} -free form of calmodulin. *FEBS Lett.* **315**, 347-352.
- Martindale, D.W., Allis, C.D. & Bruns, P.J. (1985). RNA and protein synthesis during meiotic prophase in *Tetrahymena thermophila*. *J. Protozool.* **32**, 644-649.
- Martindale, D.W. (1989). Codon usage in *Tetrahymena* and other ciliates. *J. Protozool.* **36**, 29-34.
- Martindale, D.W. (1990). A conjugation-specific gene (cnjC) from *Tetrahymena* encodes a protein homologous to yeast RNA polymerase subunits (RPB3, RPC40) and is similar to a portion of the prokaryotic RNA polymerase α subunit (rpoA). *Nucleic Acids Res.* **18**, 2953-2960.
- Martindale, D.W. & Taylor, F.M. (1988). Multiple introns in a conjugation-specific gene from *Tetrahymena thermophila*. *Nucleic Acids Res.* **16**, 2189-2201.
- Martindale, D.W., Allis, C.D. & Bruns, P.J. (1982). Conjugation in *Tetrahymena thermophila*: a temporal analysis of cytological stages. *Exp. Cell. Res.* **140**, 227-236.
- Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P.J., Bootsma, D., Hoeijmakers, J.H.J. & Hanaoka, F. (1994). Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO J.* **13**, 1831-1843.
- Mayer, A.N. & Wilkinson, K.D. (1989). Detection, resolution, and nomenclature of multiple carboxyl-terminal esterases from bovine calf thymus. *Biochemistry* **28**, 166-172.
- Mayer, R.J., Arnold, J., László, L., Landon, M. & Lowe, J. (1991). Ubiquitin in health and disease. *Biochem. Biophys. Acta* **1089**, 141-157.
- Mayo, K.M. & Orias, E. (1986). Developmental regulation of gene expression in *Tetrahymena*. *Dev. Biol.* **116**, 302-313.
- McGrath, J.P., Jentsch, S. & Varshavsky, A. (1991). UBA1: an essential yeast gene encoding ubiquitin-activating enzyme. *EMBO J.* **10**, 227-236.
- Mesco, E.R. & Timiras, P.S. (1991). Tau-ubiquitin protein conjugates in a human cell line. *Mechan. Ageing Develop.* **61**, 1-9.
- Messing, J. (1983). New M13 vectors for cloning. *Methods Enzymol.* **101**, 20-78.
- Meyer, E. (1992). Induction of specific macronuclear developmental mutations by microinjection of a cloned telomeric gene in *Paramecium primaurelia*. *Genes Dev.* **6**, 211-222.
- Meyer, F., Schmidt, H.J., Plümper, E., Hasilik, A., Mersmann, G., Meyer, H.E., Engström, Å. & Heckmann, K. (1991). UGA is translated as cysteine in pheromone 3 of *Euplotes octocarinatus*. *Proc. Natl. Acad. Sci. USA* **88**, 3758-3761.
-

- Meyers, G., Rümenapf, T. & Thiel, H.-J. (1989). Molecular cloning and nucleotide sequence of the genome of hog cholera virus. *Virology* **171**, 555-567.
- Meyers, G., Tautz, N., Dubovi, E.J. & Thiel, H.-J. (1991). Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. *Virology* **180**, 602-616.
- Mezquita, J., Lopez-Ibor, B., Pau, M. & Mezquita, C. (1993). Intron and intronless transcription of the chicken polyubiquitin gene UbII. *FEBS Lett.* **319**, 244-248.
- Mezquita, J., Pau, M. & Mezquita, C. (1988). cDNA encoding a chicken ubiquitin-fusion protein identical to the corresponding human protein. *Nucleic Acids Res.* **16**, 11838.
- Michiels, L., van der Rauwelaert, E., van Hasselt, F., Kas, K. & Merregaert, J. (1993). *fau* cDNA encodes a ubiquitin-like S30 fusion protein and is expressed as an antisense in the Finkel-Biskis-Reilly murine sarcoma virus. *Oncogene* **8**, 2537-2546.
- Mikami, K. & Hiwatashi, K. (1975). Macronuclear regeneration and cell division in *Paramecium caudatum*. *J. Protozool.* **22**, 536-540.
- Miller, H.I., Henzel, W.J., Ridgway, J.B., Kuang, W.-J., Chisholm, V., Liu, C.-C. (1989). Cloning and expression of a yeast ubiquitin-protein activity in *Escherichia coli*. *Biotechnol.* **7**, 698-704.
- Mita, S., Yasuda, H., Marunouchi, T., Ishiko, S. & Yamada, M. (1980). A temperature-sensitive mutant of cultured mouse cells defective in chromosome condensation. *Exp. Cell Res.* **126**, 407-416.
- Mitcham, J.L., Lynn, A.J. & Prescott, D.M. (1992). Analysis of a scrambled gene: the gene encoding α -telomere-binding protein in *Oxytricha nova*. *Genes Dev.* **6**, 788-800.
- Mitchell, M.J., Woods, D.R., Tucker, P.K., Opp, J.S. & Bishop, C.E. (1991). Homology of a candidate spermatogenic gene from the mouse Y chromosome to the ubiquitin-activating enzyme E1. *Nature* **354**, 483-489.
- Mitchell, M.J., Woods, D.R., Wilcox, S.A., Graves, J.A.M. & Bishop, C.E. (1992). Marsupial Y chromosome encodes a homologue of the mouse Y-linked candidate spermatogenesis gene *Ube1y*. *Nature* **359**, 528-531.
- Miyake, A. (1981). Physiology and biochemistry in ciliates. Em: Levandowsky, M. & Hutner, S.H. (eds) *Biochemistry and physiology of the protozoa*, 2^a ed., vol.4, Academic Press, New York, pp. 125-198.
- Monia, B.P., Ecker, D.J., Jonnalagadda, S., Marsh, J., Gotlib, L., Butt, T.R. & Crooke, S.T. (1989). Gene synthesis, expression, and processing of human ubiquitin carboxyl extension proteins. *J. Biol. Chem.* **264**, 4093-4103.
- Mori, S., Heldin, C.-H. & Claesson-Welsh, L. (1992). Ligand-induced polyubiquitin of the platelet-derived growth factor β -receptor. *J. Biol. Chem.* **267**, 6429-6434.
- Mori, S., Heldin, C.-H. & Claesson-Welsh, L. (1993). Ligand-induced polyubiquitin of the platelet-derived growth factor β -receptor plays a negative regulatory role in its mitogenic signalling. *J. Biol. Chem.* **268**, 577-583.
- Mori, H., Kondo, J. & Ihara, I. (1987). Ubiquitin is a component of paired helical filaments in Alzheimer's disease. *Science* **235**, 1641-1644.

- Morin, G. (1991). Recognition of a chromosome truncation site associated with α -thalassaemia by human telomerase. *Nature* **353**, 454-456.
- Morrison, A., Miller, E.J. & Prakash, L. (1988). Domain structure and functional analysis of the carboxy-terminal polyacidic sequence of the RAD6 protein of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **8**, 1179-1185.
- Mueller, R.D., Yasuda, H., Hatch, C.L., Bonner, W. & Bradbury, E.M. (1985). Identification of ubiquitinated histones 2A and 2B in *Physarum polycephalum*. *J. Biol. Chem.* **260**, 5147-5153.
- Müller-Taubenberger, A., Graack, H.-R., Grohmann, L., Schleicher, M. & Gerisch, G. (1989). An extended ubiquitin of *Dictyostelium* is located in the small ribosomal subunit. *J. Biol. Chem.* **264**, 5319-5322.
- Müller-Taubenberger, A., Hagmann, J., Noegel, A. & Gerisch, G. (1988a). Ubiquitin gene expression in *Dictyostelium* is induced by heat and cold shock, cadmium, and inhibitors of protein synthesis. *J. Cell Science* **90**, 51-58.
- Müller-Taubenberger, A., Westphal, M., Jaeger, E., Noegel, A. & Gerisch, G. (1988b). Complete cDNA sequence of a *Dictyostelium* ubiquitin with a carboxy-terminal tail and identification of the protein using an anti-peptide antibody. *FEBS Lett.* **229**, 273-278.
- Murray, A.W., Solomon, M.J. & Krishner, M.W. (1989). The role of cyclin synthesis and degradation in the control of maturation promoting factor activity. *Nature* **339**, 280-286.
- Murti, K.G., Smith, H.T. & Fried, V.A. (1988). Ubiquitin is a component of the microtubule network. *Proc. Natl. Acad. Sci. USA* **85**, 3019-3023.
- Mustaev, A., Kashlev, M., Lee, J., Polyakov, A., Lebedev, A., Zalenskaya, K., Grachev, M., Goldfarb, A. & Nikiforov, V. (1991). Mapping of the priming substrate contacts in the active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **266**, 23927-23931.
- Nanney, D.L. & McCoy, J.W. (1976). Characterization of the species of the *Tetrahymena pyriformis* complex. *Trans. Amer. Micros. Soc.* **95**, 664-682.
- Nanney, D.L. & Preparata, R.M. (1979). Genetic evidence concerning the structure of the *Tetrahymena thermophila* macronucleus. *J. Protozool.* **26**, 2-9.
- Nanney, D.L. (1953). Nucleo-cytoplasmic interaction during conjugation in *Tetrahymena*. *Biol. Bull.* **105**, 133-148.
- Nanney, D.L. (1956). Caryonidal inheritance and nuclear differentiation. *Amer. Natur.* **90**, 291-307.
- Nanney, D.L. (1964). Macronuclear differentiation and subnuclear assortment in ciliates. *Symp. Soc. Dev. Biol.* **23**, 253-273.
- Nasmyth, K.A. (1982). Molecular genetics of yeast mating type. *Ann. Rev. Genet.* **16**, 439-500.
- Nemer, M., Rondinelli, E., Infante, D. & Infante, A.A. (1991). Polyubiquitin RNA characteristics and conditional induction in sea urchin embryos. *Dev. Biol.* **145**, 255-265.
-

- Nenoi, M., Mita, K., Ichimura, S. & Cartwright, I.L. (1994). Novel structure of a Chinese hamster polyubiquitin gene. *Biochem. Biophys. Acta*. **1204**, 271-278.
- Nenoi, M. (1992). Induced accumulation of polyubiquitin gene transcripts in HeLa cells after UV-irradiation and TPA-treatment. *Int. J. Radia. Biol.* **61**, 205-211.
- Nenoi, M., Mita, K. & Ichimura, S. (1992). Evolutionary conserved structure of the 3' non-translated region of a Chinese hamster polyubiquitin gene. *Biochem. Biophys. Acta* **1130**, 247-252.
- Neves, A.M. (1991). *Estrutura e expressão da família multigénica ubiquitina em Tetrahymena pyriformis*. Tese de doutoramento, Faculdade de Ciências da Universidade de Lisboa, Lisboa.
- Neves, A., Guerreiro, P. & Rodrigues-Pousada, C. (1990). Striking changes in polyubiquitin genes of *Tetrahymena pyriformis*. *Nucleic Acids Res.* **18**, 656
- Neves, A.M., Barahona, I., Galego, L. & Rodrigues-Pousada, C. (1988). Ubiquitin genes in *Tetrahymena pyriformis* and their expression during heat shock. *Gene* **73**, 87-96.
- Neves, A.M., Guerreiro, P. & Rodrigues-Pousada, C. (1991a). The macronuclear polyubiquitin gene of the ciliate *Tetrahymena pyriformis*. *Dna Seq.-J. DNA Seq. Mapp.* **2**, 173-180.
- Neves, A.M., Guerreiro, P., Miquerol, L. & Rodrigues-Pousada, C. (1991b). Molecular cloning and expression of a *Tetrahymena pyriformis* ubiquitin fusion gene coding for a 53-amino-acid extension protein. *Mol. Gen. Genet.* **230**, 186-192.
- Nickel, B.E., Allis, C.D. & Davie, J.R. (1989). Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin. *Biochem.* **28**, 958-963.
- Niedzwiecki, A. & Fleming, J.E. (1993). Heat shock induces changes in the expression and binding of ubiquitin in senescent *Drosophila melanogaster*. *Dev. Genet.* **14**, 78-86.
- Nielsen, H., Andreassen, P.H., Dreisig, H., Kristiansen, K. & Engberg, J. (1986). An intron in a ribosomal protein gene from *Tetrahymena*. *EMBO J.* **5**, 2711-2717.
- Nishi, R., Hashimoto, H., Kidou, S.-I., Uchimiya, H. & Kato, A. (1993). Isolation and characterization of a rice cDNA which encodes a ubiquitin proteins and a 52 amino acid extension protein. *Plant Mol. Biol.* **22**, 159-161.
- Nock, A. (1981). RNA and macronuclear transcription in the ciliate *Stylonychia mytilus*. *Chromosoma* **83**, 209-220.
- Norris, S.R., Meyer, S.E. & Callis, J. (1993). The intron of *Arabidopsis thaliana* polyubiquitin genes is conserved in location and is a quantitative determinant of chimeric gene expression. *Plant Mol. Biol.* **21**, 895-906.
- Ochmani, T., Giordina, R., Shaw, D.R. & Ennis, H.L. (1989). Molecular organization of developmentally regulated *Dictyostelium discoideum* ubiquitin cDNAs. *Biochem.* **28**, 5226-5231.
- Orias, E. & Orias, J.D. (1983). Topologically invariant deformation of a cell-cell junction allows gamete pronucleus exchange in conjugating *Tetrahymena*. *J. Cell Biol.* **97**, 92a.

- Orias, E. & Hamilton, E.P. (1979). Cytogamy: an inducible alternate pathway of conjugation in *Tetrahymena thermophila*. *Genetics* **91**, 657-671.
- Orias, E. (1963). Mating type determination in variety 8 *Tetrahymena pyriformis*. *Genetics* **48**, 1509-1518.
- Orias, E. (1986). Ciliate conjugation. Em: Gall, J.G. (ed). *The molecular biology of ciliated protozoa*, Academic Press, London, pp. 45-84.
- Orias, E., Hamilton, E.P. & Flacks, M. (1979). Osmotic shock prevents nuclear exchange and produces whole-genome homozygotes in conjugating *Tetrahymena*. *Science* **203**, 660-663.
- Orias, J.D., Hamilton, E.P. & Orias, E. (1983). A microtubule meshwork associated with gametic pronucleus transfer across a cell-cell junction. *Science* **222**, 181-184.
- Osawa, S., Jukes, T.H., Watanabe, K. & Muto, A. (1992). Recent evidence for evolution of the genetic code. *Microbiol. Rev.* **56**, 229-264.
- Özkaynak, E., Finley, D., Solomon, M.J. & Varshavsky, A. (1987). The yeast ubiquitin genes: a family of natural gene fusions. *EMBO J.* **6**, 1429-1439.
- Özkaynak, E., Finley, D. & Varshavsky, A. (1984). The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein. *Nature* **312**, 663-666.
- Paolini, R. & Kinet, J.-P. (1993). Cell surface control of the multiubiquitination and deubiquitination of high-affinity immunoglobulin E receptors. *EMBO J.* **12**, 779-786.
- Parag, H.A., Dimitrovsky, D., Raboy, B. & Kulka, R.G. (1993). Selective ubiquitination of calmodulin by UBC4 and a putative ubiquitin protein ligase (E3) from *Saccharomyces cerevisiae*. *FEBS Lett.* **325**, 242-246.
- Patterson, T.A. & Dean, M. (1987). Preparation of high titer lambda phage lysates. *Nucleic Acids Res.* **15**, 6298.
- Pfeifer, K., Frank, W., Schröder, H.C., Gamulin, V., Rinkevich, B., Batel, R., Müller, I.M. & Müller, W.E.G. (1993). Cloning of the polyubiquitin cDNA from the marine sponge *Geodia cydonium* and its preferential expression during reaggregation of cells. *J. Cell Sci.* **106**, 545-554.
- Pickart, C.M. & Rose, I.A. (1985a). Functional heterogeneity of ubiquitin carrier proteins. *J. Biol. Chem.* **260**, 1573-1581.
- Pickart, C.M. & Rose, I.A. (1985b). Ubiquitin carboxyl-terminal hydrolase acts on ubiquitin carboxyl-terminal amides. *J. Biol. Chem.* **260**, 7903-7910.
- Pickart, C.M., Haldeman, M.T., Kasperek, E.M. & Chen, Z. (1992). Iodination of tyrosine 59 of ubiquitin selectively blocks ubiquitin's acceptor activity in diubiquitin synthesis catalyzed by E2_{25K}. *J. Biol. Chem.* **267**, 14418-14423.
- Picologlou, S., Brown, N. & Liebman, S.W. (1990). Mutations in *RAD6*, a yeast gene encoding a ubiquitin-conjugating enzyme, stimulate retrotransposition.
- Pieterse, C.M.J., Risseuw, E.P. & Davidse, L.C. (1991). An *in planta* induced gene of *Phytophthora infestans* codes for ubiquitin. *Plant Mol. Biol.* **17**, 799-811.
- Pollmann, L., Kampen, J. & Wettern, M. (1991). Ubiquitin in a lower plant. *Eur. J. Biochem.* **202**, 197-204.
-

- Prakash, S., Sung, P. & Prakash, L. (1990). Structure and function of *RAD3*, *RAD6* and other DNA repair genes of *Saccharomyces cerevisiae*. *The eukaryotic nucleus*, Strauss & Wilson ed., Telford Press Inc., Caldwell, N.J., p. 275-292.
- Preer Jr, J.R., & Preer, L.B. (1979). The size of macronuclear DNA and its relationship to models for maintaining genic balance. *J. Protozool.* **26**, 14-18.
- Preer Jr, J.R., Preer, L.B., Rudman, B.M. & Barnett, A.J. (1985). Deviation from the universal code shown by the gene for surface protein 51A in *Paramecium*. *Nature* **314**, 188-190.
- Preparata, R.M. & Nanney, D.L. (1977). Cytogenetics of triplet conjugation in *Tetrahymena*: origin of haploid and triploid clones. *Chromosoma* **60**, 49-57.
- Price, C.M. & Cech, T.R. (1987). Telomeric DNA-protein interactions of *Oxytricha* macronuclear DNA. *Genes Dev.* **1**, 783-793.
- Price, C.M. & Cech, T.R. (1989). Properties of the telomeric DNA-binding protein from *Oxytricha nova*. *Biochem.* **28**, 769-774.
- Price, C.M. (1990). Telomere structure in *Euplotes crassus*: characterization of DNA-protein interactions and isolation of a telomere-binding protein. *Mol. Cell. Biol.* **10**, 3421-3431.
- Price, C.M., Skopp, R., Krueger, J. & Williams, D. (1992). DNA recognition and binding by the *Euplotes* telomere protein. *Biochem.* **31**, 10835-10843.
- Putterman, D., Pepinsky, R.B. & Vogt, V.M. (1990). Ubiquitin in avian leukosis virus particles. *Virology* **176**, 633-637.
- Quail, P.H. (1991). Phytochrome: a light-activated molecular switch that regulates plant expression. *Ann. Rev. Genet.* **25**, 389-409.
- Raikov, I.B. (1982). Em: *The protozoan nucleus: morphology and evolution*, Springer-Verlag, New York.
- Ray Jr, C. (1956). Meiosis and nuclear behaviour in *Tetrahymena pyriformis*. *J. Protozool.* **3**, 88-96.
- Rechsteiner, M., (1987). Ubiquitin-mediated pathways for intracellular proteolysis. *Ann. Rev. Cell Biol.* **3**, 1-30.
- Rechsteiner, M. (ed). (1988). *Ubiquitin*. New York. Plenum Press.
- Redman, K.L. & Rechsteiner, M. (1988). Extended reading frame of a ubiquitin gene encodes a stable, conserved, basic protein. *J. Biol. Chem.* **263**, 4926-4931.
- Redman, K.L. & Rechsteiner, M. (1989). Identification of the long ubiquitin extension protein as ribosomal protein S27a. *Nature* **338**, 438-440.
- Reich, N., Evans, B., Levy, D., Fahey, D., Knight Jr., E. & Darnell Jr., J.E. (1987). Interferon-induced transcription of a gene encoding a 15-kDa protein depends on an upstream enhancer element. *Proc. Natl. Acad. Sci. USA* **84**, 6394-6398.
- Reiss, Y. & Hershko, A. (1990). Affinity purification of ubiquitin-protein ligase. *J. Biol. Chem.* **265**, 3685-3690.
- Reiss, Y., Heller, H. & Hershko, A. (1989). Binding sites of ubiquitin-protein ligase. *J. Biol. Chem.* **264**, 10378-10383.
-

- Reiss, Y., Kaim, D. & Hershko, A. (1988). Specificity of binding of NH₂-terminal residue of proteins to ubiquitin-protein ligase. Use of amino acid derivatives to characterize specific binding sites. *J. Biol. Chem.* **263**, 2693-2698.
- Reynolds, G.J. & Hooley, R. (1992). cDNA cloning of a tetraubiquitin gene, and expression of ubiquitin-containing transcripts, in aleurone layers of *Avena fatua*. *Plant Mol. Biol.* **20**, 753-758.
- Reynolds, P., Weber, S. & Prakash, L. (1985). *RAD6* gene of *Saccharomyces cerevisiae* encodes a protein containing a tract of 13 consecutive aspartates. *Proc. Natl. Acad. Sci. USA.* **82**, 168-172.
- Ribas-Aparicio, R.M., Sparkowski, J.J., Proulx, A.E., Mitchell, J.D. & Klobutcher, L.A. (1987). Nucleic acid splicing events occur frequently during macronuclear development in the protozoan *Oxytricha nova* and involve the elimination of unique DNA. *Genes Dev.* **1**, 323-336.
- Richardson, H.E., Wittenberg, C., Cross, F., Reed, S.I. (1989). An essential G₁ function for cyclin-like proteins in yeast. *Cell* **59**, 1127-1133.
- Rivett, A.J. & Sweeney, S.T. (1991). Properties of subunits of the multicatalytic proteinase complex revealed by the use of subunit-specific antibodies. *Biochem. J.* **278**, 171-177.
- Rivett, A.J. (1993). Proteasomes: multicatalytic proteinase complexes. *Biochem. J.* **291**, 1-10.
- Roberts Jr, C.T., & Orias, E. (1973). Cytoplasmic inheritance of chloramphenicol resistance in *Tetrahymena*. *Genetics* **73**, 259-272.
- Rodrigues-Pousada, C., Cyrne, M.L. & Hayes, D. (1979). Characterization of preribosomal ribonucleoprotein particles from *Tetrahymena pyriformis*. *Eur. J. Biochem.* **102**, 389-397.
- Rodrigues-Pousada, C., Marcaud, L., Portier, M.-M. & Hayes, D.H. (1975). Rapidly labelled RNA in *Tetrahymena pyriformis*. *Eur. J. Biochem.* **56**, 117-122.
- Rodriguez, J.M., Salas, M.L. & Viñuela, E. (1992). Genes homologous to ubiquitin-conjugating protein and eukaryotic transcription factor SII in African Swine Fever Virus. *Virol.* **186**, 40-52.
- Rogers, S., Wells, R. & Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-368.
- Rogers, M.B. & Karrer, K.M. (1985). Adolescence in *Tetrahymena thermophila*. *Proc. Natl. Acad. Sci. USA* **82**, 436-439.
- Romero, D.P. & Blackburn, E.H. (1992). A conserved secondary structure for telomerase RNA. *Cell* **67**, 343-353.
- Rose, I.A. & Warms, J.V.B. (1983). An enzyme with ubiquitin carboxyl-terminal esterase activity from reticulocytes. *Biochemistry* **22**, 4234-4237.
- Rose, I.A. (1988). Ubiquitin carboxyl-terminal hydrolases. Em: *Ubiquitin*, Rechsteiner, M. (ed). pp. 135-55, ed. 1988, New York, Plenum Press.
-

- Russell, R.L.Q. & Rohrmann, G.F. (1993). Nucleotide sequence of the ubiquitin-39K gene region from the *Orgyia pseudosugata* multinucleocapsid nuclear polyhedrosis virus genome. *J. Gen. Virol.* **74**, 1191-1195.
- Sadler, L.A. & Brunk, C.F. (1992). Phylogenetic relationships and unusual diversity in Histone H4 proteins within the *Tetrahymena pyriformis* complex. *Mol. Biol. Evol.* **9**, 70-84.
- Sambrook, J., Fritsh, E.F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. 2nd Ed. Cold Spring Harbor University Press, Cold Spring Harbor.
- Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci.* **74**, 5463-5467.
- Scheffner, M., Werness, B.A., Huibregtse, J.M., Levine, A. & Howley, P. (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**, 1129-1136.
- Schiedlmeier, B. & Schmitt, R. (1994). Repetitious structure and transcription control of a polyubiquitin gene in *Volvox carteri*. *Curr. Genet.* **25**, 169-177.
- Schlesinger, M. & Hershko, A., eds. (1988). *The Ubiquitin System, Current Communications in Molecular Biology*. Cold Spring Harbor. Cold Spring Harbor Lab.
- Schlesinger, D.H., Goldstein, G. & Niall, H.D. (1975). The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochem.* **14**, 2214-2218.
- Schneider, S.U., Leible, M.B. & Yang, X.-P. (1989). Strong homology between the small subunit of ribulose-1,5-biphosphate carboxylase/oxygenase of two species of *Acetabularia* and the occurrence of unusual codon usage. *Mol. Gen. Genet.* **218**, 445-452.
- Schwartz, L.M., Kocs, L. & Kay, B.K. (1990). Gene activation is required for developmentally programmed cell death. *Proc. Natl. Acad. Sci. USA* **87**, 6594-6598.
- Scott, J.M., Mikami, K., Leeck, C.L. & Forney, J.D. (1994). Non-Mendelian inheritance of macronuclear mutations is gene specific in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **14**, 2479-2484.
- Seelig, A., Kloetzel, P.M., Kuehn, L. & Dahlmann, B. (1991). Molecular interaction of the proteasome (multicatalytic proteinase): evidence that proteasome is not a constituent of the '26S' multienzyme complex. *Biochem. J.* **280**, 225-232.
- Seufert, W. & Jentsch, S. (1990). Ubiquitin-conjugating enzymes UBC4 and UBC5 mediate selective degradation of short-lived and abnormal proteins. *EMBO J.* **9**, 543-550.
- Seufert, W. & Jentsch, S. (1992). *In vivo* function of the proteasome in the ubiquitin pathway. *EMBO J.* **11**, 3077-3080.
- Seufert, W., McGrath, J.P. & Jentsch, S. (1990). UBC1 encodes a novel member of yeast ubiquitin-conjugating enzymes involved in protein degradation. *EMBO J.* **9**, 4535-4541.

- Seyfert, H.M. & Preparata, R.M. (1979). The regulation of amounts and proportions of genetic elements in the macronuclei of *Tetrahymena thermophila* strains of diverse karyotype. *J. Cell. Sci.* **40**, 111-123.
- Shanklin, J., Jabben, M. & Vierstra, R.D. (1987). Red light-induced formation of ubiquitin-phytochrome conjugates: identification of possible intermediates of phytochrome degradation. *Proc. Natl. Acad. Sci. USA* **84**, 359-363.
- Sharp, P.M. & Li, W.-H. (1987a). Molecular evolution of ubiquitin genes. *Trends Ecol. Evol.* **2**, 328-332.
- Sharp, P.M. & Li, W.-H. (1987b). The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281-1295.
- Sherman, F., Stewart, J.M. & Tsunasawa, S. (1985). Methionine or not methionine at the beginning of a protein? *BioEssays* **3**, 27-31.
- Shimogawara, K. & Muto, S. (1992). Purification of *Chlamydomonas* 28-kDa ubiquitinated protein and its identification as ubiquitinated histone H2B. *Arch. Biochem. Biophys.* **294**, 193-199.
- Shippen-Lentz, D. & Blackburn, E.H. (1990). Functional evidence for an RNA template in telomerase. *Science* **247**, 546-552.
- Siegelman, M., Bond, M.W., Gallatin, W.M., St. John, T., Smith, H.T., Fried, V.A. & Weissman, I.L. (1986). Cell surface molecule associated with lymphocyte homing is a ubiquitinated branched-chain glycoprotein. *Science* **231**, 823-829.
- Silver, E.T., Gwozd, T.J., Ptak, C., Goebel, M. & Ellison, M.J. (1992). A chimeric ubiquitin conjugating enzyme that combines the cell cycle properties of CDC34 (UBC3) and the DNA repair properties of RAD6 (UBC2): implications for the structure, function and evolution of the E2s. *EMBO J.* **11**, 3091-3098.
- Simon, E.M. & Orias, E. (1987). Genetic instability in the mating type system of *Tetrahymena pigmentosa*. *Genetics* **117**, 437-449.
- Soares, H., Cyrne, L., Barahona, I. & Rodrigues-Pousada, C. (1991). Different patterns of expression of β -tubulin genes in *Tetrahymena pyriformis* during reciliation. *Eur. J. Biochem.* **197**, 291-299.
- Soares, H., Galego, L., C3ias, R. & Rodrigues-Pousada, C. (1993). The mechanisms of tubulin messenger regulation during *Tetrahymena pyriformis* reciliation. *J. Biol. Chem.* **268**, 16623-16630.
- Soares, H., Penque, D., Mouta, C. & Rodrigues-Pousada, C. (1994). A *Tetrahymena* orthologue of the mouse chaperonin subunit CCT γ and its coexpression with tubulin during cilia recovery. *J. Biol. Chem.* **269**, 29299-29307.
- Sommer, T. & Jentsch, S. (1993). A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature* **365**, 176-179.
- Sommer, T. & Seufert, W. (1992). Genetic analysis of ubiquitin-dependent protein degradation. *Experientia* **48**, 172-178.
- Sonneborn, T.M. (1974). *Paramecium aurelia*. In *Handbook of genetics: Plants, plants viruses and protists*, ed R.C. King, vol.2, pp. 469-594, Plenum Press, New York.
-

- Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517.
- St. John, T., Gallatin, M., Siegelman, M., Smith, H.T., Fried, V.A. & Weissman, I.L. (1986). Expression cloning of a lymphocyte homing receptor cDNA: ubiquitin is the reactive species. *Science* **231**, 845-850.
- Stargell, L.A., Karrer, K.M. & Gorovsky, M.A. (1990). Transcriptional regulation of gene expression in *Tetrahymena thermophila*. *Nucleic Acids Res.* **18**, 6637-6639.
- Stargell, L.A. & Gorovsky, M.A. (1994). TATA-binding protein and nuclear differentiation in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **14**, 723-734.
- Steele, C.J., Barkocy-Gallagher, G.A., Preer, L.B. & Preer Jr, J.R., (1994). Developmentally excised sequences in micronuclear DNA in *Paramecium*. *Proc. Natl. Acad. Sci. USA* **91**, 2255-2259.
- Steinbruck, G., Haas, I., Hellmer, K. & Ammerman, D. (1981). Characterization of macronuclear DNA in five species of ciliates. *Chromosoma* **83**, 199-208.
- Stragier, P., Kunkel, B., Kroos, L. & Losick, R. (1989). Chromosomal rearrangement generating a composite gene for a development transcription factor. *Science* **243**, 507-512.
- Sullivan, M.L. & Vierstra, R.D. (1991). Cloning of a 16-kDa ubiquitin carrier protein from wheat and *Arabidopsis thaliana*. *J. Biol. Chem.* **266**, 23878-23885.
- Sun, C.-W. & Callis, J. (1993). Recent stable insertion of mitochondrial DNA into an *Arabidopsis* polyubiquitin gene by nonhomologous recombination. *Plant Cell* **5**, 97-107.
- Sung, P., Berleth, E., Pickart, C., Prakash, S. & Prakash, L. (1991). Yeast RAD6 encoded ubiquitin conjugating enzyme mediates protein degradation dependent on the N-end-recognizing E3 enzyme. *EMBO J.* **10**, 2187-2193.
- Sung, P., Prakash, S. & Prakash, L. (1988). The RAD6 protein of *Saccharomyces cerevisiae* poly-ubiquitinates histones, and its acidic domain mediates this activity. *Genes Dev.* **2**, 1476-1485.
- Sung, P., Prakash, S. & Prakash, L. (1990). Mutation of cysteine-88 in the *Saccharomyces cerevisiae* RAD6 protein abolishes its ubiquitin-conjugating activity and its various biological functions. *Proc. Natl. Acad. Sci.* **87**, 2695-2699.
- Sweetser, D., Nonet, M. & Young, R.A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proc. Natl. Acad. Sci. USA* **84**, 1192-1196.
- Swerdlow, P.S., Schuster, T. & Finley, D. (1990).. A conserved sequence in histone H2A which is a ubiquitination site in higher eukaryotes is not required for growth in *S.cerevisiae*. *Mol. Cell. Biol.* **10**, 4905-4911.
- Swindle, J., Ajioka, J., Eisen, H., Sanwal, B., Jacquemot, C., Browder, Z. & Buck, G. (1988). The genomic organization and transcription of the ubiquitin genes of *Trypanosoma cruzi*. *EMBO J.* **7**, 1121-1127.

- Taccioli, G.E., Grotewold, E., Aisemberg, G.O. & Judewicz, N.D. (1989). Ubiquitin expression in *Neurospora crassa*: cloning and sequencing of a polyubiquitin gene. *Nucleic Acids Res.* **17**, 6153-6165.
- Taccioli, G.E., Grotewold, E., Aisemberg, G.O. & Judewicz, N.D. (1991). The cDNA sequence and expression of an ubiquitin-tail gene fusion in *Neurospora crassa*. *Gene* **102**, 133-137.
- Tausta, S.L. & Klobutcher, L.A. (1989). Detection of circular forms of eliminated DNA during macronuclear development in *E.crassus*. *Cell* **59**, 1019-1026.
- Tausta, S.L. & Klobutcher, L.A. (1990). Internal eliminated sequences are removed prior to chromosome fragmentation during development in *Euplotes crassus*. *Nucleic Acids Res.* **18**, 845-853.
- Tautz, N., Meyers, G. & Thiel, H.-J. (1993). Processing of polyubiquitin in the polyprotein of an RNA virus. *Virology* **197**, 74-85.
- Thorne, A.W., Sautiere, P., Briand, G. & Robinson, C.C. (1987). The structure of ubiquitinated histone H2B. *EMBO J.* **6**, 1005-1010.
- Tobias, J.M. & Varshavsky, A. (1991). Cloning and functional analysis of the ubiquitin-specific protease gene *UBP1* of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **266**, 12021-12028.
- Tobias, J.M., Shrader, T.E., Rocap, G. & Varshavsky, A. (1991). The N-end rule in bacteria. *Science* **254**, 1374-1377.
- Toniolo, D., Persico, M. & Alcalay, M. (1988). A "housekeeping" gene on the X chromosome encodes a protein similar to ubiquitin. *Proc. Natl. Acad. Sci. USA* **85**, 851-855.
- Trausch, J.S., Grenfell, S.J., Handley-Gearhart, P.M., Ciechanover, A. & Schwartz, A.L. (1993). Immunofluorescent localization of the ubiquitin-activating enzyme, E1, to the nucleus and cytoskeleton. *Am. J. Physiol* **264** (*Cell Physiol.* **33**), C93-C102.
- Treich, I., Riva, M. & Sentenac, A. (1991). Zinc-binding subunits of yeast RNA polymerase. *J. Biol. Chem.* **266**, 21971-21976.
- Varshavsky, A. (1992). The N-end rule. *Cell* **69**, 725-735.
- Vassal, A., Boulet, A., Decoster, E. & Faye, G. (1992). QRI8, a novel ubiquitin-conjugating enzyme in *Saccharomyces cerevisiae*. *Biochem. Biophys. Acta* **1132**, 211-213.
- Vijay-Kumar, S., Bugg, C.E. & Cook, W.J. (1987a). Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **194**, 531-544.
- Vijay-Kumar, S., Bugg, C.E., Vierstra, R.D., Hatfield, P.M. & Cook, W.J. (1987b). Comparison of the three-dimensional structures of human, yeast, and oat ubiquitin. *J. Biol. Chem.* **262**, 6396-6399.
- Vijay-Kumar, S., Bugg, C.E., Wilkinson, K.D. & Cook, W.J. (1985). Three-dimensional structure of ubiquitin at 2.8 Å resolution. *Proc. Natl. Acad. Sci. USA* **82**, 3582-3585.
- Vogelstein, B & Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci.* **76**, 615-619.
-

- Wang, Z., Tarawneh, K.A. & Free, S.J. (1993). Isolation, sequencing and characterization of a gene encoding a *Neurospora* ribosomal protein. *Curr. Genet.* **13**, 330-333.
- Warner, J.R. (1989). A marriage of convenience or necessity? *Nature* **338**, 379.
- Watkins, J.F., Sung, P., Prakash, L. & Prakash, S. (1993). The *Saccharomyces cerevisiae* DNA repair gene *RAD23* encodes a nuclear protein containing a ubiquitin-like domain required for biological function. *Mol. Cell. Biol.* **13**, 7757-7765.
- Weber, P.L., Brown, S.C. & Mueller, L. (1987). Sequential ^1H NMR assignments and secondary structure identification of human ubiquitin. *Biochem.* **26**, 7282-7290.
- Weindruch, R.H. & Doerder, F.P. (1975). Age-dependent micronuclear deterioration in *Tetrahymena pyriformis*, syngen 1. *Mech. Ageing Dev.* **4**, 263-279.
- Wells, D., Bains, W. & Kedes, L. (1986). Codon usage in histone gene families of higher eukaryotes reflects functional rather than phylogenetic relationships. *J. Mol. Evol.* **23**, 224-241.
- Westphal, M., Müller-Taubenberger, A., Noegel, A. & Gerisch, G. (1986). Transcript regulation and carboxyterminal extension of ubiquitin in *Dictyostelium discoideum*. *FEBS Lett.* **209**, 92-96.
- White, T.C., El-Gewely, M.R. & Allen, S.L. (1985). Eliminated sequences with different copy numbers clustered in the micronuclear genome of *Tetrahymena thermophila*. *Mol. Gen. Genet.* **201**, 65-75.
- Wiborg, O., Pederson, M.S., Wind, A., Berglund, L.E., Marcker, K.A. & Vuust, J. (1985). The human ubiquitin multigene family: some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J.* **4**, 755-759.
- Wilkinson, K.D. & Mayer, A.N. (1986). Alcohol-induced conformational changes of ubiquitin. *Arch. Biochem. Biophys.* **250**, 390-399.
- Wilkinson, K.D., Lee, K.M., Deshpande, S., Duerksen-Hughes, P., Boss, J.M. & Pohl, J. (1989). The neuron-specific protein PGP 9.5 is a ubiquitin carboxyl-terminal hydrolase. *Science*, **246**, 670-673.
- Wilkinson, K.D., Urban, M.K. & Haas, A.L. (1980). Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J. Biol. Chem.* **255**, 7529-7532.
- Witt, I., Struub, N., Kaufer, N.F. & Gross, T. (1993). The CAGTCACA box in the fission yeast *Schizosaccharomyces pombe* functions like a TATA element and binds a novel factor. *EMBO J.* **12**, 1201-1208.
- Wolf, S., Lottspeich, F. & Baumeister, W. (1993). Ubiquitin found in the archeobacterium *Thermoplasma acidophilum*. *FEBS Lett.* **326**, 42-44.
- Wolfe, J. & Grimes, G.W. (1979). Tip transformation in *Tetrahymena*: a morphogenic response to interactions between mating types. *J. Protozool.* **26**, 82-89.
- Wolfe, J. (1982). The conjugation junction of *Tetrahymena*: its structure and development. *J. Morphol.* **172**, 159-178.
- Wolfe, J., Mpoke, S. & Tirone, S.F. (1993). Cilia, ciliary concanavalin A-binding proteins, and mating recognition in *Tetrahymena thermophila*. *Exp. Cell. Res.* **209**, 342-349.

- Wong, S. & Campbell, D.A. (1989). A polyubiquitin gene from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **37**, 147-150.
- Wong, S., Elgort, M.G., Gottesdiener, K. & Campbell, D.A. (1992). Allelic polymorphism of the *Trypanosoma brucei* polyubiquitin gene. *Mol. Biochem. Parasitol.* **55**, 187-196.
- Wong, S., Morales, T.H., Neigel, J.E. & Campbell, D.A. (1993). Genomic and transcriptional linkage of the genes for calmodulin, EF-Hand 5 protein, and ubiquitin extension protein 52. *Mol. Cell. Biol.* **13**, 207-216.
- Wöstmann, C., Tannich, E. & Bakker-Grunwald, T. (1992). Ubiquitin of *Entamoeba histolytica* deviates in six amino acid residues from the consensus of all other known ubiquitins. *FEBS Lett.* **308**, 54-58.
- Wray, C.G. & DeSalle, R. (1994). Phylogenetic utility of ubiquitin DNA sequence from 3 marine protist lineages. *Mol. Marine Biol. Biotechnol.* **3**, 13-22.
- Wu, M., Allis, C.D. & Gorovsky, M.A. (1988). Cell-cycle regulation as a mechanism for targeting proteins to specific DNA sequences in *Tetrahymena thermophila*. *Proc. Natl. Acad. Sci. USA* **85**, 2205-2209.
- Xu, M.-Q., Kathe, S.D., Goodrich-Blair, H., Nierzwicki-Bauer, S.A. & Shub, D.A. (1990). Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacterium. *Science* **250**, 1566-1570.
- Yamao, F., Muto, A., Kawauchi, M., Iwami, M., Iwagami, S., Azumi, Y. & Osawa, S. (1985). UGA is read as tryptophan in *Mycoplasma capricolum*. *Proc. Natl. Acad. Sci. USA* **82**, 2306-2309.
- Yanish-Perron, C., Vieira, J., Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.
- Yao, M.-C. & Gorovsky, M.A. (1974). Comparison of the sequences of macronuclear and micronuclear DNA of *Tetrahymena pyriformis*. *Chromosoma* **48**, 1-18.
- Yao, M.-C. & Yao, C.-H. (1981). The repeated hexanucleotide C-C-C-C-A-A is presented near free ends of macronuclear DNA of *Tetrahymena*. *Proc. Natl. Acad. Sci. USA* **78**, 7436-7439.
- Yao, M.-C. (1982). Elimination of specific DNA sequences from the somatic nucleus of the ciliate *Tetrahymena*. *J. Cell. Biol.* **92**, 783-789.
- Yao, M.-C. (1986). Amplification of ribosomal RNA genes. Em: Gall, J.G. (ed). *The molecular biology of ciliated protozoa*, Academic Press, London, pp. 179-201.
- Yao, M.-C., Choi, J., Yokoyama, S., Austerberry, C.F. & Yao, C.-H. (1984). DNA elimination in *Tetrahymena*: a developmental process involving extensive breakage and rejoining of DNA at defined sites. *Cell* **36**, 433-440.
- Yao, M.-C., Kimmel, A.R. & Gorovsky, M.A. (1974). A small number of cistrons for ribosomal RNA in the germinal nucleus of a eukaryote, *Tetrahymena pyriformis*. *Proc. Natl. Acad. Sci. USA* **71**, 3082-3086.
- Yao, M.-C., Yao, C.-H. & Monks, B. (1990). The controlling sequence for site-specific chromosome breakage in *Tetrahymena*. *Cell* **63**, 763-772.
-

- Yao, M.-C., Zheng, K. & Yao, C.-H. (1987). A conserved nucleotide sequence at the sites of developmentally regulated chromosomal breakage in *Tetrahymena*. *Cell* **48**, 779-788.
- Yarden, Y., Escobedo, J.A., Kuang, W.-J., Yang-Feng, T.L., Daniel, T.O., Tremble, P.M., Chen, E.Y., Ando, M.E., Harkins, R.N., Francke, U., Fried, V.A., Ullrich, A. & Williams, W.T. (1986). Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors. *Nature* **323**, 226-232.
- Yochem, J. & Byers, B. (1987). Structural comparison of the yeast cell division cycle gene *CDC4* and a related pseudogene. *J. Mol. Biol.* **195**, 233-245.
- Yokoyama, R.W. & Yao, M.-C. (1982). Elimination of DNA sequences during macronuclear differentiation in *Tetrahymena thermophila*, as detected by *in situ* hybridization. *Chromosoma* **85**, 11-22.
- Yu, G.-L. & Blackburn, E.H. (1991). Developmentally programmed healing of chromosomes by telomerase in *Tetrahymena*. *Cell* **67**, 823-832.
- Zakian, V.A., Runge, K. & Wang, S.-S. (1990). How does the end begin? *Trends Genet.* **6**, 12-16.
- Zeff, R.A. & Geliebter, J. (1987). Oligonucleotide probes for genomic DNA blots. *Focus* **9**, 1-2.
- Zhen, M., Heinlein, R., Jones, D., Jentsch, S. & Candido, E.P.M. (1993). The *ubc-2* gene of *Caenorhabditis elegans* encodes a ubiquitin-conjugating enzyme involved in selective protein degradation. *Mol. Cell. Biol.* **13**, 1371-1377.
- Zhuang, Z., Marks, B. & McCauley, R.B. (1992). The insertion of monoamine oxidase A into the outer membrane of rat liver mitochondria. *J. Biol. Chem.* **267**, 591-596.
- Zhuang, Z. & McCauley, R. (1989). Ubiquitin is involved in the *in vitro* insertion of monoamine oxidase B into mitochondrial outer membranes. *J. Biol. Chem.* **264**, 1485-1496.
- Ziegenhagen, R. & Jennissen, H.P. (1990). Plant and fungus calmodulins are polyubiquitinated at a single site in a Ca^{2+} -dependent manner. *FEBS Lett.* **273**, 253-256.

APÊNDICE 1

**Clusters of 5S rRNA in the intergenic region
of ubiquitin genes in *Tetrahymena pyriformis***

Clusters of 5S rRNAs in the intergenic region of ubiquitin genes in *Tetrahymena pyriformis*

Paulo Guerreiro, Ana Neves and Claudina Rodrigues-Pousada *

Laboratório de Genética Molecular, Instituto Gulbenkian de Ciência, Apartado 14, 2781 Oeiras Codex (Portugal)

(Received 20 April 1993)

(Revised manuscript received 7 July 1993)

Key words: Genome; Ubiquitin; Intergenic region; Ribosomal gene, 5S; rRNA cluster; (*T. pyriformis*)

Here, we report the molecular analysis of two independent 5S rRNA clusters found in the intergenic region of two ubiquitin genomic clones isolated from *Tetrahymena pyriformis*. Each cluster contains two 120-bp-long coding regions organized in tandem with 142/145-bp-long spacers.

The ciliate protozoa provide several interesting challenges and they present very special biological features. The most striking characteristics of ciliates are their nuclear dimorphism [1] and the special genetic code used [2–6]. Comparative analysis of the primary and secondary structures of ribosomal RNAs is a useful tool for inferring phylogenetic relationships. The 5S ribosomal RNA is a universal component of the translation machinery and is an extremely conserved molecule. This fact was confirmed by sequence analysis over 350 species with high phylogenetic divergence. However, although several 5S rRNA sequences have been determined in ciliates, few are known about their respective genes [7,8]. The only sequence reported was that of two partial 5S ribosomal genes in *Tetrahymena thermophila*. This ciliate seems to have about 30 clusters of 5S rRNA genes per haploid genome, each one with one to about 16 genes arranged in tandem, in either its macro- or micronuclei [9].

While molecular studies for *T. pyriformis* ubiquitin genes were performed [4–6], four 5S ribosomal genes were found in two ubiquitin genomic clones. Several ubiquitin genomic clones were isolated from a genomic library containing *Hind*III macronuclear DNA fragments inserted into pBR322 vector [10]. The clones TU2 [4] and TU1x (a 2.2-kb-long *Xba*I fragment of TU1 6.8-kb-long *Hind*III fragment) (unpublished) present several ubiquitin sequences. Interestingly, four 5S ribosomal genes, arranged in two clusters of two genes

each, were found in the intergenic regions of these genes (Fig. 1).

The nucleotide sequences presented in Fig. 2 correspond to the alignment of the two intergenic regions taking as reference the 5S ribosomal genes. Within each cluster we found two 120-bp-long sequences repeated in a head-to-tail tandem array, one with a spacer sequence of 142 and the other with 145 bp long. The two TU1x 5S ribosomal genes are identical, but two nucleotide differences were found between the two TU2 5S genes. Also, TU1x and TU2 5S genes are not identical, containing two to four nucleotide changes (see arrows Fig. 2).

Comparative analysis with the data reported by Preparata et al. [7] for 18 Tetrahymenine species shows that the 5S rRNAs sequences fall into six classes defined by base variation in five different positions (24, 19:59, 115, 117). According to this classification the four 5S genes of *T. pyriformis* basically fall into the

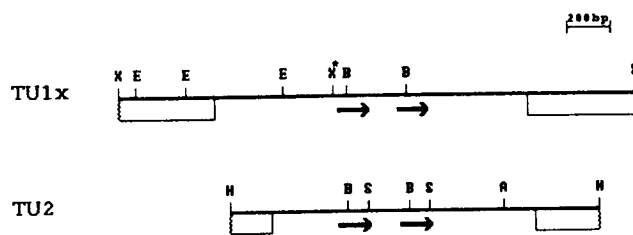


Fig. 1. Schematic representation of TU1x and TU2 ubiquitin clones. The ubiquitin clone Tu1x is a 2.2-kb-long *Xba*I fragment derived from the 6.8-kb-long *Hind*III fragment of TU1. Boxes indicated the polyubiquitin genes and the arrows are the 5S ribosomal genes. Restriction enzymes are: A, *Alu*I; B, *Bam*HI; E, *Eco*RI; H, *Hind*III; S, *Sal*I; X, *Xba*I; X*, *Xba*I (*dam* methylated).

* Corresponding author. Fax: +351 1 4431631.

TU1x	ATATAGAAATTAA	13
TU1x	TATTAAGATATTAATAAATAACATTAAAAATGATTTAATCTGTATTTACATAAATTTT	73
TU1x	TCTCATACAATTAAATGTGATTTATCATCTGACTATTAAAGTTATTTTGGCAATATGAT	133
TU1x	ACCCATAGATGATTTAATTTCTGAAATAACCTTTCTCTTATTTCTCGGATAGTTACCTACA	193
TU1x	AATAATTTATGGATGAATATCTTTGTTACAAAGTAAGATTTTGATTGATCATTTTAAAAA	253
TU2	AATAAAA	7
TU1x	ATTTAATCTAAAAATAACTGAAAAACAATCCACTCTATAATATACAAATAATGAATTCA	313
TU2	ATATTAGAAAAATATAAAAACTCTATTCAATTTAATACTTATAAATACTTCAATCTAT	67
TU1x	ATCTTTATCATTTTGTGTAATTAATAATCTTGATATCCAAAAATAGTGAAGGAAAACTTT	373
TU2	ATTTTATAAATCATTTGATCTACTCATAAAAATGATATAAAAAATTTCAATTTATCTGAAT	127
TU1x	ACTATCTTTAAATTTTACAAAGCAATTTATGATTCAGGTTATTTTAAATTTTATTTTC	433
TU2	TTTCATAATGATATTTTCATAAAATCTCCAAATAAAAAATCATAATTTACAAACAAATAAT	187
TU1x	AGTAATTTAACAATTAAGTTTATTATTATTTATTTGTTGTGTAATTTTAAAGTTTGAAGA	493
TU2	AATATTTTAAAAAATTAATTCAGATTTTAAACAAATTAAGTCTATTTATGAGTTTTCAGTTA	247
TU1x	AGAGTGAATCTTAAATTTATTTTGTATGATCTAGATTTTGTCTCATTACAAATAGATTTT	553
TU2	TTTTTAAAGTAAGGAATGGTTTGTGTAATTTTAAATACATGTTTCAATTTACTATATTTT	307
TU1x	GTTATCGGCCCATCTAAGGTGAAAAACACCGGATCCCATTCGAACTCCGAAGTTAAGCGCC	613
TU2	GTTGTCGGCCATCTAAGGTGAAAAACACCGGATCCCATTCGAACTCCGAAGTTAAGCGCC	367
TU1x	TTAAGGCTGGGTTAGTACTAAGGTGGGGGACCGCTTGGGAAGTCCCAAGTGCATTAACCT	673
TU2	TTAAGGCTGGGTTAGTACTAAGGTGGGGGACCGCTTGGGAAGTCCCAAGTGCATTAACCT	427
TU1x	TTTCTTTTTTTTATTCATATAAATCTTTTATGACTAATTAATCAATAAAAAATGCAA	732
TU2	TTTTTATTTTTTTTATGTAAGTATATTTCTTTCTCTGTAATAAATTAATTTTATGATAAAA	487
TU1x	CATCTTTAATTTACTTTATTTGATCTAATTTGAGTTAAAGCTCTTTCTGATTATCTATAAAT	792
TU2	TATGAAAAAAGAGATATAATTTTCAAAAAATTTATTTAGAAAAATTAATCTTTATATAT	547
TU1x	TATTC-ACATGAAA-GATTAATTTT	815
TU2	TATTCAGTTGAAAAATAACATTTT	572
TU1x	GTTATCGGCCCATCTAAGGTGAAAAACACCGGATCCCATTCGAACTCCGAAGTTAAGCGCC	875
TU2	GCTGTGCGCCATCTAAGGTGAAAAACACCGGATCCCATTCGAACTCCGAAGTTAAGCGCC	632
TU1x	TTAAGGCTGGGTTAGTACTAAGGTGGGGGACCGCTTGGGAAGTCCCAAGTGCATTAACCT	935
TU2	TTAAGGCTGGGTTAGTACTAAGGTGGGGGACCGCTTGGGAAGTCCCAAGTGCATTAACCT	692
TU1x	TTTTAGTTTATTTTGAATATTTATTTTAAATGGTATTATGTAGATTATGCAATTTTGA	995
TU2	TTTTTATTTTTTTTCTCTCTCTTCTATAAAAAATAGATTGCAATTTCTATAGAAATATA	752
TU1x	TCATTTTGTGTTTACAAAAATGGTTGAAAAATTAAGAAAAAAGAAATTTAAATAAATAT	1055
TU2	ATAAATATTAGAAATAAATAAATTAATTTTTCAGAAAAATAGAATTAAGATACAAAAATAT	812
TU1x	CGTTAGATTATTTTATGAGTTTGTCTCTATTTTATAGAAATCATTTTGAATAAGCA	1115
TU2	AATATAATTTTGATACAATCTCTCTCTTTGATTGTTGATCATACAATCTATTCATTAA	872
TU1x	TAGAATATTCAGTTTAAATTCACCAAAAAATAAGAACATTGAAATTTATTTTGTGTA	1175
TU2	TTAAATATGAGAACAAATTTCTCTATCAAAATTTATGATAGAGGCAAAAAATTTCTCTAT	932
TU1x	TCTACTTCAAAAAACAGATTAACTTTTCTTATCTATAAATCATAAAAATATGATATAT	1235
TU2	TTTCTGATCATAATCTTTCTCTTAAATGATCAGTTTCTCTTATTTATTTAGGAATA	992
TU1x	ACTCAAAATTTATCAAACTGATTATGTTTCTCTTTTAAACAGAAAAATTTGTGTTTG	1295
TU2	GATTCTGAAATTCCTCTAATTTTCATAGCTAATCAGATTCAATCATATCTATTTTATTTT	1052
TU1x	TTTCTCAATAAAAAATCAATAAATACAGAGATAACTTAAAAATAGTACTTAAAAAGCTGAA	1355
TU2	CCACTTTAAAAAGTCTTTTTTTACCACAGTAAAAATCATAATCAATTCATAAAAAAATAAT	1112
TU1x	TTAAAAATA	1363
TU2	AACTCTAAATAAGAAACACTCAAAATCAACACATTAACGATTTTACCAAAACAACA	1167

Fig. 2. Alignment of the nucleotide sequences of the 5S ribosomal genes present in TU1x and TU2 genomic clones. The sequences were aligned according to the coding regions of the 5S genes. In the spacer regions gaps (-) are introduced to optimize similarities.

class B, with a G at position 24, a G:C at the base-paired position 19:59, a T at position 115 and an A at position 117. Also, we found an A at positions 38 and 44, which are bases of low variability in 5S rRNA sequence according to Wolters and Erdmann [11]. In spite of the extreme conservation, we found some base heterogeneity at positions 2 and 4. This heterogeneity has also been observed in *T. thermophila* [9]. All substitutions are complementary changes in the base-paired region 2:117 and 4:115 of the 5S RNA structure [7,12], suggesting that all four genes are functional.

The spacer regions of the *T. pyriformis* 5S ribosomal genes are smaller (142/145 bp long) than the one previously reported for *T. thermophila* (166 bp long) [9]. However, both clusters reveal a high A + T content of 82% to 87%. Comparative sequence analysis of spacer regions between the two species allows the detection of some conserved stretches of T residues.

This motif is present in both 5' and 3' flanking sequences at the same relative positions. It is possible that this type of motif could be related to the transcription termination [9,13].

The transcription of the 5S rRNA genes by the RNA polymerase III is regulated by intragenic control regions (ICR), the so-called boxes A and C [14]. Box A, which is located between +50 and +70, interacts with the transcription factor TFIIB, and box C, which is located between +75 and +90, interacts with TFIIA [15,16]. Box A of the 5S genes is related to box A of other class of RNA polymerase III transcribed genes, the tRNA genes, where it is assumed to play a similar role in transcription initiation. However, sequence similarities are relatively slight, and no consensus sequence between boxes A of the 5S genes and of several tRNA genes of *Tetrahymena* [17-19] was found. Although the ICR are enough to control transcription of the 5S genes, in some organisms a region located at position -30 that contains a TATA-box like has been shown to be essential to transcription [16,20,21]. In *Tetrahymena*, the analysis of different RNA polymerase II transcribed genes suggests that the role of the TATA-box may be played by an AGAGA-box [22]. However, this box is not present in the promoter region of the 5S ribosomal genes in either *T. pyriformis* or *T. thermophila*. Curiously, a potential TATA-box was detected, at position -30, in every second 5S gene of each cluster from both ciliates. So, the possibility of an external promoter in these genes cannot be ruled. No other described RNA polymerase III external control elements [14] were found.

According to our results it seems that *T. pyriformis* contains several 5S ribosomal genes clustered in tandem of two or more repeats spread over the genome. This circumstance might be related to the fact that ribosome formation in *Tetrahymena* is a rather complex process and that many factors are likely to be involved in its regulation. Curiously, two of these clusters are flanked by ubiquitin genes. However, this organization has not a necessary and obvious functional correlation.

References

- 1 Elliot, A.M. and Kennedy, J.R. (1973) Biology of *Tetrahymena* (Elliot, A.M., ed.), Ch. 2, pp. 57-87, Dowden, Hutchinson & Ross, Stroudsburg, PA.
- 2 Hanyu, N., Kuchino, Y. and Nishimura, S. (1986) EMBO J. 5, 1307-1311.
- 3 Barahona, I., Soares, H., Cyrne, L., Penque, D., Denoulet, P. and Rodrigues-Pousada, C. (1988) J. Mol. Biol. 202, 365-382.
- 4 Neves, A.M., Barahona, I., Galego, L. and Rodrigues-Pousada, C. (1988) Gene 73, 87-96.
- 5 Neves, A.M., Guerreiro, P. and Rodrigues-Pousada, C. (1991) J. DNA Seq. Mapping 2, 173-180.
- 6 Neves, A.M., Guerreiro, P., Miquelol, L. and Rodrigues-Pousada, C. (1991) Mol. Gen. Genet. 230, 186-192.

- 7 Preparata, R.M., Meyer, E.B., Preparata, F.P., Simon, E.M., Vossbrink, C.R. and Nanney, D.L. (1989) *J. Mol. Evol.* 28, 427-441.
- 8 Kumazaki, T., Hori, H. and Osawa, S. (1983) *J. Mol. Evol.* 19, 411-419.
- 9 Pederson, D.S., Yao, M.-C., Kimmel, A.R. and Gorovsky, M.A. (1984) *Nucleic Acids Res.* 12, 3003-3021.
- 10 Barahona, I., Galego, L., Jacquet, M. and Rodrigues-Pousada, C. (1985) *FEBS Lett.* 188, 227-232.
- 11 Wolters, J. and Erdmann, V.A. (1986) *J. Mol. Evol.* 24, 152-166.
- 12 Sneath, B., Vary, C., Pavlakis, G. and Vournakis, J. (1986) *Nucleic Acids Res.* 14, 1365-1378.
- 13 Bogenhagen, D.F. and Brown, D.D. (1981) *Cell* 24, 261-270.
- 14 Kunkel, G.R. (1991) *Biochim. Biophys. Acta* 1088, 1-9.
- 15 Bieker, J.J., Martin, P.L. and Roeder R.G. (1985) *Cell* 40, 119-127.
- 16 Tyler, B.M. (1987) *J. Mol. Biol.* 196, 801-811.
- 17 Endoh, H., Nagahashi, S. and Okada, N. (1989) *Nucleic Acids Res.* 17, 10122.
- 18 Kuchino, Y., Mita, T. and Nishimura, S. (1981) *Nucleic Acids Res.* 9, 4557-4562.
- 19 Kuchino, Y., Hanyu, N., Tashiro, F. and Nishimura, S. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4758-4762.
- 20 Garcia, A.D., O'Connell, A.M. and Sharp, S.J. (1987) *Mol. Cell. Biol.* 7, 2046-2051.
- 21 Morton, D.G. and Sprague, K.U. (1984) *Proc. Natl. Acad. Sci. USA* 81, 5519-5522.
- 22 Brunk, C.F. and Sadler, L.A. (1990) *Nucleic Acids Res.* 18, 323-329.

APÊNDICE 2

**Sequenciação parcial e análise da subunidade 2
da RNA polimerase II de *Tetrahymena thermophila***

A análise das sequências a jusante do gene poliubiquitina no fragmento TTUλ3X permitiu assinalar a presença de uma segunda região codificante, orientada inversamente em relação ao gene ubiquitina (ver Fig. IV.2). Uma pesquisa nos bancos de dados de sequências EMBL/GenBank e SwissProt permitiu identificá-la como correspondendo a um gene ortólogo dos genes codificantes para a segunda maior subunidade da RNA polimerase II em eucariotas, *RBP2*.

A sequência nucleotídica e respectiva tradução, apresentada na Fig.A2.1, revela que apenas os últimos 123 resíduos deste gene se encontram clonados, sendo o resíduo Gly(36) interrompida por um intrão de 67 nucleótidos. A presença de intrões nestes genes é usual, já que três dos quatro genes conhecidos contêm intrões: um nos genes *RPB2* de *S.pombe* e de *E.octocarinaus*, e dois no gene de *D.melanogaster*. Contudo, nenhum destes intrões se localiza numa posição equivalente à do intrão em *T.thermophila*. As sequências flanqueantes do intrão, GTA../.TAG, são perfeitamente conservadas em relação às encontradas noutros intrões de *T.thermophila* e *T.pyriformis* (Csank *et al.*, 1990; Soares *et al.*, 1994).

	S R G P V V A I V R Q P T H G R S R G G	
1	TCTAGAGGTCCAGTTGTAGCAATTGTCAGACAACCTACTCACGGTCGTTCCAGAGGTGGT	60
	G L R F G E M E R D C I I A H	
61	GGTTTACGTTTTGGTGAAATGGAAAGAGATTGTATCATTGCTCACGgtaaattttatctca	120
		G T S
121	taattctattttaaatcttggtttttattaaatgatattattttaaaatttactagGTACCAG	180
	N F L K E R T F G V S D Q F R T H V C S	
181	TAATTTCTTAAAAGAAAGAACATTTGGCGTTTCTGATCAGTTCAGAACCCATGTTTGCTC	240
	R C G L F A K A D V T N H K Y D C T G C	
241	AAGATGTGGTTTGTGTTGCTAAAGCAGATGTTACAAATCACAAGTATGACTGTACTGGCTG	300
	S Q G D K K H Y K I V Q V Y L P Y A A K	
301	CTCTTAAGGTGATAAGAAGCACTACAAAATCGTCCAAGTGTATCTCCCATATGCAGCTAA	360
	Q L I Q E L M A M H I V P R I K V D V T	
361	GCAACTTATTCAAGAACTTATGGCTATGCATATTGTTCCAAGAATTAAAGTTGATGTCAC	420
	S K K Y V *	
421	TAGTAAAAAGTATGTTTGATAaaatctaatacattttaaaaatatcatttttatggtccttagt	480
481	ccatctatccataataaattgatataaaaaaattgctaataaaatattttttaaaattaagc	540
541	tttgctctgtgtttcgtatattaaaatttaaaatttccttggttttagttgatttttagagtta	600
601	attaaaattttattttattttatttaattgaattttattttgtttttgaaaaatttaa	660
661	ttcaaaaatgttttttttagaattgtaatttttatgaaactaaataataaaaaagaaaata	720
721	tatatattataaaaaaattttaactacactaaagatttttaaaactttaactaaaaagaaa	780
781	aaatattacattaatgtgttattagctgttagtttattataattaattaattgaaaag	840
841	caggaagatttcaattattttatttttaggtgtttaattgaattttatttttaaaatga	900
901	aaaattagttaaaaatctatttttaaaaagcaaaactttaatgataaaaaagtttaataaa	960
961	gttttagcaataatagtaaaattcatagttttaatttccttaatttttaaaaatttaatat	1020
1021	ttaaaagaaaaaaaataatttcagcatattttatttactgctaattaaagactattaat	1080
1081	tattatttagctttattttagtaatcagaagatttaaaattaattcgcttaggttatttaa	1140
1141	ttaaatttgatttttacatgaataacttgattaaaaaaatatttgctttattaggaaaat	1200

Fig.A2.1. Sequência do gene codificante para a segunda maior subunidade da RNA polimerase II de *T.thermophila*, *RBP2*. A sequência corresponde a uma região de 1200 nucleótidos do clone pTUλ3X localizada a jusante do gene poliubiquitina que codifica para região C-terminal do gene *RBP2* de *T.thermophila*, apresentada como complementar inversa devido à orientação oposta de ambos os genes (ver Fig.IV.2). O intrão e a região não traduzida localizada a jusante do gene são apresentadas em letra pequena.

O alinhamento da sequência conhecida em *T.thermophila* com as regiões C-terminais equivalentes das outras RBP2 conhecidas revela um elevado grau de homologia, característico das regiões C-terminais destas enzimas (Fig.A2.2). A região apresentada inclui dois dos 18 domínios definidos por Acker e colaboradores (1992), por comparação entre nove organismos que incluem eucariotas, procariotas e cloroplastos: o domínio Q (incompleto, de Ser(1) a Ser(49)), e o domínio R (de Ala(96) a Ile(109)). Estes dois domínios, que se encontram igualmente muito conservados na sequência obtida em *T.thermophila*, parecem ser importantes para o funcionamento da enzima, já que mutações aí localizadas afectam a pausa e terminação da RNA polimerase bacteriana (Landick *et al.*, 1990). O domínio Q, em particular, é o mais extenso dos 18 domínios definidos e um dos quatro que são significativamente hidrofílicos, localizando-se igualmente aí um dos possíveis locais de ligação de nucleótidos durante a iniciação da transcrição (Mustaev *et al.*, 1991).

A região compreendida entre estes dois domínios é, a par da extremidade C-terminal, a mais divergente entre as sequências analisadas na Fig.A2.2. Esta região inclui, em *T.thermophila*, uma inserção de 4 aminoácidos que constitui o dado mais assinalável na sequência encontrada neste organismo, já que não é observada na de outros organismos, entre as quais a do ciliado *Euplotes octocarinatus*. A inserção segue-se à última Cys de um motivo "zinc finger" do tipo CX₂CX₉₋₁₅CX₂C, constituído pelas Cys (57), (60), (75) e (78), que é conservado entre as RNA polimerases eucariotas e arqueobacterianas (Acker *et al.*, 1992). Idêntico motivo foi igualmente encontrado na região N-terminal da maior subunidade da RNA polimerase II, RBP1, sugerindo que as duas maiores subunidades desta enzima podem interagir através deste tipo de motivo (Treich *et al.*, 1991).

RBP2_TTHER	SRGPVVAIVRQPTHTGRSRGGGLRFGEMERDCIIAHGTSNFLKERTFGVSDQFRTHVCSRCGL	(62)
RBP2_EUPLO	LQILT E H MVS AAR L D CYTV RI	1142
RBP2_SCHPO	A QILT VE D Q S C SV R L DC AY VI DI	1157
RBP2_YEAST	A MQVLT VE D M AAS LMEA A V I GI	1168
RBP2_DROME	A QIL ME A D Q S AAQ R L E PY V I NF	1126
RBP2_HUMAN	A IQILN ME D Q AAQ R L EA PYQV NL I	1124

RBP2_TTHER	FA-KADVNTNHKYDCTGCSQGDKKHYKIVQVYLPYAAKQLIQELMAMHIVPRIKVDVTSKKYV*	(123)
RBP2_EUPLO	IC-E NLRQQ L R ----QNSTN C G C L F Q A MITS S*	1194
RBP2_SCHPO	I -I SYKKDS E RS ----QNRTRFS L F S N A LFTKNHK*	1210
RBP2_YEAST	MTVI KLNHNQFE K ----DNKID Y IHI L F N T LYT RSRDF*	1224
RBP2_DROME	I -I NLR NTFE K ----NKTQ S R L F S N A LM T*	1176
RBP2_HUMAN	M -I NTRT T E R ----RNKTQ SL RM C L F S S A MMSV*	1174

Fig.A2.2. Alinhamento das regiões C-terminais das RBP2 de eucariotas. Apenas as diferenças para a sequência de *T.thermophila* são indicadas; o traço (-) representa deleções na sequência, e a estrela (*) o codão de terminação. Os domínios Q e R (Acker *et al.*, 1992) encontram-se sublinhados, e os resíduos Cys do motivo "zinc finger" encontram-se a negrito. TTHER, *T.thermophila*; EUPLO, *Euplotes octocarinatus* (Kaufmann *et al.*, 1992); SCHPO, *Schizosacchchromyces pombe* (Kawagishi *et al.*, 1993); YEAST, *Saccharomyces cerevisiae* (Sweetser *et al.*, 1987); DROME, *Drosophila melanogaster* (Falkenburg *et al.*, 1987); HUMAN, *Homo sapiens* (Acker *et al.*, 1992).

Oferta
3.000x100
96.05.29

BIBLIOTECA
DO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS
"ABEL SALAZAR"