

UNIVERSITÉ DU QUÉBEC

MÉMOIRE

PRÉSENTÉ

À L'INSTITUT ARMAND-FRAPPIER

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN

VIROLOGIE ET IMMUNOLOGIE

PAR

BENOÎT HÉBERT

IDENTIFICATION DE QUATORZE TRANSCRITS
CODANT POUR DES PROTÉINE TYROSINE KINASES
CHEZ LES CELLULES E-5;
CARACTÉRISATION PARTIELLE DE TROIS
NOUVEAUX TRANSCRITS

MAI 1993^o

The cloning of humans is on most of the lists of things to worry about from Science, along with behavior control, genetic engineering, transplanted heads, computer poetry and the unrestrained growth of plastic flowers.

- Lewis Thomas

What me worry ?

- Alfred E. Newman

*The great tragedy of Science -
the slaying of a beautiful
hypothesis by an ugly fact.*

*- T.H. Huxley
viii Biogenesis
and Abiogenesis*

TABLE DES MATIÈRES

TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES ABRÉVIATIONS	xv
SOMMAIRE	xvii
REVUE BIBLIOGRAPHIQUE	1
1. Physiologie du thymus murin	2
1.1 Morphologie du thymus murin	2
1.2 Organisation du stroma thymique	2
1.2.1 Cellules épithéliales thymiques	2
1.2.2 Macrophages et cellules dendritiques thymiques	4
2. Différenciation intrathymique des lymphocytes T	6
2.1 Aperçu des étapes de différenciation et de sélection ..	6
2.2 Nature et rôle des interactions lympho-stromales dans la différenciation	9
3. Le contrôle du stroma par les cellules lymphoïdes	12
3.1 Avant-propos	12
3.2 Restauration de la médulla thymique chez les souris SCID	13
3.3 Modèle de la régulation du stroma de la moelle osseuse	14
3.4 Un modèle <i>in vitro</i> permettant l'étude du contrôle du stroma	16
3.4.1 Avant-propos	16
3.4.2 Le modèle <i>in vitro</i> E-5 pour étudier les <i>signaux</i> <i>inversés</i>	16
4. Les protéines kinases	20
4.1 Avant-propos	20
4.2 Régulation métabolique par les protéine kinases	21
4.3 Caractéristiques structurales des protéines kinases	23
4.3.1 Généralités	23

4.3.3	L'évolution des protéines kinases	25
4.3.4	Motifs conservés des domaines catalytiques . . .	26
4.3.5	Caractéristiques moléculaires des protéine tyrosine kinases	36
4.3.6	Choix des amorces et amplification des domaines catalytiques des protéine tyrosine kinases	36
OBJECTIFS DE LA RECHERCHE		40
MATÉRIEL & MÉTHODES		42
1.	Les cellules E-5	43
1.1	Conditions de culture des cellules E-5	43
2.	Purification du complexe gp23/45/90	44
2.1	Préparation de la fraction membranaire de cellules E-5 .	44
2.2	Préparation des glycoprotéines membranaires	45
2.2.1	Solubilisation des membranes cellulaires	45
2.2.2	Purification des glycoprotéines membranaires par chromatographie sur Sepharose 4B/Lentil- Lectin	45
2.3	Co-immunoprécipitation du complexe gp23/45/90	46
2.3.1	Marquage des glycoprotéines membranaires à l'iode radioactif	46
2.3.2	Immunoprécipitations et analyse par SDS-PAGE .	47
2.4	Purification du complexe gp23/45/90 par chromatographie d'affinité	47
2.4.1	Couplage d'anticorps au support Sepharose-4B activé au CNBr	47
2.4.2	Purification par chromatographie d'affinité	48
3.	Caractérisation moléculaire des cellules E-5	49
3.1	Purification du RNA total des cellules E-5	49
3.2	Amplification des domaines catalytiques de protéine tyrosine kinases	51
3.2.1	Liste des amorces	51
3.2.2	Synthèse du DNA complémentaire (cDNA)	51
3.2.2.1	Synthèse à l'aide de MuLV-RT	52
3.2.2.2	Synthèse à l'aide de Superscript™ RT	52
3.2.3	Amplification par PCR des motifs VI-IX des domaines catalytiques des protéine tyrosine kinases	53
3.3	Analyse sur gel d'agarose	56

3.4	Stratégies de clonage	56
3.4.1	Vecteurs utilisés	56
3.4.2	Clonage des motifs VI - IX des domaines catalytiques de protéine tyrosine kinases	56
3.4.2.1	Préparation des produits de PCR	56
3.4.2.2	Préparation du vecteur	57
3.4.2.3	Ligation	57
3.5	Transformation des cellules bactériennes	58
3.5.1	Souche bactérienne utilisée	58
3.5.2	Préparation de bactéries compétentes	58
3.5.3	Sélection des clones	59
3.5.4	Mini-préparations par PCR	60
3.5.5	Mini-préparation de plasmides (Magic Mini- Preps™)	61
4.	Séquençage	62
4.1	Liste des amorces pour le séquençage	62
4.2	Réactions de séquençage manuel	62
4.3	Électrophorèse de séquence	63
4.4	Séquençage par la méthode automatique: A.L.F. (<i>automated light fluorescence</i>)	65
5.	Analyse des séquences obtenues	67
5.1	<i>Protein Kinase Catalytic Domain Database</i> (PKCDD)	67
5.2	Logiciels utilisés	67
5.2.1	A.L.F. Manager™	67
5.2.2	ReadSeq	68
5.2.3	DNASIS et PROSIS	68
5.2.4	<i>Sequence Analysis Software Package</i> ™ (S.A.S.P.) de Genetics Computer Group (GCG)	69
5.2.4.1	Organisation du logiciel S.A.S.P. de GCG	69
5.2.4.2	FastA	69
5.2.4.3	PileUP	70
5.2.5	Alignements multiples à l'aide du logiciel <i>MACAW (Multiple Alignment and Construction Analysis Workbench)</i>	70
RÉSULTATS		72
1.	Avant-Propos	73
2.	Caractérisation biochimique du complexe gp23/45/90	75
2.1	Purification de la gp90 par chromatographie d'affinité	75

2.2	Purification de la gp90 par co-immunoprécipitation . . .	76
3.	Caractérisation moléculaire des protéine tyrosine kinases des cellules E-5	82
3.1	Amplification des régions VI - IX des domaines catalytiques des protéine tyrosine kinases à partir de cDNA de cellules E-5.	82
3.2	Clonage des fragments de domaines catalytiques obtenus par PCR	85
3.3	Criblage par PCR	86
3.4	Séquençage: établissement du catalogue des mRNAs PTK-spécifiques présents chez les cellules E-5.	91
3.5	Nouvelles protéine tyrosine kinases	93
3.5.1	Clone 109c -- une protéine tyrosine kinase de type récepteur homologue à <i>tek</i>	93
3.5.2	Clone 060e -- une protéine kinase membre de la sous-famille <i>src</i>	100
3.5.3	Clone 113e -- une protéine kinase membre de la famille <i>eph</i>	106
	DISCUSSION	118
1.	Stratégie d'amplification	119
1.1	Avant-propos	119
1.2	Amplification par PCR	120
1.2.1	Choix des amorces	120
1.2.2	Amplification des séquences codant pour les motifs VI-IX des domaines catalytiques de protéine tyrosine kinases.	122
1.3	Adaptations techniques	124
1.3.1	Augmentation de la sensibilité et de la spécificité de l'amplification par PCR suite à l'utilisation de paraffine comme barrière d'évaporation.	124
1.3.2	Criblage par PCR	125
1.3.3	Complications du séquençage: structures secondaires et compressions	126
2.	Les cellules E-5 transcrivent au moins onze gènes codant pour des PTKs déjà répertoriées.	132
2.1	Les banques de cDNA par PCR pour l'identification des PTKs	132

2.2	<i>RYK</i> : une PTK de type récepteur qui présente des idiosyncrasies dans trois motifs du domaine catalytique.	136
2.3	La présence de certaines PTKs peut être liée à la transformation cellulaire.	138
2.4	PTKs de type récepteur agissant comme récepteurs de facteurs de croissance	139
2.5	La présence de <i>tyro3</i> suggère un lien entre le système nerveux et le système immunitaire.	139
3.	Les cellules E-5 transcrivent au moins trois gènes codant pour des PTKs non-répertoriées.	141
3.1	Le processus décisionnel permettant la découverte de nouvelles PTKs.	141
3.2	Une nouvelle PTK de type récepteur: le clone 109c est homologue aux membres de la famille du FGF-R.	144
3.3	Un nouveau membre de la sous-famille <i>src</i> : le clone 060e présente une homologie élevée à <i>c-yes</i>	146
3.4	Un nouveau membre de la sous-famille <i>eph</i>	151
3.4.1	Les membres de la sous-famille <i>eph</i> sont exprimés spécifiquement dans certains tissus.	151
3.4.2	Particularités moléculaires du clone 113e	153
4.	Résumé et conclusion générale	157
	REMERCIEMENTS	160
	BIBLIOGRAPHIE	162
	APPENDICE I	181
	APPENDICE II	197

LISTE DES TABLEAUX

MATÉRIEL & MÉTHODES

TABLEAU I	Séquence des amorces utilisées pour la synthèse du cDNA et pour l'amplification des motifs VI à IX de domaines catalytiques de PTKs.	51
TABLEAU II	Concentration des amorces utilisées pour la synthèse du cDNA.	53
TABLEAU III	Conditions d'électrophorèse pour le séquençage automatique A.L.F.	66

RÉSULTATS

TABLEAU IV	Conditions de lavage, d'élution et de neutralisation utilisées pour la mise au point de la purification du complexe gp23/45/90 par chromatographie d'affinité.	75
TABLEAU V	Liste des protéine kinases identifiées suite au clonage des régions VI-IX des domaines catalytiques	92

LISTE DES FIGURES**REVUE BIBLIOGRAPHIQUE**

FIGURE 1	Profil de similarité des domaines catalytiques de protéines kinases.	27
FIGURE 2	Alignement multiple de protéine tyrosine kinases représentatives des principales sous-familles connues	29
FIGURE 3	La matrice PAM250: matrice de probabilité MDM de substitutions d'acides aminés.	30
FIGURE 4	Domaines catalytiques de la famille des protéine-tyrosine kinases.	35
FIGURE 5	Alignement des motifs VI et IX des domaines catalytiques utilisés pour l'élaboration des amorces de PCR.	39

MATÉRIEL ET MÉTHODES

FIGURE 6	Cycles de PCR pour l'amplification des domaines catalytiques des PTKs.	55
----------	--	----

RÉSULTATS

FIGURE 7	Séparation par SDS-PAGE de fractions obtenues suite à l'élution de la colonne d'affinité C3C12, anti-gp23.	79
FIGURE 8	Immunoprécipitation de glycoprotéines membranaires et de fractions de protéines membranaires.	81
FIGURE 9	Amplification des motifs VI-IX des domaines catalytiques de PTKs à partir de différents cDNAs.	84

LISTE DES FIGURES (suite)

RÉSULTATS (suite)

FIGURE 10	Stratégie utilisée pour le clonage des motifs VI-IX des domaines catalytiques de PTKs.	88
FIGURE 11	Minipréparations par PCR des clones contenant des fragments de domaines catalytiques.	90
FIGURE 12	Alignement des séquences nucléotidiques des clones 109c, 111c et 089c.	95
FIGURE 13	Séquence nucléotidique et motifs conservés du clone 109c.	97
FIGURE 14	Recherche d'homologie du clone 109c.	99
FIGURE 15	Séquence nucléotidique et motifs conservés du clone 060e.	102
FIGURE 16	Alignement multiple des motifs VI à IX entre le clone 060e et trois membres de la sous-famille <i>src</i> .	104
FIGURE 17	Alignement de la séquence nucléotidique et peptidique déduite du clone 060e avec <i>c-yes</i> .	105
FIGURE 18	Séquence nucléotidique et motifs conservés du clone 113e.	109
FIGURE 19	Alignement multiple des motifs VI à IX entre le clone 113e et des membres de la sous-famille <i>eph</i> .	111
FIGURE 20	Alignement de la séquence nucléotidique et peptidique déduite du clone 113e et <i>Cek-5</i> .	112

LISTE DES FIGURES (suite)

RÉSULTATS (suite)

FIGURE 21	Recherche d'homologie de la séquence nucléotidique du clone 113e avec les principaux membres de la sous-famille <i>eph</i> .	
	a- Identité du clone 113e avec <i>erk</i> et <i>eck</i> .	114
	b- Identité du clone 113e avec <i>elk</i> et <i>eek</i>	115
	c- Identité du clone 113e avec <i>eph</i> et <i>eek</i>	116
FIGURE 22	Recherche d'identité entre la séquence peptidique déduite du clone 113e et les membres de la famille <i>eph</i> .	117

DISCUSSION

FIGURE 23	Structure secondaire suggéré par l'algorithme de Zuker et Stiegler ainsi qu'un exemple de compressions observées lors du séquençage manuel par la méthode de Sanger.	128
FIGURE 24	Compressions causées par les structures secondaires lors du séquençage automatique. (a), le clone 085e séquencé avec l'amorce RP. (b), le clone 085e séquencé en direction inverse avec l'amorce UP.	130-131
FIGURE 25	Alignement multiple de huit clones correspondant à des PTKs déjà connues.	134
FIGURE 26	Recherche d'identité du clone 024e à l'aide du logiciel FASTA.	145
FIGURE 27	Alignement multiple des domaines catalytiques des principaux membres de la famille <i>src</i> .	148

LISTE DES FIGURES (suite)**DISCUSSION (suite)**

FIGURE 28	Substitutions silencieuses au motif VIII du clone 113e.	155
FIGURE 29	Alignements des motifs VIII entre le clone 113e et les principaux membres de la sous-famille <i>eph</i> .	156

LISTE DES ABRÉVIATIONS

Les abréviations sont conformes avec la nomenclature adoptée par le Conseil de recherches médicales du Canada dans le lexique de terminologie *Vocabulaire du génie génétique*, publié en 1990.

ATP	:	Adénosine triphosphate
BSA	:	Albumine sérique bovine (<i>Bovine serum albumin</i>)
cDNA	:	DNA complémentaire (<i>complementary DNA</i>)
CMH	:	Complexe majeur d'histocompatibilité
CTL	:	Lymphocyte T cytotoxique (<i>Cytotoxic T lymphocyte</i>)
CTP	:	Cytosine triphosphate
DAG	:	Diacylglycérol
D.O.	:	Densité optique
DNA	:	Acide désoxyribonucléique (<i>Deoxyribonucleic acid</i>)
E-5gp	:	Extrait de glycoprotéine de cellules E-5
E-5m	:	Lysat de membranes plasmiques de cellules E-5
EDTA	:	Acide éthylènediaminetétraacétique
EGF	:	Facteur de croissance épidermique (<i>Epidermal growth factor</i>)
<i>g</i>	:	Unité de gravité
GTP	:	Guanosine triphosphate
HBSS	:	<i>Hanks' balanced salt solution</i>
IFN	:	Interféron
Ig	:	Immunoglobuline
IL	:	Interleukine
kD	:	Daltons (1000)
mAb	:	Anticorps monoclonal (<i>Monoclonal antibody</i>)
MEM	:	<i>Minimum essential medium</i>
MNNG	:	<i>N</i> -methyl- <i>N'</i> -nitro- <i>N</i> -nitrosoguanidine
mRNA	:	ARN messenger (<i>Messenger RNA</i>)
ND	:	Non déterminé
NP-40	:	Nonidet P-40
PAGE	:	électrophorèse sur gels de polyacrylamide (<i>Polyacrylamide gel electrophoresis</i>)
PBS	:	<i>Phosphate-buffered saline</i>
PDGF	:	Facteur de croissance des plaquettes (<i>Platelet-derived growth factor</i>)
PMSF	:	Fluorure de phénylméthylsulfonyl
PSI	:	Livres par pouce carré (<i>Pounds per square inch</i>)
PTK	:	Protéine tyrosine kinase
PSTK	:	Protéine sérine thréonine kinase
PVC	:	Chlorure de polyvinyl

R	:	Récepteur (e.g., IL-2R, FGF-R)
RNA	:	Acide ribonucléique (<i>Ribonucleic acid</i>)
SDS	:	Dodécyl-sulfate de sodium (<i>Sodium dodecyl sulfate</i>)
TCR	:	Récepteur de l'antigène des cellules T
Tris	:	Tris(hydroxyméthyl)aminométhane
TTP	:	Thymidine triphosphate

SOMMAIRE

La différenciation intra-thymique des lymphocytes T est dépendante des cellules du stroma. La sécrétion de cytokines et les interactions moléculaires sont responsables de son contrôle. L'importance des interactions moléculaires a été démontrée à l'aide de modèles *in vitro*, puis tout récemment, chez des souris transgéniques. La caractérisation biochimique de plusieurs molécules d'adhésion a permis d'établir une classification fonctionnelle et structurale. Bien qu'une multitude de molécules d'adhésion existe, les profils d'activation se ressemblent. La caractérisation moléculaire des profils d'activation a démontré que les mécanismes d'amplification de signaux sont intimement reliés soit par des mécanismes communs soit par des effecteurs intracellulaires similaires.

Au cours des dernières années, l'étude de la transduction de signaux intra-thymiques a été effectuée presque uniquement chez les thymocytes. Par contre, la régulation de l'activité des cellules stromales par les thymocytes a été démontrée récemment. Pour la première fois, un modèle *in vitro* d'interactions lympho-stromales a permis de démontrer que les thymocytes peuvent transmettre des signaux aux cellules stromales. Cette interaction est possible par l'intermédiaire du complexe gp23/45/90 retrouvé chez certaines cellules stromales. Ce complexe a été purifié et partiellement caractérisé à partir de la lignée cellulaire d'épithélium thymique E-5.

L'interaction entre des thymocytes et les cellules E-5 déclenche spécifiquement, entre autres, la phosphorylation de résidus tyrosine de trois protéines des cellules E-5. Compte tenu de l'implication d'une protéine tyrosine kinase dans le modèle d'activation *in vitro* des cellules E-5 et de l'importance des protéine tyrosine kinases dans les signaux de transduction,

nous avons entrepris de répertorier les transcrits PTK-spécifiques des cellules E-5 afin d'approfondir la caractérisation moléculaire de l'interaction E-5 - thymocytes.

Sachant que les protéine tyrosine kinases possèdent des caractéristiques moléculaires communes, il est possible d'amplifier sélectivement, par PCR, des domaines catalytiques distincts. Nous avons procédé à l'amplification et au clonage des domaines catalytiques des PTKs des cellules E-5. Par le séquençage des clones obtenus, nous avons identifié onze transcrits différents codant pour des PTKs connues. De plus, nous avons démontré la présence de trois nouveaux transcrits de PTKs. Le premier présente une homologie à *tek*, le deuxième à *c-yes*, et le troisième, à *Cek-5*, un membre de la sous-famille de protéine tyrosine kinase *eph*. Ce mémoire présente les séquences nucléotidiques et peptidiques (déduites) ainsi que l'homologie des nouvelles protéine tyrosine kinases avec celles déjà connues. Le rôle potentiel de ces protéine tyrosine kinases est également discuté.

REVUE BIBLIOGRAPHIQUE

1. Physiologie du thymus murin

1.1 Morphologie du thymus murin

Le thymus est un organe lymphoïde primaire responsable du développement des lymphocytes T. Localisé près du cœur, il est recouvert d'une capsule et est formé de deux lobes, chacun composé de lobules. Chaque lobule est caractérisé par une région corticale (externe) et une région médullaire (interne). La trame du thymus est composée essentiellement de cellules épithéliales provenant de la troisième cavité brachiale ectodermique ainsi que du troisième feuillet pharyngien (Auerback, 1961; van Ewijk, 1984). Ces cellules épithéliales sont responsables du développement des lymphocytes T.

Contrairement au cortex, la médulla est de densité cellulaire plus faible et les cellules ne possèdent pas de réseau étendu de prolongements cytoplasmiques (van Ewijk, 1988). Les deux régions principales sont donc structurellement différentes et possèdent des types cellulaires qui leur sont propres. L'agencement de ces différents types cellulaires permet une organisation structurale modulaire du thymus, d'où la notion de micro-environnements.

1.2 Organisation du stroma thymique

1.2.1 Cellules épithéliales thymiques

Les cellules épithéliales forment la population majeure de la trame du thymus. Plusieurs phénotypes de cellules épithéliales sont retrouvés dans des compartiments distincts.

Les cellules épithéliales du cortex sont majoritairement caractérisées par de longs prolongements cytoplasmiques reliés par des desmosomes (type I). Ces ramifications forment un filet tridimensionnel au travers duquel les thymocytes interagissent durant leur différenciation (van Ewijk, 1988 et revue en 1991). Contrairement aux cellules du cortex, les cellules épithéliales de la médulla sont d'apparence plutôt sphérique et présentent des prolongements cytoplasmiques plus courts (type II).

Les cellules épithéliales diffèrent également par l'expression sélective d'antigènes de surface. Les cellules épithéliales corticales expriment un niveau élevé de molécules de classe II du CMH sans exprimer pour autant de molécules de classe I. D'autre part, les cellules épithéliales médullaires expriment fortement les molécules de classe I et II du CMH (van Ewijk *et al.*, 1980; Barclay and Mayrhofer, 1981).

En plus de critères morphologiques, l'expression spécifique d'antigènes de surface permet également d'établir une classification. Les cellules épithéliales du cortex peuvent se distinguer des cellules de la médulla à l'aide d'anticorps monoclonaux. Entre autres, les travaux du groupe de van Ewijk ont démontré la réactivité des anticorps de la série ER-TR avec différentes populations de cellules épithéliales. Les cellules épithéliales corticales sont reconnues par l'anticorps ER-TR4 alors que les cellules épithéliales médullaires réagissent avec l'anticorps ER-TR5 (van Vliet 1984; 1985). D'autres anticorps peuvent également être utilisés afin de caractériser les différentes populations de cellules épithéliales. Les cellules corticales présentent un antigène également exprimé chez les cellules B dont la présence est détectée par le mAb 6C3 (Adkins *et al.*, 1988). D'autre part, le mAb G8.8 réagit avec les cellules épithéliales médullaires en plus de réagir avec les cellules épithéliales sous-capsulaires (Farr *et al.*, 1991). Une autre classification peut être établie selon

l'expression différentielle de cytokératine (Farr et Braddy, 1989).

L'aspect morphologique permet également une classification dans le cas des cellules "nurse" thymiques (TNC; Wekerle et Ketelson, 1980; Wekerle *et al.*, 1980). Ces cellules peuplent la région sous-capsulaire et peuvent englober jusqu'à 200 thymocytes. Les TNCs ont un phénotype similaire aux cellules corticales (van Vliet *et al.*, 1984) et constituent un exemple unique de micro-environnement où les thymocytes se différencient en étroite association avec des cellules épithéliales spécialisées (Andrews et Boyd, 1985; de Waal-Malefijt, 1986).

1.2.2 Macrophages et cellules dendritiques thymiques

Le thymus est également peuplé de macrophages et de cellules dendritiques que l'on retrouve dans le cortex mais ils sont plus abondants dans la médulla (Duijvestijn et Hoefsmit, 1981). De plus, les cellules dendritiques sont présentes à la jonction cortico-médullaire où elles ont possiblement un rôle clé dans la différenciation des thymocytes, soit dans l'induction de tolérance (Lo, 1986), soit dans la sélection négative (Speiser *et al.*, 1989).

Les macrophages et les cellules dendritiques sont d'origine myéloïde. Contrairement aux macrophages, les cellules dendritiques ne sont pas matures à leur arrivée dans le thymus (Fairchild et Austyn, 1990). De plus, la migration de cellules dendritiques de la périphérie vers le thymus n'a jamais été démontrée. Les travaux récents du groupe de Shortman ont démontré que les cellules dendritiques ainsi que les thymocytes pourraient provenir d'un précurseur commun (Ardavin *et al.*, 1993). Le développement des cellules dendritiques est entièrement intrathymique.

L'expression de certains antigènes de surface permet de différencier phénotypiquement les macrophages des cellules dendritiques. Les macrophages et les cellules dendritiques expriment les molécules de classe I du CMH (van Ewijk, 1984) alors que seulement ces dernières expriment les molécules de classe I et II (Duijvestijn *et al.*, 1983). D'autre part, les cellules dendritiques réagissent avec le mAb ER-TR6 (van Vliet *et al.*, 1984) alors que les macrophages réagissent avec le mAb Mac-1 (Springer *et al.*, 1979) et à l'occasion avec le mAb ER-TR6.

2. Différenciation intrathymique des lymphocytes T

2.1 Aperçu des étapes de différenciation et de sélection

Le thymus est colonisé par des vagues de pré-thymocytes (Pénit et Vasseur, 1989). Bien que les précurseurs aient déjà colonisé le thymus au jour 11, la prolifération de ceux-ci est dépendante de la présence d'éléments stromaux entièrement développés (van Ewijk *et al.*, 1982; van Vliet 1985).

Les thymocytes corticaux immatures constituent environ 5% de la population totale de thymocytes. Ces thymocytes immatures sont CD4^{lo}, CD8⁻ et CD3⁻ (Wu *et al.*, 1991). Ces cellules sont connues historiquement sous le nom de "doubles négatives" (DN), les thymocytes CD4⁻ CD8⁻ constituant la population la plus immature connue avant les travaux de Wu en 1991. Les thymocytes DN n'expriment pas de TCR à leur surface (Bluestone *et al.*, 1987) mais les réarrangements des gènes de TCR ont lieu à ce stade (Raulet *et al.*, 1985). Ces précurseurs prolifèrent et se différencient pour donner lieu à la population "double positive" CD4⁺ CD8⁺ (DP) par l'intermédiaire d'un stade CD4⁻ CD8⁺ CD3⁻ ou CD4⁺ CD8⁻ CD3⁻ (Bluestone *et al.*, 1987; MacDonald *et al.*, 1988; Arase *et al.*, 1990; Hugo *et al.*, 1990 et 1991). La population des cellules DP est majoritaire; elle représente environ 75 à 80% des thymocytes. Environ 50% des cellules DP expriment le TCR $\alpha\beta$ /CD3⁺ (Roehm *et al.*, 1984). La différenciation de ces précurseurs mène à l'apparition des populations effectrices matures simple positives (SP) CD4⁻ CD8⁺ CD3⁺ ou CD4⁺ CD8⁻ CD3⁺ (Hugo *et al.*, 1990).

Le pourcentage élevé de cellules DP du thymus suggère que cette étape est cruciale dans le développement des lymphocytes T et seulement

une fraction parvient à maturité. En effet, le système immunitaire doit s'assurer de la capacité des lymphocytes T matures à reconnaître un antigène dans la périphérie et à répondre convenablement à celui-ci. Il y a donc une distinction à faire entre les antigènes du "soi" et du "non-soi".

Plus précisément, la sélection négative est responsable de la destruction des clones réactifs aux antigènes du soi présentés dans le contexte du CMH autologue (Kisielow *et al.*, 1988; Scott *et al.*, 1989). Les thymocytes exprimant un TCR non-fonctionnel sont également éliminés (Blackman *et al.*, 1990). D'autre part, le mécanisme de sélection positive s'assure de la destruction des clones ne reconnaissant pas le CMH autologue. La sélection positive est probablement sous le contrôle des cellules épithéliales corticales (Benoist et Mathis, 1989), à un stade SP au lieu du stade DP tel qu'anticipé. Par contre, la sélection négative serait sous le contrôle des cellules dendritiques corticales CMH⁺ (Speiser *et al.*, 1989) et s'effectuerait probablement au niveau du stade DP (Fowlkes *et al.*, 1988). La sélection négative est responsable de l'élimination intrathymique de la majorité des cellules DP (Scollay *et al.*, 1980, 1984). En effet, seulement 1-5% des thymocytes parviennent à maturité (Shortman et Jackson, 1974).

Le mécanisme général de sélection négative semble être l'induction d'une mort cellulaire programmée (apoptose) des thymocytes immatures autoréactifs (Cohen, 1991; Jenkinson *et al.*, 1989). L'induction résulte de la transduction de signaux suite à l'activation via le TCR. Trois observations supportent ce schéma: 1) l'activation via le TCR, soit par un mAb anti-TCR, soit par un ligand naturel CMH-antigène, entraîne une augmentation de l'apoptose des thymocytes immatures (DN) isolés (Smith *et al.*, 1989; Swat *et al.*, 1991); 2) l'injection intraveineuse de doses élevées d'ovalbumine à des souris transgéniques exprimant un TCR spécifique à l'ovalbumine

entraîne une apoptose intrathymique des DNs (Murphy *et al.*, 1990); 3) la cyclosporine A, qui inhibe la transduction de signaux, inhibe la sélection négative *in vivo* ainsi que l'apoptose induite *in vitro* par un mAb anti-TCR (Shi *et al.*, 1991; Jenkins *et al.*, 1988).

Les mécanismes précis de sélection négative restent à élucider même si plusieurs groupes ont tenté de les expliquer. Sachant que les thymocytes immatures ne démontrent presque pas de réactivité via leur TCR (Adkins *et al.*, 1987), que les molécules CD4 et CD8 participent dans la sélection intrathymique (Fowlkes *et al.*, 1988; MacDonald *et al.*, 1988) et que l'activité enzymatique de la PTK p56^{lck} est augmentée suite à la liaison de CD4 dans les cellules DNs (Veillete *et al.*, 1989), il a été proposé que la PTK p56^{lck} est impliquée dans la sélection négative. Par contre, des travaux plus récents ont démontré que la sélection négative chez des souris transgéniques est possible en présence de génistéine, un inhibiteur de PTKs (Nakayama et Loh, 1992)

Le concept de la sélection positive est plus problématique. En effet, l'utilisation des mêmes récepteurs tant pour la sélection négative que la sélection positive n'a pas encore été expliqué. Sachant que les lymphocytes T matures sont activés par un antigène présenté dans le contexte du soi et jamais par le CMH seul, il est paradoxal que la sélection positive puisse être efficace en absence d'un antigène.

Deux théories distinctes tentent d'expliquer le paradoxe de la sélection intra-thymique. Premièrement, la sélection positive et négative pourraient être contrôlées par des populations distinctes de cellules stromales qui livrent des signaux qualitativement différents. Par exemple, les cellules épithéliales pourraient être responsables de la sélection positive alors que les cellules dendritiques (d'origine myéloïde) pourraient être

responsables de la sélection négative. Un corollaire à cette théorie est que les cellules épithéliales expriment, dans le contexte du CMH du soi, une série de peptides utiles à la sélection. Ces peptides auraient pour but de mimer l'univers des peptides antigéniques sans pour autant être similaires à des peptides que les cellules matures pourraient rencontrer dans la périphérie. La deuxième théorie est basée sur des différences quantitatives de transmission de signaux qui pourraient être perçues par les thymocytes immatures. La sélection positive résulterait de la liaison de faible affinité entre le TCR et le CMH du soi, alors que la sélection négative résulterait d'une liaison de plus grande affinité.

2.2 Nature et rôle des interactions lympho-stromales dans la différenciation

Nous avons vu précédemment que les interactions $TCR\alpha\beta CD3$ et $CD4/8$ avec le CMH(soi)/peptide sont responsables de la sélection positive et négative. Par contre, il serait difficile d'assumer que ces interactions gouvernent tous les événements responsables des étapes de différenciation des thymocytes. L'observation que la différenciation ne peut être complétée seulement en présence de cytokines (Andrews *et al.*, 1985) suggère que d'autres molécules d'adhésion pourraient être impliquées dans la différenciation. De plus, l'observation et l'isolation de complexes lympho-stromaux variant morphologiquement (Kyewski, 1987) appuie cette hypothèse. En effet, ces observations ont démontré que les thymocytes immatures interagissent en premier lieu avec des macrophages corticaux (Ia^-) suivi d'interactions avec des cellules épithéliales corticales (Ia^+) et des TNCs puis finalement avec des cellules dendritiques médullaires (Ia^+).

L'utilisation de lignées de cellules stromales thymiques a permis de caractériser, au niveau moléculaire, les interactions lympho-stromales. De

plus, ces interactions *in vitro* donnent lieu à des réponses biologiques ayant possiblement une contrepartie *in vivo*. C'est d'ailleurs avec un système *in vitro* que la sécrétion de facteurs solubles par une lignée de cellules épithéliales corticales a été démontrée lorsqu'elles interagissent avec des thymocytes (Mizutani *et al.*, 1987).

D'autre part, il a été démontré que plusieurs de ces lignées peuvent reproduire *in vitro* des étapes spécifiques de la différenciation des lymphocytes T *in situ*. Par exemple, les travaux de Denning *et al.* ont démontré que des cultures de cellules épithéliales purifiées peuvent induire la prolifération de thymocytes immatures (Denning *et al.*, 1987) ou matures (Denning *et al.*, 1988). D'autres travaux ont clairement démontré l'importance des interactions lympho-stromales. Le rôle des interactions lympho-stromales a également été mis en évidence par le groupe de Fujiwara. Ils ont démontré que l'interaction entre une lignée de cellules stromales (MRL104.8a) et des thymocytes immatures DN permet la prolifération et la différenciation des cellules DN (Fujiwara *et al.*, 1990). Par ailleurs, les cellules de la lignée épithéliale IT 79 MTM C3 peuvent induire, en présence d'IL-2, la prolifération et la différenciation des thymocytes foetaux avec lesquels ils interagissent (Itoh *et al.*, 1988).

Outre la différenciation des cellules immatures, les résultats obtenus avec certaines lignées cellulaires suggèrent que la sélection positive et négative est également contrôlée par des interactions lympho-stromales. Les travaux subséquents de Fujiwara ont démontré que la même lignée cellulaire MRL104.8a peut induire la sélection négative en présence d'interféron- γ (Kosaka *et al.*, 1989; Matsubara *et al.*, 1990). De plus, ces interactions peuvent induire des changements morphologiques. La lignée de cellules épithéliales thymique B/c.TEC-L1 peut induire la prolifération ou l'apoptose de thymocytes syngéniques suite à la formation de TNCs *in vitro* (Hiramine

et al., 1990). Aucun de ces travaux n'a permis d'identifier clairement quelles sont les molécules d'adhésion impliquées dans les interactions étudiées.

Les travaux qui démontrent une sélection négative *in vitro* par des lignées cellulaires impliquent la participation du TCR dans ces événements. Par contre, la prolifération ou la sécrétion de cytokines résultant des interactions lympho-stromales n'est pas nécessairement reliée aux interactions TCR-dépendantes. De plus, la nécessité d'interactions via le TCR n'exclut pas la possibilité que d'autres interactions ligand-récepteur soient essentielles pour engendrer une réponse. En effet, il a été démontré que les interactions LFA-1 / ICAM-1 sont nécessaires en plus des interactions via le TCR pour que la sélection négative par les cellules dendritiques corticales soit efficace (Carlow *et al.*, 1992).

3. Le contrôle du stroma par les cellules lymphoïdes

3.1 Avant-propos

Il a été démontré, *in vitro*, que l'ajout d'interféron- γ à des cultures de cellules épithéliales module positivement l'expression du CMH de classe-II (Papiernik *et al.*, 1986; Lo *et al.*, 1986; Matsubara *et al.*, 1990). Sachant que les thymocytes constituent la source intrathymique principale d'interféron- γ (Farr *et al.*, 1989), il est probable que les thymocytes activés jouent un rôle primordial dans la régulation de l'expression de ces molécules. D'autre part, il a été démontré que la sécrétion d'IL-1 (Le *et al.*, 1987), de GM-CSF, d'IL-6 et d'IL-8 par des cellules épithéliales influence la différenciation et la prolifération des thymocytes (Le *et al.*, 1989; Galy *et al.*, 1990). La régulation de la production de ces cytokines est possiblement sous le contrôle des thymocytes qui sécrètent notamment l'IL-4 et l'interféron- γ (Galy et Spits, 1991).

Bien que les cytokines peuvent être en partie responsables de la régulation et du maintien de l'homéostasie thymique, l'absence de preuve formelle démontrant l'effet d'une incapacité à produire ou utiliser une cytokine précise ne nous permet pas d'en faire un modèle. Effectivement, l'étendue et la spécificité d'action de ces cytokines reste à démontrer *in vivo*.

L'importance des interactions lympho-stromales est apparente par l'étude *in vitro* des modèles présentés auparavant. Par contre, les modèles *in vivo* réservent souvent des surprises. En effet, le modèle de colonisation du thymus, tel que présenté auparavant, fut quelque peu ébranlé par les travaux de Shores, van Ewijk et Singer en 1991. Leurs travaux ont

démontré l'importance des thymocytes normaux dans l'implantation des cellules stromales au niveau du thymus. D'autre part, les travaux de Jacobsen, Kincade et Osmond (1992) ont démontré des interactions similaires dans la moelle osseuse.

3.2 Restauration de la médulla thymique chez les souris SCID

Les souris SCID (*severe combined immunodeficiency*) ne possèdent pas de lymphocytes T ou B normaux à cause de l'absence de réarrangements fonctionnels des gènes du TCR ou d'immunoglobulines respectivement (Bosma *et al.*, 1983; Dorshkind *et al.*, 1984).

Histologiquement, le thymus des souris SCID ne présente pas de région médullaire par l'absence de régions positives pour l'expression d'antigènes de surface reconnus par le mAb ER-TR5 (Shores *et al.*, 1991; van Vliet *et al.*, 1984). Seulement quelques cellules positives pour ER-TR5 sont éparpillées dans le thymus entier sans organisation apparente.

Les travaux de Shores *et al.* (1991) avaient comme objectif de permettre d'observer l'organisation du thymus après une injection de moelle osseuse (sans cellules T) dérivée de souris AKR à des SCID. L'apparition de thymocytes AKR TCR⁺ dans le thymus des SCIDs a été vérifiée de 4 à 10 semaines après l'injection. Chez les souris AKR → SCID, le profil de colonisation ainsi que la distribution des cellules Thy-1⁺ était presque indiscernable de celui des souris normales. Finalement, le thymus des souris AKR → SCID montre une région médullaire identique à celles de souris normales (Shores *et al.*, 1991).

Bien que les travaux aient été effectués chez les souris SCID, les auteurs proposent que les mécanismes de formation de la trame du thymus

chez les souris normales est étroitement relié à celui présenté. Sachant que l'épithélium médullaire thymique apparaît dès le 13^{ième} jour de la gestation (van Vliet *et al.*, 1985), ils proposent que les thymocytes exprimant un TCR γ/δ qui colonisent le thymus embryonnaire sont responsables de l'induction de la maturation et de l'organisation de l'épithélium médullaire thymique réactif à l'anticorps ER-TR5⁺.

3.3 Modèle de la régulation du stroma de la moelle osseuse

Le développement des précurseurs des lymphocytes B dans la moelle osseuse s'effectue en étroite association avec les cellules stromales d'une façon similaire à celles que l'on trouve dans le thymus (Osmond, 1986; Kincade *et al.*, 1989). C'est afin de mieux caractériser ces associations entre cellules stromales et lymphocytes pré-B ainsi que pour comprendre les étapes de différenciation des lymphocytes B que fut développé la LTBM (*long-term bone marrow culture*; Dexter et Lajtha *et al.*, 1974). L'utilisation du LTBM a permis de démontrer *in vitro* que les cellules stromales de la moelle osseuse sont nécessaires pour la majeure partie des étapes de différenciation des précurseurs (Kincade *et al.*, 1989). Cette nécessité est reflétée par l'étroite association observée entre les cellules stromales et les précurseurs (Kincade *et al.*, 1989).

Afin de mieux caractériser ces interactions, plusieurs clones de cellules stromales ont été développés. Ces clones peuvent supporter le développement des précurseurs jusqu'à maturité des lymphocytes B (Hunt *et al.*, 1987). Les travaux récents de Jacobsen *et al.* (1992) ont démontré l'expression *in vivo* d'un antigène membranaire (KM16) restreint aux cellules stromales. Cette molécule se retrouve à l'interface des associations entre les cellules stromales et les précurseurs de lymphocytes B.

Ce modèle, analogue à celui proposé pour les interactions lympho-stromales thymiques, suggère que le contrôle de la différenciation est orchestré d'une façon similaire pour toutes les populations de cellules effectrices du système immunitaire.

3.4 Un modèle *in vitro* permettant l'étude du contrôle du stroma

3.4.1 Avant-propos

Même si les travaux de Shores *et al.* (1991) suggèrent que les cellules TCR⁺ sont en partie responsables de la régulation de la prolifération des cellules stromales, aucune évidence moléculaire ne peut appuyer ces hypothèses. De plus, aucun modèle permettant l'étude des signaux bi-directionnels ou au moins, unidirectionnels des thymocytes vers les cellules stromales n'existe.

L'observation, par Potworowski et son équipe, en 1986, de l'adhérence spécifique de thymocytes à une lignée de cellules épithéliales médullaires n'avait rien d'inhabituel. En effet, les travaux de Patrice Hugo au sein du même laboratoire, ont démontré l'adhérence spécifique des thymocytes DP (majoritairement) à la lignée E-5 (Hugo et Potworowski, 1990). Par contre, l'observation que les thymocytes adhérents ne subissaient aucune modification phénotypique majeure, ne proliféraient pas, ou n'étaient pas sélectionnés, a soulevé certaines questions (Potworowski *et al.*, 1989). C'est alors que les travaux ont été orientés afin de déterminer, dans un premier temps, la (les) molécule(s) des cellules épithéliales responsable de cette adhérence. Un deuxième objectif était de déterminer le rôle de l'interaction.

3.4.2 Le modèle *in vitro* E-5 pour étudier les *signaux inversés*

La lignée de cellules épithéliales thymique provient d'une culture primaire de cellules stromales immortalisées par un carcinogène (Potworowski *et al.*, 1986). Elles expriment les déterminants reconnus par les anticorps ER-TR5 et ER-TR6 spécifiques aux cellules épithéliales

médullaires (van Vliet *et al.*, 1984) en plus d'exprimer la cytokératine.

Par un test d'adhérence *in vitro*, il a été démontré que les cellules E-5 peuvent interagir avec des thymocytes (Potworowski *et al.*, 1986) ainsi qu'avec des cellules de la lignée de lymphôme Ti-6 (Hugo *et al.*, 1989). Une caractérisation plus approfondie a permis de déterminer que les thymocytes qui adhèrent aux cellules sont principalement des blastes DP et qu'ils expriment faiblement le TCR $\alpha\beta$ /CD3 (Hugo et Potworowski, 1990).

Les premières observations ont indiqué que 50% des thymocytes adhérents se détachent spontanément et perdent l'habilité à adhérer de nouveau à d'autres cellules E-5. Ces résultats suggéraient que suite au contact, les thymocytes accédaient à un stade de différenciation tardif et qu'ils n'exprimaient plus le récepteur essentiel à l'interaction suite à cette différenciation (Hugo et Potworowski, 1990). Par contre, l'interaction ne démontrait pas de prolifération des thymocytes alors que la seule modification phénotypique est une faible diminution du taux d'expression de CD8 (Potworowski *et al.*, 1989). D'autre part, le groupe de Potworowski a conclu que la participation de la cellule E-5 dans des événements liés à la sélection était peu probable; l'interaction thymocytes - E-5 n'étant pas dépendante de l'haplotype du CMH de classe I (Hugo *et al.*, 1988).

Bien que les interactions thymocytes - E-5 n'entraînent pas de changements majeurs chez les thymocytes, la caractérisation préliminaire de l'interaction a révélé des différences importantes lorsque comparées aux autres modèles d'interactions lympho-stromales. En effet, les complexes observés sont formés plus efficacement à 4 °C qu'à 37 °C et ne requièrent pas de calcium ni de magnésium (Hugo et Potworowski, 1989). La molécule d'adhésion responsable de l'interaction n'avait pas d'homologie fonctionnelle à celles connues. De plus, les travaux de Hugo *et al.* ont

démontré que la molécule d'adhésion potentiellement impliquée est une glycoprotéine membranaire sensible à la trypsine dont la glycosylation est sans conséquence pour l'interaction (Potworowski *et al.*, 1986; Hugo *et al.*, 1989).

Les travaux subséquent de Clément Couture (thèse de Ph.D.) ont permis non-seulement une caractérisation biochimique de la molécule d'adhésion, mais également une caractérisation fonctionnelle de l'interaction.

La découverte d'un anticorps monoclonal bloquant l'interaction entre les thymocytes et les cellules E-5 a permis d'identifier le ligand des cellules E-5 (Couture *et al.*, 1990). L'anticorps L-43 (ou son sous-clone C3C12) permet l'immunoprécipitation d'une glycoprotéine de 23 kDa (gp23). De plus, l'immunoprécipitation effectuée en absence de détergent démontre la présence de plus d'une protéine. En effet, la co-immunoprécipitation d'un trimère a démontré l'association non-covalente de trois protéines: gp23, gp45 et gp90. Une caractérisation *in vitro* a permis d'établir que seules les gp23 et gp45 sont nécessaires à l'adhésion de thymocytes au complexe (Couture *et al.*, 1990). Ces résultats ne permettent pas d'associer un rôle, *a priori*, à la gp90 même si par analogies, un rôle dans la transduction de signaux est proposé (Couture *et al.*, 1990).

La caractérisation fonctionnelle de l'interaction a révélé des mécanismes encore inédits. En effet, pour la première fois, il a été démontré que l'interaction observée induit une phosphorylation chez les cellules stromales (Couture *et al.*, 1992). Cette phosphorylation est bloquée par l'anticorps anti-gp23 (C3C12); elle est spécifique à l'interaction via la gp23/45/90. De plus, la co-culture des cellules E-5 et des thymocytes où les contacts cellulaires sont bloqués par une fine membrane poreuse a

démontré la nécessité du contact pour la phosphorylation ainsi que l'impossibilité que seules des cytokines soit responsables du déclenchement de la phosphorylation (Couture *et al.*, 1992).

La phosphorylation suivant une co-stimulation de 30 minutes est observée principalement chez trois protéines: pp15, pp20, pp90 (Couture *et al.*, 1992). Il restait à déterminer si la pp90 ainsi observée est analogue à la gp90 du complexe. Une expérience de phosphorylation *in vitro* a permis d'établir la corrélation. Suite à une co-culture de 30 minutes de thymocytes et de cellules E-5 privées de phosphate, l'immunoprécipitation a été effectuée. Un essai kinase sur le complexe immunoprécipité a démontré que la gp90 était phosphorylée *in vitro* (Couture *et al.*, 1992). Les travaux ultérieurs ont démontré par une analyse des acides aminés phosphorylés que seuls des résidus tyrosine étaient phosphorylés. De plus, la génistéine, un inhibiteur non-spécifique des protéine tyrosine kinases, bloque entièrement la phosphorylation observée *in vitro* (Couture *et al.*, 1992).

Les dernières observations ont donc démontré l'implication d'une protéine tyrosine kinase suite à l'interaction entre les cellules E-5 et les thymocytes. Afin de vérifier si une PTK déjà connue pouvait répondre aux critères biochimiques établis par les expériences, il a été entrepris d'interroger les banques de données afin de déterminer si, par analogie, la gp90 était déjà connue. Cette recherche fut infructueuse.

4. Les protéines kinases

4.1 Avant-propos

La première protéine kinase isolée et caractérisée fut la phosphorylase kinase en 1959 (Krebs *et al.*, 1959). Les travaux de Krebs et Fischer leur ont valu le prix Nobel de Physiologie et Médecine en 1992. En 1968, la deuxième protéine kinase fut isolée par Walsh. Il avait purifié la protéine kinase cAMP-dépendante (Walsh *et al.*, 1968). Il s'agissait dans les deux cas de protéine sérine-thréonine kinases. La première protéine tyrosine kinase fut isolée par Erikson en 1979. Il s'agissait de la désormais célèbre pp60^{v-src}. La deuxième, l'EGF-R fut isolée par Cohen en 1982. Celle-ci semblait différer considérablement de la première par leurs poids moléculaires et leurs structures biochimiques.

La présence d'une activité kinase similaire chez des protéines différentes suggère l'existence d'une homologie dans la région du domaine catalytique même si la spécificité de substrat est différente. Les travaux de Barker et Dayhoff (1982) ont démontré l'existence d'une région de 300 acides aminés formant le domaine catalytique des protéines kinases qui présente une homologie très élevée.

Deux facteurs ont un impact sur le nombre de nouvelles PKs qui font leur apparition dans les banques de données. D'une part, l'intérêt purement fondamental de la biochimie des protéines kinases a suscité l'élaboration de méthodes qui permettent d'isoler des nouvelles PKs à partir d'homologies tirées de l'études des PKs connues. Ces méthodes s'appuient sur les caractéristiques suivantes: la grande spécificité de substrat permet la purification par chromatographie d'affinité; les propriétés biochimiques

communes des kinases exprimées *in vitro*; la présence d'un domaine catalytique très fortement conservé permet l'utilisation de techniques de clonage spécifiques aux familles et sous-familles; la présence de résidus phosphorylés contre lesquels des anticorps monoclonaux ont été produits. D'autre part, le rôle d'oncogène potentiel de plusieurs protéine kinases a fait des protéine kinases la famille d'enzymes probablement la plus étudiée (Cantley *et al.*, 1991; Weiss, 1993).

4.2 Régulation métabolique par les protéine kinases

Le nombre élevé de protéine kinases répertoriées reflète leur rôle central dans les mécanismes de régulation des différents systèmes cellulaires: cycle cellulaire, régulation métabolique et catabolique, signaux de transduction. La complexité des réseaux de régulation est d'autant plus grande du fait que les protéines kinases possèdent généralement plusieurs substrats. En effet, la régulation de plusieurs circuits métaboliques dépend des événements de phosphorylation et de déphosphorylation. C'est d'ailleurs pourquoi certains assument que pour chaque kinase, il y aurait une phosphatase (Hardie, 1989).

La régulation métabolique de l'activité kinase est principalement effectuée de quatre façons: premièrement, par des interactions entre kinase et phosphatase, deuxièmement, par des interactions kinase / kinase, troisièmement, par l'autophosphorylation et quatrièmement, dans le cas d'enzymes sous-unitaires, par l'interaction entre une sous-unité régulatrice et une sous-unité catalytique. Connaissant l'importance des événements liés à la transduction par les protéines kinases (Ullrich et Schlessinger, 1990), il est compréhensible que plus d'un de ces mécanismes peut intervenir dans la régulation de l'activité d'une même kinase. Par exemple, les protéine tyrosine kinases Lck et PDGF-R possèdent toutes deux une

activité d'autophosphorylation *in vitro* (Veillette *et al.*, 1989; Kazlauskas et Cooper, 1989). La phosphorylation de ces résidus module positivement ou négativement l'activité enzymatique (Bergman *et al.*, 1992; Helin *et al.*, 1991; Cooper *et al.*, 1986).

La régulation, tant positive que négative, est principalement effectuée par la phosphorylation autologue ou hétérologue de certains sites; en plus de leur activité d'autophosphorylation, Src, Lyn et Fyn peuvent s'associer à l'isoforme beta du récepteur de PDGF (PDGF-R β) *in vivo* et être phosphorylées (Kypta *et al.*, 1990). La signification ou l'importance, *in vivo*, de la phosphorylation autologue (autophosphorylation) reste encore à démontrer même si une telle activité fut clairement démontrée *in vitro*. Par exemple, pp60^{c-src} possède une activité d'autophosphorylation *in vitro* mais d'autres travaux ont démontré la possibilité qu'il en soit tout autrement *in vivo* par l'implication d'une autre protéine tyrosine kinase dans son réseau de régulation (Cooper *et al.*, 1986). Dans le cas du récepteur de l'insuline et de récepteur de l'EGF, on soupçonne que cette activité d'autophosphorylation soit importante (Bertics et Gill, 1985). Finalement, l'autophosphorylation d'un résidu tyrosine conservé (Tyr857) au sein du domaine catalytique de PDGF-R semble impliquée dans la modulation de l'activité enzymatique alors que l'autophosphorylation d'un résidu tyrosine (Tyr751) dans une région insérée dans le domaine catalytique serait responsable de moduler les interactions entre le récepteur activé et des protéines cellulaires (Kazlauskas et Cooper, 1989). Ces résultats suggèrent que l'autophosphorylation de certains résidus est un mécanisme autologue de régulation négative tout comme l'autophosphorylation sur d'autres sites semble altérer la structure tertiaire du domaine catalytique ou du moins, d'enlever une contrainte structurale empêchant l'activité catalytique (Honegger *et al.*, 1988a, 1988b).

L'importance de la régulation autologue négative et positive a été démontrée directement. L'analyse des mécanismes de régulation de l'EGF-R a permis de démontrer que le site de liaison du substrat est en compétition avec le site d'autophosphorylation (Honegger *et al.*, 1988a, 1988b). La prévalence de l'activité d'autophosphorylation chez la famille des PTKs les distingue des protéine sérine thréonine kinases. En effet, dans plusieurs cas, il semble difficile de trouver des substrats de PTKs autres que les PTKs elles-mêmes.

D'autre part, la régulation hétérologue est fréquemment observée. L'activité protéine tyrosine kinase de pp60^{c-src} est modulée par une PSTK. En effet, la sérine thréonine kinase p34^{cdc2} phosphoryle pp60^{c-src} selon les étapes du cycle cellulaire (Morgan *et al.*, 1989). De plus, l'activité de l'EGF-R est modulée négativement par la PKC (Cochet *et al.*, 1984).

4.3 Caractéristiques structurales des protéines kinases

4.3.1 Généralités

Alors qu'il est presque impossible d'imposer une classification des protéines kinases selon leur(s) substrat(s), il est recommandé par le Comité de Nomenclature de l'Union Internationale des Biochimistes de les classer selon le type de résidus qu'elles phosphorylent. En plus des deux classes mentionnées auparavant, soit les classes de kinases spécifiques aux résidus tyrosines (E.C. 2.7.11) ou aux sérines et thréonines (E.C. 2.7.10), s'ajoutent la classe des histidine kinases (E.C. 2.7.12), des cystéine kinases (E.C. 2.7.13) et des aspartyl ou glutamyl kinases (E.C. 2.7.14) (revue par Hunter, 1991). Bien que l'existence des histidine kinases et des aspartyl kinases ait été démontrée (Smith *et al.*, 1974), ces dernières ont fait l'objet de peu d'études et leurs structures demeurent inconnues à ce jour.

Les deux grands types de kinases, à cause de leur diversité, ont été regroupés en familles et sous-familles. Dans le cas des sérine thréonine kinases, le mode de régulation caractérise la famille alors que les sous-familles sont composées soit d'isoformes soit d'homologues de séquence ou d'homologues provenant d'une autre espèce. Pour les protéine tyrosine kinases, la situation est quelque peu différente. Les familles sont caractérisées par la topologie transmembranaire ou intracellulaire de la protéine. Les sous-familles sont composées de membres ayant une très grande homologie de séquence au niveau du domaine catalytique.

Il existe donc deux familles de protéine tyrosine kinases; les PTKs de type récepteur ou de type cytoplasmique. Le prototype de ce dernier est, sans aucun doute, la protéine tyrosine kinase la mieux connue, pp60^{c-src}. Nous allons d'ailleurs utiliser ce modèle pour illustrer quelques caractéristiques des protéine tyrosine kinases.

Toutes les protéine tyrosine kinases connues possèdent une homologie de séquence très élevée sur une longueur d'environ 300 acides aminés. Cette région a été définie comme étant le domaine catalytique (Hanks *et al.*, 1988). La sous-famille *src* comporte également des régions d'homologies à l'extérieur du domaine catalytique. Il ont été nommés SH-2 et SH-3 (*src homology domain 2 et 3*) (revue par Margolis, 1992; Pawson et Gish, 1992; Kock *et al.*, 1991). Les domaines non-catalytiques SH-2 ont pour rôle de lier des résidus phosphotyrosines (Matsuda *et al.*, 1990) tandis que les domaines SH-3 seraient impliqués dans la liaison des protéines G analogues à *Ras* (Cicchetti *et al.*, 1992; Clark *et al.*, 1992).

La formation de complexes stables entre les récepteurs d'hormones de croissance, tel PDGF-R β , et les protéines contenant un domaine SH-2 dépend, au préalable, de l'activation du récepteur par son ligand et d'une

autophosphorylation subséquente (Kazlauskas & Cooper, 1989; Kazlauskas *et al.*, 1990). Ces observations proviennent d'expériences démontrant que les domaines SH-2 *in vitro* peuvent se lier de façon synergique au récepteur du PDGF- β (Anderson *et al.*, 1990). Les domaines SH-3 (Mayer *et al.*, 1988) sont distincts et pourraient être impliqués, conjointement avec les domaines SH-2, dans la modulation des interactions de protéines phosphorylées avec le cytosquelette et les membranes (Rodaway *et al.*, 1990). Entre autre, il a été démontré récemment qu'un domaine SH-2 intact est nécessaire pour l'activité de *abl* *in vitro* et *in vivo* (Mayer *et al.*, 1992).

Outre les domaines SH-2 et SH-3 qui sont responsables de certaines activités biologiques, la région N-terminale distale est nécessaire pour la myristylation à la membrane (Buss et Sefton, 1985). Cette myristylation est indispensable pour l'activité biologique (Cross *et al.*, 1985; Kamps *et al.*, 1985).

4.3.3 L'évolution des protéines kinases

Le pourcentage élevé d'homologie retrouvé dans les domaines catalytiques, la présence de motifs conservés ainsi que l'existence de modes de régulation très similaires chez des protéines kinases ayant des rôles différents suggèrent une évolution de cette famille à partir de la duplication d'un gène ancestral. Effectivement, l'analyse phylogénétique développée par Fitch et Margoliash (1967), utilisée selon la méthode de Feng et Doolittle (1987) permet d'observer graphiquement les relations entre les kinases. Il est intéressant de noter que les protéines sérine-thréonine kinases sont apparues plus tôt dans l'évolution. Ainsi, on observe ni d'activité kinase ni de phosphorylation chez *E. coli*, alors que chez les levures, bien qu'une activité kinase ait été détectée (Schieven *et al.*, 1986; Levin *et al.*, 1987), aucune enzyme de cette spécificité n'a été isolée malgré

la tentative d'identifier des gènes homologues à l'aide de sondes dérivées de gènes de mammifères (Hunter, 1987).

L'apparition tardive des protéine tyrosine kinases dans l'évolution est justifiée par leur importance primordiale en tant que récepteurs et transducteurs de signaux pour les systèmes hormonaux. Cette fonction est largement inutile chez les organismes unicellulaires. Leur acquisition, chez les systèmes plus évolués, coïncide probablement avec l'apparition des interactions multicellulaires d'où l'apparition d'un rôle déterminant dans la communication intercellulaire.

4.3.4 Motifs conservés des domaines catalytiques

Le profil évolutif des protéine kinases (PK) suggère que l'homologie fonctionnelle est reflétée au niveau de la séquence peptidique et nucléotidique. L'homologie retrouvée chez les PKs est localisée dans des domaines bien précis. Comme nous l'avons discuté au préalable, trois régions sont conservées, l'importance principale étant accordée au domaine catalytique (DC) même si les domaines SH-2 et SH-3 sont importants dans la reconnaissance de certains substrats des PKs.

Les domaines catalytiques varient de 250 à 300 acides aminés, ce qui correspond à environ 30 kDa. Les limites des domaines catalytiques ont été identifiées fonctionnellement par l'analyse d'enzymes tronquées (Levinson *et al.*, 1985; Wilkerson *et al.*, 1985) ainsi que par une analyse des séquences conservées (Barker et Dayhoff, 1982; Hanks *et al.*, 1988).

Structurellement, les protéine kinases diffèrent largement dans leur régions C-terminales et N-terminales. Bien que la position du DC soit variable, il est retrouvé, dans la majorité des cas, près de la région C-

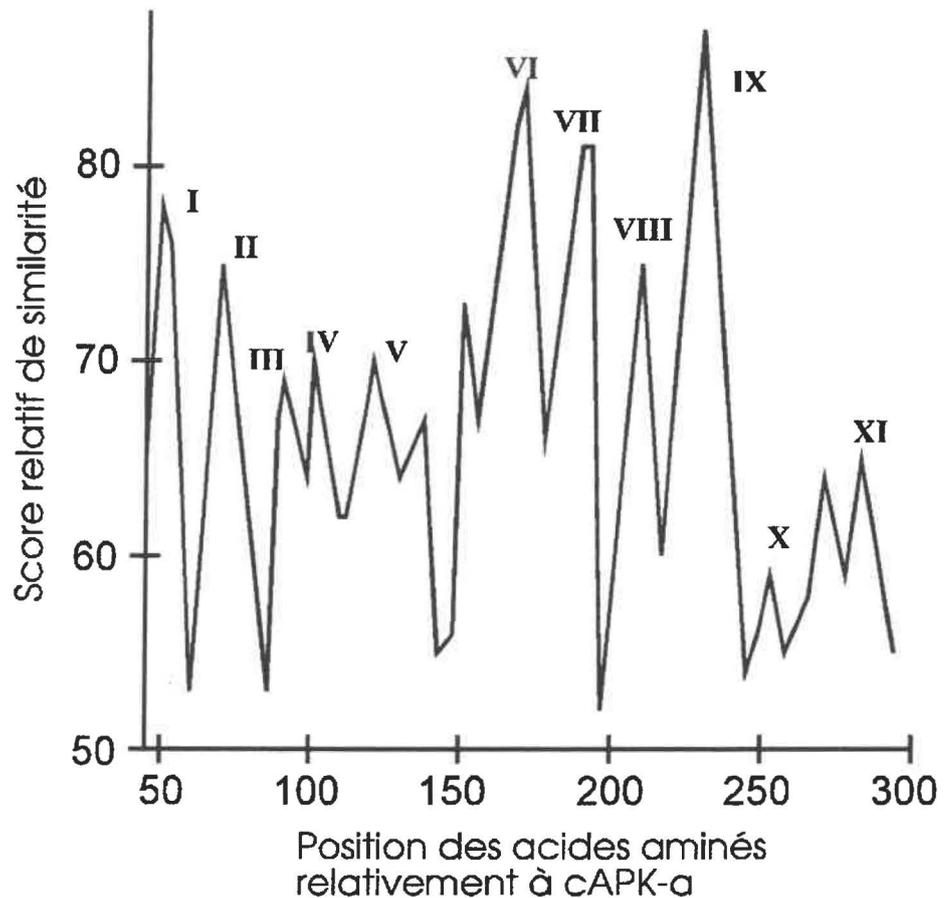


Figure 1: Profil du degré de similarité des domaines catalytiques de protéines kinases. A partir d'un alignement multiple de 65 protéines kinases représentatives des différentes sous-familles, un score de similarité relative a été déterminé selon une matrice de similarité d'acides aminés établie par Feng et Doolittle (1985). Les scores de similarité ont été calculés selon la somme de toutes les combinaisons possibles de paires entre tous les acides aminés puis exprimés en pourcentage du score le plus élevé possible (i.e. lorsque toutes les PKs possèdent un résidu identique à une position donnée). Les chiffres romains indiquent les motifs conservés. La position relative des acides aminés, basée sur la séquence de la protéine kinase cAMP-dépendante, est approximative. (repris de Hanks *et al.*, 1988)

Figure 2: Alignement multiple de protéine tyrosine kinases représentatives des principales sous-familles connues. L'alignement, généré à l'aide du logiciel MACAW, a été effectué à partir des motifs VI-IX des domaines catalytiques de onze PTKs qui définissent des sous-familles. On peut constater que les motifs comportent certaines variations de séquences mais les changements observés sont toujours sans conséquence (motif VIII). Le motif VII est important dans la liaison de l'ATP et est presque toujours conservé.

C-SRC	...----ASGMAYVErmnyVHRDLRAANilvgenlvckvaDFGlarliedneytarqqakf--PIKWTAPE	166
C-ABL	...----SSAMEYLEkknfIHRDLAARNclvgenhlvkvaDFGlsrlmtgdtytahagakf--PIKWTAPE	168
INS.R	...----ADGMAYLNakkfVHRDLAARNcmvahdftvkigDFGmtrdiyetydyrkggkgl1-PVRWMAPE	184
DFPS	...----AAGMRYLEskncIHRDLAARNclvdlehsvkisDFGmsreeeeyivsdgmkqi---PVKWTAPE	168
JAK1	...----CKGMDYLGsrqyVHRDLAARNvlvesehqvkigDFGltkaietdkeyytkddrdsPVFWYAPE	179
EPH	...----ASGMNYLSnhnyVHRDLAARNilvnqnlcckvsDFGltrllddfdgtyetqggki-PIRWTAPE	173
BEK	...----ARGMEYLAsqkcIHRDLAARNvlvtennvmkiaDFGlardinnidyykkttngrl-PVKWMAPE	193
FLG	...----ARGMEYLAskkcIHRDLAARNvlvtednvmkiaDFGlardihhidyykkttngrl-PVKWMAPE	193
EGF-R	...----AKGMNYLEdrllVHRDLAARNvlvktpqhvkitDFGlakllgaekeyhaeggkv-PIKWMAL	173
PDGFR-A	...tyqvARGMEFLAskncVHRDLAARNvllagqkivkicDFGlardimhdsnyvskgstfl-PVKWMAPE	273
tyro3	...----RCGMEYLSsrnfIHRDLAARNcmlaedmtvcvaDFGlsrkiysgdyyrqgcaskl-PVKWLALE	179

VI

VII

VIII

C-SRC	AALygrftikSDVWSFGIilltelttkgrvpyppgmvnrevldqvergyrmpcpece.....	250
C-ABL	SLAynkfsikSDVWAFGVllwei atygmspygidpsqvyellekdyrmkrpegcpe.....	252
INS.R	SLKdgvfttsSDMWSFGVlweitslaeqpyqglsneqvlkfvmdggyl dqpdcpe.....	268
DFPS	ALNfgkytslCDVWSYGIImweifskgdt pysgmtnsrareridtgymptpkstpe.....	263
JAK1	CLMqskfyiaSDVWSFGVtlhelltycdsdsspmalflkmigpthgqmtvtrlvntl.....	282
EPH	AIAhrifttaSDVWSFGIvmwevlsfgdkpygemsnqevmksiedgyrlpppvdcpa.....	257
BEK	ALFdrvythqSDVWSFGVlmweiftlgspypgpiveelfkllkeghrmdkpanctn.....	292
FLG	ALFdriythqSDVWSFGVllweiftlgspypgpveelfkllkeghrmdkpsnctn.....	277
EGF-R	SILhriythqSDVWSYGVtvwelmtfgskpydgipaseissilekgerlpqppicti.....	257
PDGFR-A	SIFdnlyttlSDVWSFGIillweifslggtpypgmmvdstfynkiksgyrmakpdhat.....	358
tyro3	SLAdnlytvhSDVWAFGVtmweimtrgqtpyagi enaeiynyliggnrlkqppcme.....	367

IX

terminale alors que la région N-terminale est vouée à un rôle de régulation ou de spécificité du substrat (e.g. *src*). Même si les domaines catalytiques des protéines kinases sont conservés au sein des deux grandes classes (PTK, PSTK), il existe des motifs conservés permettant de classer les PKs sans connaître leur activité métabolique (Hanks *et al.*, 1988).

La présence de motifs conservés a été démontrée à l'aide d'alignements multiples et graphiquement à l'aide d'un profil de similarité (figure 1). Le profil de similarité présenté à la figure 1 a été établi à partir de l'alignement multiple de 65 protéines kinases (Hanks *et al.*, 1988). Le profil permet, à première vue, d'apprécier l'existence de sous-domaines au sein du domaine catalytique. Une analyse plus détaillée montre qu'il existe onze sous-domaines (I à XI) et que certains possèdent une densité d'homologie plus élevée.

L'alternance de régions d'homologie plus ou moins élevée est caractéristique des protéines globulaires homologues (Chothia et Lesk, 1986) et suggère un rôle potentiel des sous-domaines. La conservation de ces régions reflète leur importance dans le maintien de l'intégrité des domaines catalytiques. Les sous-domaines sont potentiellement essentiels pour les fonctions enzymatiques, soit directement comme partie intégrale du site actif, soit indirectement en contribuant à la formation du site actif en imposant des contraintes structurales (structure secondaire). Les régions conservées sont probablement localisées dans des boucles responsables du rapprochement des motifs conservés. La variation ponctuelle observée dans ces domaines est sans conséquence majeure pour le maintien de l'intégrité du domaine catalytique.

L'alignement multiple des séquences de PK permet d'identifier les séquences homologues. Les domaines catalytiques possèdent des

séquences qui présentent une homologie élevée dans leur ensemble mais des régions ponctuelles de forte homologie sont également retrouvées. Le profil de similarité à la figure 1 permet d'identifier 4 régions où l'homologie est très élevée, soit les motifs VI, VII, VIII et IX. L'alignement multiple des motifs VI à IX de onze protéine tyrosine kinases montre clairement des régions d'homologie (figure 2).

Les alignements multiples permettent d'aligner des séquences qui ne sont pas parfaitement homologues. L'identité entre les séquences est calculée à l'aide de logiciels qui utilisent une matrice de remplacement d'acide aminés (figure 3). Cette matrice de remplacement attribue une valeur aux substitutions selon l'implication qu'elles peuvent avoir sur la structure secondaire. Dans certains cas, la substitution d'un acide aminé est sans conséquence pour l'activité de la protéine. En effet, la substitution d'un acide aminé par un autre de même type (e.g. hydrophile à polarité positive) n'entraîne pas de changements importants dans la structure secondaire. Par contre, la substitution d'un acide aminé par un autre dont les propriétés physico-chimiques diffèrent peut entraîner des modifications tant au niveau de l'activité enzymatique que de la spécificité. De plus, si le changement s'avérait néfaste pour l'activité enzymatique, la mutation pourrait être létale. L'alignement de séquence de la figure 2 montre que le remplacement d'un résidu valine par une isoleucine au sein du motif VI est presque sans conséquences. Alors que les scores de la matrice PAM250 varient de -8 (substitution Tyr → Cys) à +17 (Tyr conservé), le score attribué au remplacement Val → Ile est de 4.

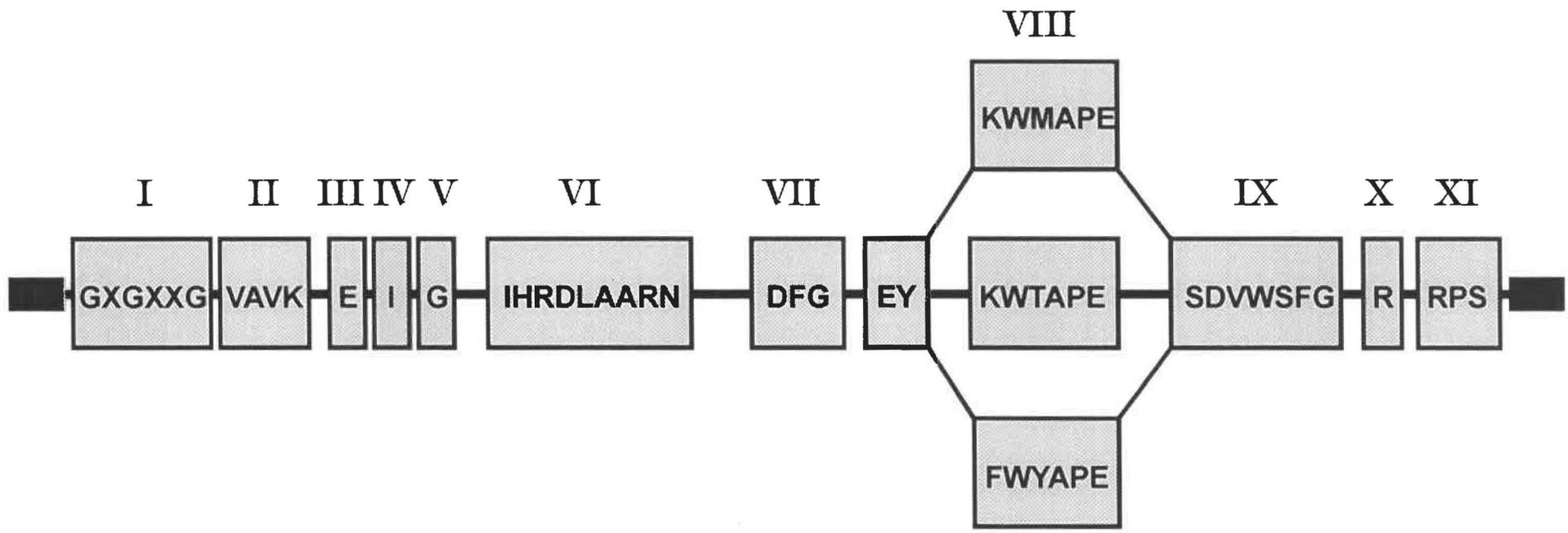
L'homologie de séquences illustrée à la figure 2 n'est pas optimale; les différentes PTKs utilisées dans l'alignement présentent des particularités associées à leur sous-famille respective et sont moins apparentées les une aux autres que les membres de la même sous-famille. Les domaines

catalytiques présentent en tout onze motifs conservés (figure 4).

Certains sous-domaines (motifs) ont un rôle direct dans la catalyse. Ce rôle a été démontré soit par mutagénèse, soit par analogie avec d'autres enzymes ayant une ou des fonctions similaires. Ainsi, le motif I -GxGxxG- (figure 4) se retrouve chez plusieurs protéines (autres que les PKs) liant des nucléotides (Wierenga et Hol, 1983) et est impliqué directement dans la liaison de celui-ci (Sternberg et Taylor, 1984).

Le résidu lysine invariant du motif II est certainement le résidu le mieux caractérisé des protéines kinases. Ce résidu est possiblement impliqué directement dans le transfert du groupement phosphate (Kamps et Sefton, 1986). Toutes les substitutions effectuées par mutagénèse dirigée, même le remplacement par un résidu arginine, ont provoqué une perte complète de l'activité catalytique (Chen *et al.*, 1987). Les motifs VI et VII ont également été impliqués dans la liaison de l'ATP. De plus, les résidus du motif VII -DFG- sont également présents chez plusieurs phosphotransférases bactériennes (Brenner, 1987). Ce triplet est d'autant plus important qu'il représente le motif ininterrompu le plus conservé du domaine catalytique (Hanks *et al.*, 1988). Le motif VIII contient le triplet Ala-Pro-Glu (-APE-) qui est une indication presque infaillible des domaines catalytiques de PKs. Des études de mutagénèse dirigée ont démontré la présence essentielle de chacun de ces résidus dans l'activité catalytique de *v-src* (Bryant et Parsons, 1983, 1984). Les motifs VI et VIII sont d'autant plus importants qu'ils permettent d'établir une classification des PKs en PTK ou PSTK. Cette spécificité de séquence à une famille suggère que ces motifs sont en partie essentiels dans la reconnaissance de l'acide aminé à phosphoryler. Aucun rôle n'a encore été proposé pour les motifs III, IV, V, IX, X et XI même si le motif IX est un des plus conservés (figure 1).

Figure 4: Domaines catalytiques de la famille des protéine tyrosine kinases. Les motifs conservés du domaine catalytique de la famille des PTKs sont représentés schématiquement. Les lignes épaisses représentent les régions hors du domaine catalytique alors que les lignes fines joignant les motifs représentent les régions non-conservées du domaine catalytique. Les motifs sont numérotés selon Hanks *et al.* (1988). Les motifs VIII sont spécifiques à des sous-familles. (Adapté de Wilks, 1991).



4.3.5 Caractéristiques moléculaires des protéine tyrosine kinases

L'indicateur le plus représentatif de la spécificité enzymatique est le motif VI. Ce motif est situé entre deux résidus invariant Asp et Asn qui sont impliquées dans la liaison de l'ATP (Brenner, 1987). Le motif -Asp-Leu-Lys-Pro-Glu-Asn- (-DLKPGN-) est caractéristique des PSTKs alors que le motif des PTKs est soit -Asp-Leu-Arg-Ala-Ala-Asn- (-DLRAAN-) pour les membres de la sous-famille *src* soit -Asp-Leu-Ala-Ala-Arg-Asn- (-DLAARN-) pour les autres (Hanks *et al.*, 1988).

La région amino-terminale du triplet -APE- du motif VIII permet également de déterminer la classe de PK. Ainsi, la séquence Gly-Thr/Ser-X-X-Tyr/Phe-X-Ala-Pro-Glu- est caractéristique des PSTKs alors que le motif -Pro-Ile/Val-Lys/Arg-Trp-Thr/Met-Ala-Pro-Glu- (-P[I/V]-[R/K]-W-[T/M]-A-P-E-) est caractéristique des PTKs. La motif VI des PTKs montre des différences marquées avec celui des PSTKs. Il a été proposé d'utiliser ce motif afin de concevoir un amorce de PCR qui permet d'amplifier spécifiquement une classe de PK.

4.3.6 Choix des amorces et amplification des domaines catalytiques des protéine tyrosine kinases

Les caractéristiques moléculaires partagées entre les protéine tyrosine kinases permet l'amplification par PCR des domaines catalytiques. À partir des alignements multiples des séquences nucléotidiques, des régions séquentielles d'homologie sont apparentes. Le motif VI, tel que présenté à la section précédente, en est un exemple parfait. En plus de présenter un niveau élevé d'homologie, il permet de cibler les PTKs ou les PSTKs si désiré. Le motif VIII présente également une séquence homologue contigüe

de longueur importante qui permet de développer une amorce qui lui est spécifique.

La sélection d'une région présentant une homologie peptidique n'implique pas que la séquence nucléotidique est la même. En effet, la sélection d'une amorce au motif VI (figure 5) entraîne l'apparition de dégénérescence à cause de la présence de substitution ou encore de l'utilisation de plus d'un codon pour un même acide aminé retrouvé chez plusieurs PTKs. Le choix de la deuxième amorce est basé sur deux critères. Le premier est que la séquence entre les deux amorces doit générer assez d'informations pour permettre de déterminer la validité de la séquence clonée ainsi que le type de PK dont il s'agit; le deuxième est que la séquence doit permettre l'amplification spécifique du plus grand nombre de PTKs. Le premier critère a sûrement incité Wilks (1989) à ne pas choisir le motif VIII alors que le choix du motif IX a sûrement été influencé par le deuxième critère que nous avons établi.

Figure 5: Alignement des motifs VI et IX des domaines catalytiques utilisés pour l'élaboration des amorces de PCR. Les séquences de huit PTKs représentatives de plusieurs sous-familles sont alignées. La région utilisée pour la synthèse des amorces est indiquée. À cause de la présence de différents acides aminés ou de l'utilisation de codons différents, les amorces incorporent une dégénérescence indiquée par: **M** ≡ A/C; **S** ≡ C/G; **R** ≡ A/G; **Y** ≡ C/T; **W** ≡ A/T; **N** ≡ A/C/G/T. Les sites de restriction sont également indiqués.

OBJECTIFS DE LA RECHERCHE

Les interactions lympho-stromales jouent un rôle central dans le maintien de l'homéostasie thymique. Comme nous l'avons discuté, ces interactions engendrent des signaux intracellulaires qui ont pour effet d'amplifier le signal initial résultant en des réponses physiologiques mesurables telles que la prolifération cellulaire ou la sécrétion de cytokines. L'étude *in vitro* de ces systèmes permet de caractériser les interactions au niveau moléculaire. Le modèle E-5/thymocytes permet de mesurer l'activation des cellules E-5 suite au contact, via le complexe gp23/45/90.

En soit, l'activation des cellules stromales est un phénomène nouvellement décrit; il a l'énorme avantage de se prêter à l'étude des signaux intracellulaires telle la phosphorylation de la gp90. Afin d'ouvrir la porte sur les événements post-phosphorylation, il est impératif de déterminer quelles sont les protéine tyrosine kinases impliquées dans le maintien de l'homéostasie thymique.

Le but du projet est donc de caractériser les transcrits spécifiques aux PTKs dans les cellules E-5. Cette information permettra de déterminer si ces transcrits codent pour de nouvelles PTKs ou encore, pour des PTK connues mais dont le rôle n'est pas encore déterminé.

MATÉRIEL & MÉTHODES

1. Les cellules E-5

1.1 Conditions de culture des cellules E-5

La lignée cellulaire E-5 est dérivée d'une culture primaire de thymus provenant de souris C57Bl/6. Ces cellules ont été immortalisées à l'aide d'un traitement par un carcinogène, le *N*-méthyl-*N'*-nitro-*N*-nitrosoguanidine (Potworowski et al., 1986). Le phénotype des cellules de la lignée E-5, déterminé par la réactivité à une batterie d'anticorps monoclonaux, est caractéristique des cellules épithéliales médullaires thymiques (Hugo *et al.*, 1988). Les cellules ont été cultivées dans un milieu contenant un mélange 1:1 de MEM-Eagle additionné de sels d'Earle (Gibco) et de milieu 199 additionné de sels de Hanks (Gibco), le tout contenant 10% de NuSerum (Collaborative Research, Lexington, MA). Les cellules ont été maintenues à 37°C dans une atmosphère contenant 5% de CO₂. Selon le nombre de cellules désirées, elles ont été cultivées dans des flacons de 25, 75 ou 150 cm².

2. Purification du complexe gp23/45/90

2.1 Préparation de la fraction membranaire de cellules E-5

Toutes les étapes de la préparation de la fraction membranaire de cellules E-5 ont été effectuées à 4°C. Brièvement, chaque lot de cellules a été préparé à partir de 60 flacons de 150 cm² où les cellules ont été cultivées jusqu'à 80% de confluence. Le milieu a été vidangé et les feuillets ont été rincés à l'aide de 10 mL de PBS froid contenant 2 mM EDTA. Un volume d'environ 2 mL par flacon a été laissé puis le volume résiduel a été décanté. Les cellules ont été incubées en présence du 2 mL résiduel de PBS/EDTA, à température ambiante, avec une agitation rotative (70 rpm) afin de les détacher de la surface du flacon. Elles ont ensuite été récoltées puis centrifugées pendant 5 min à 800 x *g*. Une culture de cellules E-5 maintenue dans les mêmes conditions (150 cm²) a été trypsinée afin de déterminer le nombre de cellules totales. Le culot cellulaire a été resuspendu à raison de 1x10⁷ cellules/mL de milieu STEM froid (Sucrose 200 mM, Tris-HCl 10 mM, MgCl₂ 2 mM, EGTA 1 mM, pH 7.4; Islam *et al.*, 1983) pour un maximum de 2x10⁹ cellules. Afin de prévenir la dégradation protéique suite à la lyse cellulaire, 50 μM de PMSF a été ajouté aux cellules resuspendues.

Avec une agitation légère, la rupture des membranes a été effectuée par cavitation sous une atmosphère d'azote, à 4°C. La bombe à pression (Parr Instruments, modèle 5248) a été maintenue à une pression constante de 650 PSI pendant 15 min. L'homogénat récolté sur glace a été centrifugé à 1 000 x *g* (rotor Beckman JA-20) pendant 10 min à 4°C, suivi d'une centrifugation à 5 000 x *g* pendant 10 min à 4°C. Le surnageant a été conservé et les débris cellulaires au culot ont été jetés. Les membranes ont été recueillies au culot par ultracentrifugation à 100 000 x *g* (rotor Beckman

60Ti) pendant 60 min à 4°C. Afin d'augmenter le rendement de la production de membranes, un deuxième surnageant obtenu à 5 000 x *g* a été ajouté au culot de membranes obtenu par ultracentrifugation. Une ultracentrifugation subséquente a permis de recueillir ces membranes en suspension et de les joindre au culot précédent. Les membranes purifiées ont été conservées pour l'étape de solubilisation.

2.2 Préparation des glycoprotéines membranaires

2.2.1 Solubilisation des membranes cellulaires

Les culots de membranes cellulaires ont été resuspendus dans un maximum de 14 mL de tampon Tris-HCl 50 mM, NaCl 150 mM, pH 8.0 (froid) à l'aide d'un homogénéisateur manuel en verre dépoli préalablement refroidi. La concentration en protéines a été déterminée par la méthode BCA (Smith *et al.*, 1985). Les membranes ainsi resuspendues ont été solubilisées à l'aide de 1-o-n-octyl- β -D-glucopyranoside 0.25% (Ammonyx, Boehringer Mannheim) en conservant un ratio détergent:protéines de 2:1 dans du tampon Tris-HCl 50mM, NaCl 150mM pH 8.0. L'homogénat de membranes a été agité délicatement pendant la nuit à 4°C. Les membranes ainsi solubilisées ont été ultracentrifugées à 100 000 x *g* pendant 60 min à 4°C afin de retirer la fraction insoluble. Généralement, la fraction insoluble représentait à peine 2% des membranes recueillies. Le surnageant a ensuite été utilisé pour la purification des glycoprotéines.

2.2.2 Purification des glycoprotéines membranaires par chromatographie sur Sepharose 4B/Lentil-Lectin

Une colonne de 20cm (C Column System, Pharmacia) contenant du Sepharose 4B auquel a été couplé du lentil-lectin (Pharmacia) a été utilisée

pour la purification des glycoprotéines. La colonne a été équilibrée avec deux volumes de tampon de solubilisation (Tris-HCl 50 mM, NaCl 150 mM, Ammonyx 0.25%, pH 8.0). L'échantillon a été recirculé à un débit d'environ 0.5 mL/min à l'aide d'une pompe péristaltique pendant 12 à 48hrs (selon le volume et le contenu en protéines). La colonne a ensuite été lavée avec le tampon de solubilisation jusqu'à ce que la densité optique soit de retour à la ligne de base. Cette étape éliminait la majorité des protéines adsorbées non-spécifiquement.

La colonne a ensuite été lavée à l'aide de 10 volumes de Tris-HCl 50 mM, NaCl 150 mM, pH 8.0 afin de se débarrasser du détergent. La colonne a été éluée à l'aide d'un tampon Tris-HCl 50 mM, NaCl 150 mM, Glucose (α -dextrose) 0.5 M, pH 8.0. Dès la détection des premières protéines, le système a été fermé pour la nuit afin de re-équilibrer la colonne. Le lendemain, la pompe a été redémarrée et les protéines ont été éluées à un débit de 0.3 mL/min jusqu'au retour à la ligne de base.

Entre les purifications, la colonne était régénérée par des lavages (10 volumes au moins) à l'aide d'un tampon Tris-HCl 50 mM, NaCl 1.0 M pH 8.5 puis re-équilibrée avec 5 volumes de tampon de solubilisation. Après quelques utilisations, le premier traitement était suivi d'un traitement similaire avec du tampon NaOAc 0.1 M, NaCl 0.5 M pH 4.5, puis ré-équilibrée. Chaque colonne pouvait être utilisée de sept à dix fois.

2.3 Co-immunoprécipitation du complexe gp23/45/90

2.3.1 Marquage des glycoprotéines membranaires à l'iode radioactif

Un aliquot de glycoprotéines membranaires a été dialysé contre du tampon (phosphate de sodium) PO_4 0.1 M, pH 7.5 et 50 μg a été marqué

à l'iode radioactif ^{125}I (ICN Biomedical) à l'aide de la méthode Iodo-gen (Salacinski et al., 1981). L'efficacité de marquage a été déterminée pour les fractions recueillies sur colonnes PD-10, à l'aide d'un compteur gamma (Beckman Gamma 7000).

2.3.2 Immunoprécipitations et analyse par SDS-PAGE

Afin de retirer le matériel pouvant s'adsorber à la protéine G, les glycoprotéines marquées ont été incubées, à trois reprises, avec de la protéine G-Sepharose (Pharmacia) pendant 60 min à 4°C avec rotation. La fraction clarifiée a ensuite été incubée avec l'anticorps (C3C12 ou contrôle idiotypique) pendant 90 min à 4°C avant d'ajouter 150 μL de protéine G-Sepharose pour une incubation d'une nuit à 4°C avec rotation. Les complexes ont été lavés à quatre reprises à l'aide de PBS froid et élués par ébullition dans du tampon d'échantillon pour électrophorèse (Tris-HCl 125 mM pH 6.8, glycérol 10%, SDS 2%, 2-mercaptoéthanol 5%). Les échantillons ont été analysés par électrophorèse sur gel de polyacrylamide selon la méthode de Laemmli (1970) en conditions réductrices. Après l'électrophorèse, les gels ont été fixés à l'aide d'une solution aqueuse de méthanol (50%) et d'acide acétique (10%) pendant 20 min. Par la suite, le gel a été rincé pendant 30 min dans une solution aqueuse de méthanol (50%) et de glycérol (10%). Après séchage, les gels étaient exposés sur pellicules Kodak XAR-2 à -70°C.

2.4 Purification du complexe gp23/45/90 par chromatographie d'affinité

2.4.1 Couplage d'anticorps au support Sepharose-4B activé au CNBr

La résine Sepharose 4B pré-activée au CNBr a été utilisée pour immobiliser les anticorps. Deux anticorps monoclonaux ont été utilisés, soit

un anticorps de rat IgG1 anti-gp23 murin et un anticorps isotypique anti-CD4 murin qui a été utilisé pour la pré-colonne. Le protocole utilisé est basé sur les recommandations du fabricant (Pharmacia). Brièvement, 10mg d'anticorps a été utilisé pour chaque mL de résine gonflée. L'anticorps a été ajouté à 5 mL de tampon de couplage (NaHCO_3 0.1 M, NaCl 0.5 M pH 8.3) par gramme de résine sèche. Suite au couplage, la résine a été lavée puis conservée dans du Tris-HCl 0.1 M, NaCl 0.5 M pH 8.0, contenant 0.02% de NaN_3 .

2.4.2 Purification par chromatographie d'affinité

Les échantillons de ^{125}I -E-5gp ont été appliqués sur la colonne de sépharose 4B-C3C12 et une étape de lavage a été effectuée à l'aide de 10 volumes de tampon Tris-HCl 20 mM, pH 7.6. Les conditions optimales d'élution, mesurées par la quantité de radioactivité libérée, ont été obtenues en présence de triéthylamine 0.1 M à pH 11.5. L'éluat a été neutralisé dans 0.05 volume de tampon PO_4 1.0 M à pH 6.8. Les fractions recueillies ont été analysées par SDS-PAGE (sect. 2.3.2).

3. Caractérisation moléculaire des cellules E-5

3.1 Purification du RNA total des cellules E-5

L'isolation et la purification des RNA totaux de cellules E-5 ont été effectuées à l'aide de la méthode de Chomczynski et Sacchi (1987). Cette méthode est basée sur l'utilisation du thiocyanate de guanidinium (GuSCN) et du 2-mercaptoéthanol lors de la lyse du feuillet cellulaire afin de produire une dénaturation rapide des protéines (Chirgwin *et al.*, 1979). Ceci a pour effet de réduire l'action des ribonucléases libres. De plus, pour limiter le taux d'hydrolyse du RNA par les ribonucléases, l'homogénéisation des tissus dans la solution de GuSCN se faisait sur la glace (Han *et al.*, 1987).

Après la lyse cellulaire, le RNA devait être sélectivement séparé du DNA contaminant, des protéines et des débris cellulaires. La séparation a été effectuée selon les coefficients de partition de chacun de ces composés en présence de phénol additionné d'un volume d'un mélange de chloroforme et d'alcool isoamylique (49:1). Le RNA se retrouvait dans la phase aqueuse alors que le DNA et les protéines se retrouvaient à l'interface et dans la phase organique.

Le matériel utilisé pour l'isolation du RNA ainsi que l'eau entrant dans la composition des tampons étaient traités au DEPC pour inactiver les RNases (Sambrook *et al.*, 1989). De préférence, des tubes en plastique stérile étaient utilisés.

Brièvement, des cellules E-5 (45 passages) contenues dans deux flacons de 250 cm² (Corning) ont été débarassées de leur milieu de culture pour être ensuite lysées directement avec 4 mL de solution dénaturante (guanidinium thiocyanate 4 M, NaCitrates 25 mM pH 7.0, 2-mercaptoéthanol 0.1 M, *N*-lauroylsarcosine 0.5%) sans être préalablement rincées. L'homogénat a été recueilli dans 12 tubes Eppendorf (1.5 mL) auxquels nous avons ajouté 0.1 volume d'acétate de sodium (2.0 M, pH 4.0) et mélangé vigoureusement. Nous avons ensuite ajouté un volume (par rapport au volume initial de solution dénaturante) de phénol (saturé d'eau) à chaque tube. Après une agitation vigoureuse, 0.2 volume de solution 49:1 de chloroforme et d'alcool isoamylique a été ajouté. La solution a été incubée pendant 15 min à 4°C. Après une centrifugation à 10 000 x g pendant 30 min (4°C), la phase aqueuse a été prélevée. Le RNA a été précipité à l'aide d'un volume d'isopropanol 100% pendant 60 min à -20°C puis centrifugé pendant 15 min à 10 000 x g à 4°C. Le culot de RNA a été resuspendu dans 0.3 volume de solution dénaturante et précipité avec un volume égal d'isopropanol 100% pendant 60 min à -20°C, ceci suivi d'une autre centrifugation. Le culot a été resuspendu dans 500 µL d'éthanol 75%, vortexé, puis incubé pendant 15 à 20 min à température ambiante afin de permettre la solubilisation du GuSCN résiduel. Dans certains cas, le RNA a été chauffé jusqu'à 45-50°C pour accélérer le processus de solubilisation. La solution a finalement été centrifugée pendant 10 min à température ambiante à 10 000 x g puis le culot séché au SpeedVac pendant 5 min. Le culot a été resuspendu dans 25-40 µL d'eau DEPC.

La qualité du RNA a été déterminée par spectrophotométrie en diluant 200 fois un aliquot. Le spectre d'absorbance entre 200 et 300 nm nous a permis d'établir le rapport 260/230 qui est un indicateur de la présence GuSCN résiduel ainsi que le rapport 260/280 qui reflète la pureté du RNA.

3.2.2.1 Synthèse à l'aide de MuLV-RT

Généralement, 0.2 à 3.0 μg de RNA ont été utilisés pour la réaction de transcription inverse. Cette réaction a été effectuée dans 20 μL de tampon PCR (50 mM KCl, 10 mM Tris-HCl, 1.5 mM MgCl_2 , pH 8.3), contenant 1 mM dNTPs (Pharmacia), 50 pmole de l'amorce "aval", 10 U/mL RNAguard (Pharmacia) et 20 U de MuLV transcriptase inverse (Pharmacia). Ce tampon s'est avéré tout aussi efficace que le tampon suggéré pour la MuLV-RT et facilite les étapes subséquentes d'amplification par PCR. Avant d'ajouter l'enzyme à la réaction, le tube était chauffé 30 sec à 94°C, 7 min à 65°C et refroidi rapidement sur la glace pendant 2 min. Ceci dans le but de dénaturer les structures secondaires possibles du RNA. Après l'addition de l'enzyme, la solution était incubée 60 min à 42°C.

3.2.2.2 Synthèse à l'aide de Superscript™ RT

Cette méthode utilise la transcriptase inverse Superscript™ (Gibco). Cette enzyme a été modifiée par génie génétique afin d'éliminer l'activité RNase H intrinsèque aux RT conventionnelles. La dégradation des mRNAs est ainsi prévenue et la synthèse permet d'obtenir des cDNAs de pleines longueurs en plus grand nombre. De plus, cette enzyme n'étant inhibée ni par la présence de RNA ribosomal ni par celle de RNA de transfert, son utilisation avec du RNA total permet d'obtenir une synthèse efficace du premier brin de cDNA.

Brièvement, 3 à 5 μg de RNA total de cellules E-5 a été utilisé pour les synthèses de cDNA. Le RNA a été ajouté à de l'eau-DEPC pour un volume total de 13 μL auquel nous avons ajouté 1 μL d'amorces. Différentes amorces ont été utilisées (tableau II). Lors de l'utilisation d'amorces spécifiques pour la synthèse de cDNA, la concentration joue un

rôle peu important d'où l'utilisation d'un volume fixe de 1 μ L.

Le mélange RNA-amorces a été incubé à 70°C pendant 10 min puis rapidement refroidi sur la glace pendant 1 min avant d'y ajouter 2 μ L de tampon de synthèse 10x (Tris-HCl 200 mM (pH 8.4), KCl 500 mM, MgCl₂ 25 mM, BSA 1 mg/mL), 1 μ L de mélange de dNTPs 10 mM, 2 μ L de DTT 0.1 M et 1 μ L d'enzyme Superscript RT (200 U/ μ L). Le mélange réactionnel a été incubé à 42°C pendant 60 min et la réaction a été arrêtée par une incubation de 5 min à 90°C et refroidie sur de la glace pendant 2 min. Afin de minimiser l'interférence causée par la présence du RNA durant les étapes d'amplification par PCR, la réaction de cDNA a été traitée avec 2 unités de RNase H pendant 20 min à 37°C.

Oligonucléotide	Concentration	Provenance
Oligo dT	0.5 μ g/ μ L	Gibco/BRL
6-mers aléatoires	50 ng/ μ L	Gibco/BRL
RNA-tag	127 pmol/ μ L	IAF
PTK-c9R	124 pmol/ μ L	IAF
PTK-c9Rb	153 pmol/ μ L	U. de Calgary

Tableau II: Concentration des amorces utilisées pour la synthèse du cDNA

3.2.3 Amplification par PCR des motifs VI-IX des domaines catalytiques des protéine tyrosine kinases

Généralement, 1 μ L résultant de la réaction de synthèse de cDNA a été utilisé pour les réactions de PCR. La méthode d'amplification utilisée consiste en une modification du "hot-start" (D'aquila *et al.*, 1991) que nous avons élaborée

dans notre laboratoire (Hébert *et al.* 1993; appendice I). Au cDNA ont été ajoutés les dNTPs, le MgCl₂ et le tampon de PCR. À ce mélange, une pastille de paraffine a été ajoutée comme barrière d'évaporation. Le tout a été incubé à 95°C pendant 5 min puis à 75°C de 8 à 10 min avant l'ajout du mélange réactionnel contenant le tampon de PCR, les amorces et l'enzyme *Taq* polymérase. Les conditions finales de l'amplification étaient: tampon de PCR 1X (KCl 50 mM; Tris-HCl 10 mM; pH 9.0; 0.1% Triton X-100, Bio/Can Scientific), MgCl₂ 3.0 mM, 1 pmol/μL (volume final) de l'amorce "amont" et "aval" puis 0.5 U de *Taq* polymérase (Bio/Can Scientific).

L'amplification par PCR a été effectuée par une incubation consistant en 4 cycles préliminaires suivis de 35 cycles et de 10 min d'élongation dans un thermocycleur programmable (Perkin Elmer Cetus). Les quatre cycles préliminaires consistaient en une amplification à faible température d'appariement (4 cycles de 1 minute à 94°C, 1 minute à 37°C et 1 minute à 72°C) avant l'amplification proprement dite. L'étape d'élongation permet d'avoir des amplicons de pleine longueur afin d'augmenter l'efficacité de clonage. Les temps requis pour passer d'un plateau de température à un autre ont également été contrôlés (figure 6).

Lors des cycles préliminaires, la région complémentaire entre les amorces et le DNA représentait seulement 14 et 17 nucléotides pour PTK-c6F et PTK-c9R respectivement. La présence de dégénérescence ainsi que du site de restriction et des nucléotides de stabilisation en position 5' des sites de restriction défavorise la stabilité des duplexes durant les premiers cycles (Compton, 1990). L'incorporation de quatre cycles préliminaires d'amplification a permis de produire quelques copies d'amplicons ayant des extrémités 100% complémentaires aux amorces. Ces amplicons servent par la suite de matrice pour l'amplification de 30 cycles à une température d'hybridation de 55°C (Compton, 1990). Il a été démontré que de telles conditions favorisent la stabilité des duplexes amorces-DNA et par le fait même augmente l'efficacité du PCR (Mack et Sninsky, 1988).

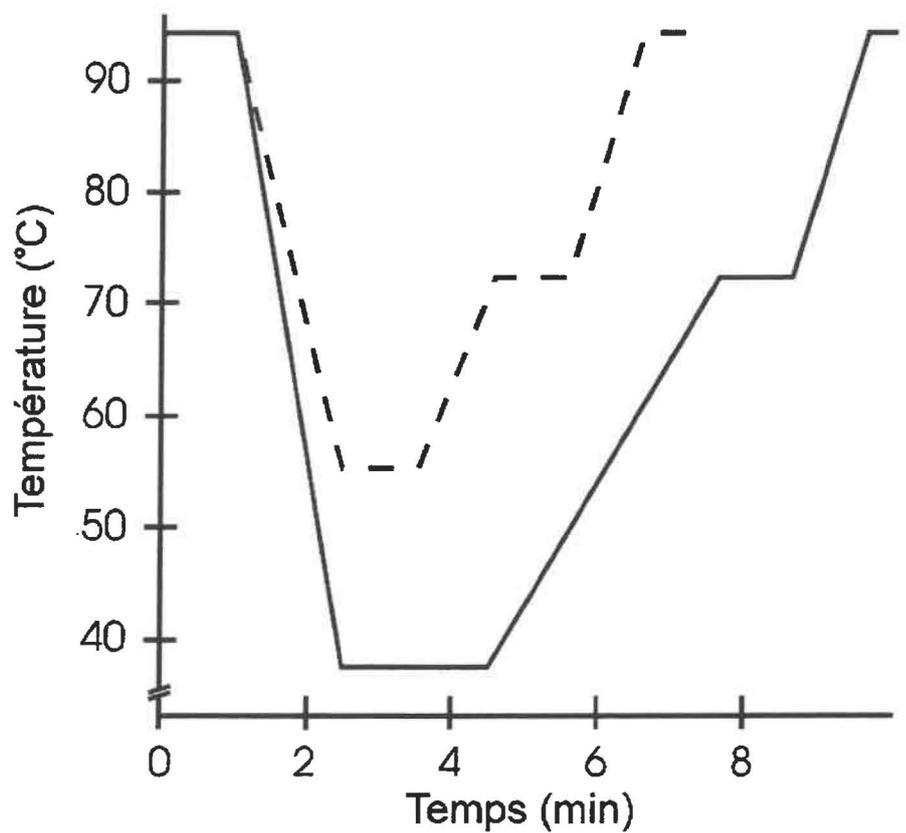


Figure 6: Cycles de PCR pour l'amplification des domaines catalytiques des PTKs. (—) les quatre premiers cycles de PCR sont effectués à 37°C afin de maximiser la stabilité de l'appariement; (----) la température d'appariement est élevée à 55°C pour les 30 cycles suivants.

Les produits de l'amplification par PCR étaient analysés par électrophorèse sur gel d'agarose (0.8 à 1.2%) tel que décrit ci-après.

3.3 Analyse sur gel d'agarose

Les produits de PCR, le vecteur (pBluescript II KS+), ainsi que les recombinants ont été analysés par électrophorèse sur gel d'agarose dans une solution de TBE 1X (Tris-OAc 89 mM, acide borique 89 mM, EDTA 2 mM; Sambrook *et al.*, 1989).

3.4 Stratégies de clonage

3.4.1 Vecteurs utilisés

Le phagemide pBluescript II KS+ a été utilisé pour le clonage des motifs VI - IX des domaines catalytiques de PTKs (Alting-Meese et Short, 1989).

3.4.2 Clonage des motifs VI - IX des domaines catalytiques de protéine tyrosine kinases

3.4.2.1 Préparation des produits de PCR

L'incorporation de sites de restriction dans les amorces a grandement facilité le clonage des amplicons. Les produits de PCR étaient d'abord débarrassés des amorces libres, des nucléotides libres et des dimères d'amorces par chromatographie sur une colonne "PrimeErase Quick" (Stratagene). Le DNA était libéré de la colonne par l'ajout de 500 μ L de tampon d'élution. On récupérait ensuite le DNA par une précipitation avec 2 volumes d'éthanol en présence de 40 μ L de polyacrylamide linéaire 0.25% pour aider la précipitation (Gaillard et Strauss, 1990). Une fois resuspendu dans 50 μ L, le DNA ainsi que 2 μ g de phagemide

pBluescript II KS+ étaient digérés par les enzymes de restriction. Les enzymes étaient éliminées par l'addition de 7 μ L de résine "Strataclean" (Stratagene). Cette résine a la particularité de lier les protéines en solution. Cette extraction était répétée deux autres fois avant de procéder à la ligation.

3.4.2.2 Préparation du vecteur

Le vecteur pBluescript II KS+ a été préparé de la façon suivante: le phagemide (5 μ g) a d'abord été linéarisé par une digestion avec l'enzyme *Cla*I (10 U) dans un volume final de 20 μ L, puis par l'enzyme *Bam*HI (10 U) dans les conditions recommandées (Gibco/BRL).

Afin de réduire le bruit de fond dû à la recircularisation du phagemide digéré uniquement par une enzyme, le phagemide linéarisé a été traité par la phosphatase alcaline de veau (CIP) pour éliminer le phosphate terminal en position 5'. Pour ce faire, les 5 μ g de phagemide digérés précédemment ont été resuspendus dans un volume final de 50 μ L de tampon One-Phor-All-PLUS (Tris-OAc 10 mM, MgOAc 10 mM, KOAc 50 mM et CIP 0.1U; Pharmacia). Ce mélange était incubé à 37°C pour une période de 30min. L'enzyme était ensuite inactivée par une incubation de 15 min à 85°C. Suite aux digestions et à la déphosphorylation, les vecteurs digérés ont été traités à la résine Strataclean (Stratagene) afin de retirer toute trace de contaminant protéique. De cette préparation, 2.5 μ g de vecteur ont été utilisés pour le clonage des produits de PCR.

3.4.2.3 Ligation

Les ligations des différents fragments obtenus après digestion par les endonucléases de restriction ont été effectuées selon une modification de la méthode de ligation classique décrite par Sambrook et al. (1989). Le produit de PCR ou le vecteur digéré ont été combinés dans des rapports molaires

vecteur:insert de 1:3 et 1:5. La ligation a été effectuée dans du tampon de ligation (Tris-HCl 20 mM, EDTA 1 mM, dithioérythritol 5 mM, KCl 60 mM, glycérol 50% (v/v), pH 7.6; Boehringer Mannheim) en présence de 2 U de T4 DNA ligase pendant 6 à 10 heures à 10°C.

3.5 Transformation des cellules bactériennes

3.5.1 Souche bactérienne utilisée

Pour les transformations, la souche d'*Escherichia coli* XL1-Blue a été utilisée. Cette souche produit moins de recombinaison que les souches conventionnelles comme JM-109 et JM-103 (Bullock *et al.*, 1987). Elle possède le génotype suivants :

XrecA- (*rec A1*, *lac^c*, *end A1*, *gyr A96*, *thi*, *hsd R17*, *sup E44*, *rel A1*, {F' *pro AB*, *lac I^s*, *lac ZΔM15*, Tn 10})

3.5.2 Préparation de bactéries compétentes

Pour rendre les bactéries susceptibles à la transformation par le DNA des différents vecteurs, la souche d'*Escherichia coli* XL-1 Blue a été traitée au CaCl₂ (Hanahan, 1983; Sambrook *et al.*, 1989). Généralement, 100 mL de bouillon LB (10 g de Bactotryptone, 5 g d'extrait de levure et 10 g de NaCl par litre de H₂O) étaientensemencés avec un mL de culture de 16h. La culture bactérienne était placée dans un agitateur orbital (Queue Orbital) à 37°C avec une agitation de 250 rpm. Quand la culture atteignait une D.O._{550nm} de 0.3 à 0.4, les bactéries étaient récoltées par une centrifugation de 5 min à 2 000 x *g*. Une fois le surnageant décanté, le culot de bactéries était resuspendu dans 50 mL d'une solution glacée de CaCl₂ à 50 mM. Les bactéries en suspension étaient conservées sur glace pour une période de 30 min, recentrifugées et resuspendues dans 10 mL de CaCl₂. Pour la transformation, 10 ng de DNA provenant du mélange de ligation

étaient déposés dans un tube de polypropylène (Falcon) préalablement refroidi. Après l'addition de 200 μL de bactéries compétentes, les tubes étaient conservés sur glace pendant 30 à 40 min. Pour permettre au DNA de pénétrer, les bactéries étaient alors soumises à un choc thermique de 60 secondes à 42°C. La régénération de la paroi cellulaire se faisait ensuite par l'addition d'un mL de bouillon 2x YT et d'une incubation de 60 min à 37°C. Les bactéries transformées (75 μL) étaient alors étalées sur une gélose 2x YT ou LB (12 g de bacto-agar par litre de bouillon 2x YT ou LB selon le cas) contenant 100 $\mu\text{g}/\text{mL}$ d'ampicilline, 3 $\mu\text{L}/\text{mL}$ d'IPTG à 100 mM et 13 $\mu\text{L}/\text{mL}$ d'une solution de X-Gal à 2% dans du DMF. Les géloses étaient incubées à 37°C en position inversée pendant la nuit.

3.5.3 Sélection des clones

Grâce à la présence du gène de résistance à l'ampicilline chez le vecteur, seules les bactéries transformées croissent sur des géloses contenant cet antibiotique. Le système de sélection des clones repose sur la présence fonctionnelle de l'enzyme β -galactosidase. En présence d'un inducteur de l'opéron *lac* comme l'IPTG, les bactéries *E. coli* transformées par un vecteur de la famille pUC produiront une enzyme β -galactosidase fonctionnelle grâce à la complémentation de type alpha. De telles cellules hydrolyseront le substrat X-Gal pour ainsi produire un chromophore bleu, le bromochloro-indole d'où la coloration des colonies résultantes. L'insertion d'un fragment de DNA dans un site approprié du vecteur interfère, normalement, avec la production de β -galactosidase dans la cellule transformée d'où l'obtention de colonies blanches. Les colonies blanches ont été récoltées etensemencées dans du bouillon LB ou 2xYT contenant 100 $\mu\text{g}/\text{mL}$ d'ampicilline. Après une incubation à 37°C pour la nuit, les cultures étaient analysées par mini-préparation par PCR ou conservées à -70°C en présence de 20% de glycérol.

3.5.4 Mini-préparations par PCR

Deux variations ont été utilisées dans le choix de la source de matériel à amplifier. La première utilise directement les colonies d'intérêt alors que la deuxième utilise 1 μL de bouillon de culture.

Pour la première, chacune des colonies d'intérêt a été mélangée à 50 μL de NaCl 0.85% dans les puits d'un plateau à ELISA stérile (fond conique). Les réactions de PCR ont été effectuées en utilisant 5 μL de ce mélange dans les conditions décrites plus bas. La deuxième méthode consistait à inoculer 3 mL de milieu 2x YT avec les colonies potentiellement recombinantes, à laisser croître pour la nuit puis à utiliser 1 μL de bouillon de culture pour le PCR. Cette deuxième méthode a été utilisée lorsque les colonies étaient de petite taille.

Dans les deux cas, le protocole de PCR a été légèrement adapté au criblage. Les réactions ont été effectuées en présence de tampon d'amplification 1X (section 3.2.3), de 1 pmol/ μL d'amorces amont et aval, de MgCl_2 3.2 mM et de 0.5 U de *Taq* polymérase. À un volume de 25 μL de tampon de réaction (avec l'enzyme), l'échantillon (5 μL de mélange colonie/NaCl ou 1 μL de culture) a été ajouté. La programmation du thermo-cycleur, appelé "turbo-PCR", consiste en une première période de dénaturation de 2 min à 94°C suivit de 35 cycles où chaque plateau de température dure 1 min. Les temps inter-plateaux sont de durée minimum. Les conditions sont les suivantes: dénaturation à 94°C; appariement à 40°C; élongation à 72°C .

Les amplicons ont été séparés sur gels d'agarose de 15 cm x 10 cm tel que décrit à la section 3.3.

3.5.5 Mini-préparation de plasmides (Magic Mini-Preps™)

Pour l'analyse des recombinants par séquençage, le DNA des plasmides était extrait de mini-préparations selon une modification commerciale (Magic Mini-Preps™; Promega) de la méthode de lyse alcaline décrite par Birnboim et Doly (1977) et modifiée par Sambrook *et al.* (1989). Brièvement, on inoculait 3 mL de milieu 2x YT, contenant 100 µg/mL d'ampicilline, avec une colonie bactérienne isolée. Ces tubes étaient incubés pendant la nuit à 37°C avec une agitation de 250 rpm. Le lendemain, on transvidait 1,5 mL de la culture dans des microtubes de type Eppendorf. Les tubes étaient ensuite centrifugés deux minutes dans une microcentrifugeuse. Le surnageant était aspiré et un deuxième volume de culture était ajouté au tube puis re-centrifugé. Le culot a été resuspendu dans 200 µL de la solution de "resuspension" (EDTA 10 mM, Tris-HCl 25 mM pH 8.0). La lyse était effectuée avec 200 µL de la solution de lyse (NaOH 0.2 N, SDS 1%). Les tubes étaient agités par inversions rapides jusqu'à éclaircissement de la solution. Ensuite, 200 µL de la solution de neutralisation (KOAc 2.55 M, pH 4.8) étaient ajoutés et les tubes étaient agités vigoureusement. Une centrifugation de 10 min permettait de se débarrasser des débris cellulaires et le surnageant était transféré dans un nouveau tube. À ce tube a été ajouté 1 mL de "résine magique" (*Magic™ resin*; Promega). Le contenu du tube a été transféré dans une seringue de 2 mL à laquelle avait été ajoutée une mini-colonne. La résine a été poussée dans la colonne puis lavée avec 2 mL de solution de lavage (EtOH 50%, NaCl 200 mM, Tris-HCl 20 mM pH 7.5, EDTA 5 mM). Une première centrifugation de 20 sec, à l'aide d'une microfuge, a été effectuée pour éliminer le surplus de solution de lavage emprisonné dans la résine. Par la suite, une deuxième centrifugation a été effectuée en ajoutant 50 µL d'eau distillée à la colonne pour l'élution. Le DNA plasmidique ainsi recueilli a été congelé à -20°C jusqu'au séquençage.

4. Séquençage

4.1 Liste des amorces pour le séquençage

Pour les réactions de séquençage nous avons principalement utilisé les amorces commerciales. L'amorce universelle de séquençage (5' GTAAAACGACG-GCCAGT 3') a été obtenue de Pharmacia, tandis que l'amorce inverse de séquençage (5' AACAGCTATGACCATG 3') a été synthétisée. Toutes les amorces ont été synthétisées à l'aide d'un "Gene Assembler DNA Synthesizer" de Pharmacia et la concentration des oligonucléotides a été déterminée par spectrophotométrie.

4.2 Réactions de séquençage manuel

Le séquençage des produits de PCR clonés a été réalisé selon la méthode "dideoxy" décrite par Sanger *et al.* (1977) en utilisant deux trousse commerciales de séquençage. Les trousse, "Sequenase sequencing kit" (USB) et le "T7 Sequencing kit" (Pharmacia) étaient utilisées selon les recommandations du fabricant. La réaction d'appariement de l'amorce à la matrice se faisait dans 10 μL d'une solution composée de Tris-HCl 40 mM, pH 7.5, MgCl_2 10 mM et NaCl 50 mM à laquelle ont été ajoutés 5 μg de DNA et 2.0 pmol d'amorce. Le mélange était chauffé à 37°C pendant 20 à 30 min et refroidi lentement, sur une période de 30 min, à la température ambiante. Aux 10 μL de solution contenant le complexe matrice-amorce, on ajoutait 1 μL de [α - ^{35}S]dATP (Amersham, 37 MBq, 1 mCi, 1000 Ci/mmol), 1 μL de 0.1 M DTT, 2 μL d'une solution contenant 1.5 μM de chacun des nucléotides (dGTP, dCTP, dTTP) et 3 U de polymérase. Le mélange était incubé à la température ambiante 5 à 10 min. On répartissait ensuite 4.5 μL de ce mélange dans 4 tubes identifiés G, A, T, et C. Ces tubes contenaient 2.5 μL de mélange de terminaison. Chacun des mélanges de terminaison contenait 80 μM

de chaque déoxynucléotide (dATP, dGTP, dCTP, dTTP) et 8 μM d'un des quatre didéoxynucléotides (ddATP, ddGTP, ddCTP, ou ddTTP). Le tube identifié "G" contenait par exemple 80 μM dGTP, 80 μM dATP, 80 μM dCTP, 80 μM dTTP et 8 μM ddGTP. Ces mélanges réactionnels étaient incubés à 37°C environ 5 min. La réaction de polymérisation était ensuite stoppée par l'addition de 4 μL de solution d'arrêt (95% formamide, 20 mM EDTA, 0.05% bromophénol bleu, 0.05% xylène cyanol FF). Ces mélanges réactionnels pouvaient être conservés à -20°C jusqu'au moment de l'électrophorèse. Pour résoudre le problème des structures secondaires associées au séquençage de certaines matrices bicaténares, nous avons, à quelques reprises, remplacé le dGTP par le 7-deaza dGTP ou dITP.

4.3 Électrophorèse de séquence

Toutes les électrophorèses de séquençage conventionnel ont été réalisées sur des gels de 35 x 42.5 cm, d'une épaisseur non-uniforme de 0.2 à 0.8 mm à l'aide d'un appareil IBI modèle STS 45. Parce que nous avons utilisé le ^{35}S comme isotope radioactif, les gels de séquence devaient être séchés pour augmenter la sensibilité. Nous avons choisi de sécher le gel directement sur la vitre après l'électrophorèse. Pour permettre l'adhérence du gel, la plus grande des deux vitres était pré-traitée au Silane A174 (Pharmacia) selon les indications du manufacturier. Pour être enlevée plus facilement, la petite vitre était vaporisée d'anti-adhésif végétal de type PAM et l'excès essuyé à l'aide d'un coton à fromage juste avant la préparation du montage. A cause de la faible épaisseur (0.2 mm) de la partie supérieure du gel, il était très difficile de couler le mélange d'acrylamide de la façon conventionnelle. Une méthode alternative pour couler les gels a déjà été mise au point dans le laboratoire de Peter Tijssen. On fixait d'abord, à l'aide de ruban adhésif, le long de la plaque de verre traitée au Silane les deux espaceurs en biseau (0.2 à 0.8 mm) et au bas de la vitre un espaceur uniforme de 0.8 mm. On déposait ensuite la petite vitre graissée de façon à la superposer sur environ 2 cm du bas de la grande vitre. Une partie du mélange d'acrylamide était alors déposée

sur la grande vitre et on glissait la petite vitre vers le haut. On ajoutait du mélange au fur et à mesure que la vitre était glissée. Un puits était formé en insérant un peigne (0.2 mm, 64 puits, dents de requin) en position inversée. Le montage était maintenu par 6 pince-notes jusqu'à la polymérisation complète de l'acrylamide (30 à 60 min). Après la polymérisation, le peigne et l'espaceur du bas ont été enlevés. Les puits ont été rincés abondamment pour enlever toute trace de polyacrylamide et d'urée puis le peigne a été réinséré dans la position adéquate. Le gel a été soumis à une pré-électrophorèse de 30 min pour augmenter sa température. Après un dernier rinçage des puits, on déposait 2 μ L des réactions G,A,T,C (dans l'ordre) dans quatre puits contigus. L'électrophorèse a été effectuée à voltage constant de 1 800 volts (49 mA, 60 W) et a été arrêtée 20 min après que le bleu de bromophénol ait atteint le bas du gel (2,5 heures). La petite vitre graissée, le peigne et les espaceurs étaient enlevés à l'aide d'une spatule. Le gel sur la grande vitre était fixé pendant 10 min dans une solution de 10% méthanol, 10% acide acétique et rincé 2 fois 10 min avec 2 litres d'eau. On séchait ensuite le gel à l'aide de deux séchoirs à cheveux (30-60 min) et on déposait directement sur ce dernier un film pour autoradiographie de type X-Omat RP (Kodak).

De façon générale les gels de séquence étaient constitués d'un mélange de polyacrylamide de 19:1 (acrylamide:bisacrylamide) dans du tampon TBE. Pour un gel standard on préparait 150 mL d'une solution de 5% acrylamide, 7 M urée, 5 μ L/mL d'ammonium persulfate 10% et 1 μ L/mL de TEMED. Ce type de gel permettait d'identifier, après une électrophorèse de 2,5 heures, environ 275 bases.

4.4 Séquençage par la méthode automatique: A.L.F. (*automated light fluorescence*)

Afin d'augmenter la rapidité de l'analyse des clones, plusieurs gels de séquence ont été effectués à l'aide d'un séquenceur automatique: A.L.F. (*automated light fluorescence*; A.L.F. Sequencer, Pharmacia). De plus, le séquenceur automatique permet de résoudre certains problèmes de compression observés sur les gels conventionnels par le remplacement du dGTP par du deaza-dGTP dans la trousse AutoRead™ (Pharmacia).

Toutes les réactions de séquençage ont été effectuées à l'aide de la trousse AutoRead™ selon les directives du fabricant. Brièvement, 5 μg de DNA ont été ajoutés à 5 pmoles d'amorces fluorescentes et 1.5 μL de NaOH 1 N dans un volume final de 11.5 μL . Après une incubation de 5 min à 65°C, les tubes ont été transférés à 37°C; 1.5 μL de HCl 1 N et 2 μL de tampon d'appariement ont été ajoutés (Pharmacia). Les tubes ont été incubés à 37°C pour une période de 10 min, centrifugés brièvement puis laissés à la température de la pièce pour 10 min. Après cette incubation, la réaction a été démarrée par l'ajout de 1 μL de tampon d'extension (Pharmacia), de 3.5 μL de DMSO et de 2 μL de T7 DNA Polymérase (4 U/ μL). Un volume de 5.4 μL de cette réaction a été réparti dans 4 tubes identifiés G, A, T, et C. Ces tubes contenaient 3 μL de mélange de terminaison. Chacun des mélanges de terminaison contenait chaque déoxynucléotide (dATP, dCTP, dTTP, c⁷dGTP) et un des quatre didéoxynucléotides (ddATP, ddGTP, ddCTP, ou ddTTP). Par exemple, le tube identifié "G" contenait du ddGTP ainsi que les quatre dNTPs (à des concentrations non-divulguées par le fabricant). Ces réactions étaient incubées à 37°C pendant 5 min. La réaction de polymérisation était ensuite stoppée par l'addition de 6 μL de solution d'arrêt (95% formamide, 20 mM EDTA, 0.05% xylène cyanol FF). Ces mélanges réactionnels pouvaient être conservés à -20°C jusqu'au moment de l'électrophorèse.

Les gels d'électrophorèse ont été effectués selon les conditions établies par Pharmacia. Afin d'assurer le maximum de performance du séquenceur automatique, nous avons utilisé des mélanges de gels préparés "ReadyMix™ Gel" (Pharmacia). Selon notre expérience, l'utilisation de gels Pharmacia augmente la performance du A.L.F.

Deux types de gels d'électrophorèse ont été utilisés, soit des gels d'épaisseur conventionnelle soit des gels minces. Les résultats obtenus dans le laboratoire montrent que les gels conventionnels permettent de lire des fragments de plus grande taille alors que l'avantage principal des gels minces est la diminution du temps requis pour effectuer la migration des séquences (tableau III).

Paramètres	Gels conventionnels	Gels minces
épaisseur (mm)	0.5	0.3
% acrylamide-bis (19:1)	6.0	6.0
urée (M)	7.0	7.0
tampon TBE	0.6x	0.6x
voltage (V)	1500	1500
courant (mA)	45	45
puissance (W)	34	45
température (°C)	48	48
puissance du laser (mW)	3	5
échantillonnage (sec ⁻¹)	2	1.25
temps d'électrophorèse	5.5 heures	3.5 heures
nombre de bases lues	450-500+	350+

Tableau III: Conditions d'électrophorèse pour le séquençage automatique A.L.F. Les conditions des gels conventionnels et des gels minces sont comparées.

5. Analyse des séquences obtenues

5.1 *Protein Kinase Catalytic Domain Database (PKCDD)*

Cette banque de données est maintenue par Steve Hanks et Anne Marie Quinn au Salk Institute for Biological Studies. Elle contient les séquences en acides aminés des domaines catalytiques de tous les membres de la famille des protéine kinases. Les nouvelles séquences sont acheminées directement par courrier électronique. Les séquences équivalentes retrouvées chez plus d'une espèce d'invertébrés ne sont pas toutes répertoriées. Seulement une séquence équivalente est répertoriée pour chaque groupe d'équivalence.

La banque est mise à jour mensuellement et est accessible par le réseau InterNet (appendice II).

5.2 Logiciels utilisés

5.2.1 A.L.F. Manager™

Le logiciel A.L.F. Manager™ (version 2.13; Pharmacia) est nécessaire pour le fonctionnement du séquenceur automatique. La principale fonction d'analyse du logiciel est l'interprétation des chromatogrammes d'électrophorèse. Selon des algorithmes élaborés par les laboratoires de Pharmacia, le logiciel procède à un traitement mathématique des données d'intensité en plusieurs étapes:

- 1- Réduction du bruit de fond;
- 2- Soustraction du bruit de fond des détecteurs;
- 3- Normalisation de l'amplitude des intensités de fluorescence;
- 4- Mesure de la séparation des pics;
- 5- Mesure géométrique de la forme des pics;
- 6- Optimisation des pics afin d'augmenter leur séparation;
- 7- Attribution d'un nucléotide aux différents pics obtenus.

De plus, le logiciel permet l'analyse détaillée des chromatogrammes par l'utilisateur. Il est possible de modifier les séquences au besoin. Par exemple, les ambiguïtés ont été résolues manuellement en comparant le profil des courbes dans les 5 minutes avant et après la séquence non-déterminée par le logiciel.

Le logiciel est conçu pour la plate-forme OS/2 mais les séquences peuvent être exportées vers d'autres plates-formes comme DOS, Windows, UNIX et VAX/VMS.

5.2.2 ReadSeq

Ce logiciel a été conçu par Don Gilbert (Indiana University, Bloomington) et permet la conversion des séquences nucléotidiques et peptidiques entre plusieurs formats. Le logiciel détecte automatiquement le format des fichiers à convertir. La conversion peut être effectuée entre dix-huit formats courant dont: IG/Stanford, Genbank/GB, NBRF, EMBL, GCG, Pearson/FastA, PIR/CODATA, ASN.1, MSF, etc.

La version la plus récente de ce logiciel est disponible via le réseau InterNet (appendice II).

5.2.3 DNASIS et PROSIS

Ces logiciels commerciaux sont distribués par HIBIO (Hitachi) et permettent l'analyse de séquences nucléotidiques (DNASIS) et peptidiques (PROSIS). DNASIS a été utilisé pour les profils de restriction, la détermination des cadres de lecture ainsi que pour la traduction des séquences. Il a été également utilisé pour la prédiction des structures secondaires des mRNAs à l'aide de la méthode de Chou et Fasman (1978). Principalement, ce logiciel a permis d'effectuer une recherche d'homologie préliminaire avec GenBank sur CD-ROM avant d'effectuer les

recherches avec FastA sur le système VAX.

PROSIS a été utilisé pour la recherche d'homologie entre les clones obtenus (séquence peptidique déduite) et les domaines catalytiques de protéine kinases de la banque PKCDD. PROSIS utilise l'algorithme de Lipman et Pearson (1985). Cette recherche a permis d'effectuer un tri préliminaire avant d'effectuer les recherches avec FastA sur le système VAX.

5.2.4 *Sequence Analysis Software Package*™ (S.A.S.P.) de Genetics Computer Group (GCG)

5.2.4.1 Organisation du logiciel S.A.S.P. de GCG

Le logiciel S.A.S.P. version 7.2 (Devereux *et al.*, 1984) comprend une multitude d'applications permettant la modification et l'analyse de séquences nucléotidiques ou peptidiques. De plus, S.A.S.P. permet la recherche d'homologie avec GenBank/PIR et EMBL ainsi que la création d'alignements multiples. Le logiciel est conçu pour le système VAX/VMS. La majorité des analyses a été effectuée sur un ordinateur VAX 6410 avec le système d'exploitation VMS version 5.5 à l'Université Laval par l'intermédiaire du réseau de l'Université du Québec et du réseau R.I.S.Q. D'autre part, le système VAX/VMS (v5.4-2) de l'Institut de Recherche Clinique de Montréal a également été utilisé, principalement pour les analyses avec l'application PileUp.

5.2.4.2 FastA

L'application FastA a été utilisée pour toutes les recherches d'homologie. Cette application utilise les algorithmes de Pearson et Lipman (Pearson et Lipman, 1988) pour déterminer la similarité entre une séquence cible et les composantes

des banques de données visées. Les paramètres par défaut n'ont pas été changés car ils convenaient à nos besoins. Les principaux sont:

WordSize = 6 : longueur des séquences initialement homologues
 PAMfactor = 250 : matrice d'identité utilisée pour les acides aminés

5.2.4.3 PileUP

L'application PileUp génère un alignement multiple à partir d'un groupe de séquences homologues en utilisant progressivement des alignements pairés. Le logiciel utilise une version simplifiée de la méthode d'alignements progressifs de Feng et Doolittle (1987). La méthode est similaire à celle décrite par Higgins et Sharp (1989).

En plus de la liste des séquences à aligner, trois paramètres sont requis. Leur valeur par défaut sont:

GAPweight = 3.0 : pénalité imposée pour les brèches dans la séquence
 LENGHTweight = 0.1 : pénalité variant selon la longueur de chaque brèche
 Density = 100 : nombre de séquences par 100 alignements

5.2.5 Alignements multiples à l'aide du logiciel MACAW (*Multiple Alignment and Construction Analysis Workbench*)

Le logiciel MACAW (version 1.06) a été conçu par une équipe du *National Center for Biotechnology Information* (NCBI) au *National Institute of Health* (NIH, Bethesda). Ce logiciel a été conçu pour la construction interactive d'alignements multiples (Schuler *et al.*, 1991). Il permet à l'utilisateur de localiser, d'analyser (statistiquement), de modifier et de combiner des "blocs" de séquences alignées. Ce logiciel a l'avantage de pouvoir établir des alignements multiples locaux significatifs. Ceci en fait un des plus sensibles aux motifs qui apparaîtraient

statistiquement non-significatifs par les méthodes analysant les similarités pairées (Schuler *et al.*, 1991).

Le logiciel a été conçu pour l'environnement Windows™ version 3.1 et requiert l'utilisation de fonctions mathématiques avancées par l'intermédiaire d'un co-processeur mathématique. Il a été exploité sur un ordinateur compatible-IBM. Le système est équipé d'un processeur Intel i486Dx-33 avec 16MB de mémoire vive et utilise les capacités graphiques étendues (sVGA et 8514a) de l'environnement Windows. La dernière version du logiciel est disponible gratuitement au NCBI par le réseau InterNet (appendice II).

RÉSULTATS

1. Avant-Propos

Afin de connaître les événements déclenchés par le contact entre les thymocytes et le ligand gp23/45/90 des cellules E-5, il est nécessaire d'établir les caractéristiques biochimiques de ce complexe. En particulier, des connaissances plus approfondies concernant les événements de phosphorylation étaient considérés prioritaires. Ces événements jouent un rôle clé dans la transmission de signaux entre les thymocytes et les cellules E-5. L'interaction entre thymocytes et cellules E-5 se traduit par la phosphorylation de trois substrats chez les cellules E-5, dont une phosphoprotéine de 90 kDa. Par ailleurs, des expériences de phosphorylation *in vitro* à l'aide du complexe 23/45/90 isolé ont démontré que la gp90 est phosphorylée sur des résidus tyrosines. Par analogie, ces résultats suggèrent que la phosphorylation détectée proviendrait d'une autophosphorylation.

Le manque d'anticorps monoclonaux ou polyclonaux spécifiques à la gp90 en empêche la purification directe. La seule méthode disponible, jusqu'à maintenant, est la co-immunoprécipitation. La détection par autoradiographie de la gp90 isolée par co-immunoprécipitation, suite à l'analyse du profil de phosphorylation dépendant du contact entre les thymocytes et les cellules E-5, ne reflète pas le niveau d'expression de celle-ci. En fait, l'énergie élevée des émissions β du ^{32}P permet de déceler un niveau de phosphorylation très faible et l'analyse autoradiographique est limitée par un seuil de saturation peu élevé. De plus, l'analyse en SDS-PAGE montre qu'il existe plusieurs protéines ayant une masse moléculaire relative de 90 kDa. C'est donc l'incapacité d'identifier la gp90 parmi les autres glycoprotéines de même taille qui nous a contraint à ne pas utiliser l'électrophorèse en 2-D pour séparer une fraction de glycoprotéines de masse moléculaire relative de 90 kDa obtenue par électroélution suite à une séparation par SDS-PAGE.

Les travaux antérieurs ont démontré la possibilité de co-immunoprécipiter le complexe gp23/45/90 en absence de détergents. Afin de purifier le complexe en quantité appréciable, il a été envisagé de modifier cette méthode pour l'adapter à un système de chromatographie d'affinité. Cette purification initiale permettrait d'obtenir une fraction enrichie de gp90.

L'impossibilité d'obtenir le complexe ou la gp23 par cette méthode nous a obligé à reviser nos objectifs à court terme. Outre les molécules (ou complexes) transmembranaires agissant directement dans la transduction de signaux, des effecteurs internes sont responsables de l'amplification et de la régulation des signaux. Tel que nous l'avons discuté, la phosphorylation d'acides aminés clés chez certains récepteurs transmembranaires entraîne une régulation de l'activité enzymatique. Il est donc impératif de connaître les molécules impliquées dans les événements de phosphorylation des cellules E-5 afin de déterminer la nature physiologique des événements. D'une part, le peu de connaissances que nous avons des molécules impliquées dans les signaux de transduction chez les cellules E-5 et d'autre part l'impossibilité d'associer, avec un minimum de certitude, une tyrosine kinase connue à la phosphorylation observée chez les cellules E-5, nous ont permis de définir des objectifs vers lesquels la majorité du présent travail a été orientée.

2. Caractérisation biochimique du complexe gp23/45/90

2.1 Purification de la gp90 par chromatographie d'affinité

Les conditions de lavage et d'élution du système composé d'une pré-colonne avec un anticorps isotypique à laquelle a été jumelée, en série, une colonne anti-gp23, ont premièrement été mises au point. Des fractions de glycoprotéines membranaires marquées à l'iode-¹²⁵I ont été utilisées comme matériel à purifier. Les différentes conditions de lavage et d'élution sont indiquées au tableau IV.

Tampon de lavage	Tampon d'élution	Tampon de neutralisation
PO ₄ 10 mM, pH 8.0	triéthylamine 100 mM, pH 11.5	1/20 vol. PO ₄ 1 M, pH 6.8
PO ₄ 10 mM, pH 6.8	glycine 100 mM, pH 2.5	1/20 vol. PO ₄ 1 M, pH 8.0
PO ₄ 10 mM, pH 6.8	glycine 100 mM, pH 1.5	1/20 vol. PO ₄ 1 M, pH 8.0
PO ₄ 10 mM, pH 7.2	LiCl 5 M	----
PO ₄ 10 mM, pH 7.2	LiCl 8 M	----
PO ₄ 10 mM, pH 7.2	MgCl ₂ 3.5 M *	----
Tris-HCl 20 mM, pH 7.6	eau déionisée	----
Tris-HCl 20 mM, pH 7.6	triéthylamine 100 mM, pH 11.5	1/20 vol. PO ₄ 1 M, pH 6.8

* MgCl₂ 3.5 M dans du tampon PO₄ 10 mM, pH 7.2

Tableau IV: Conditions de lavage, d'élution et de neutralisation utilisées pour la mise au point de la purification du complexe gp23/45/90 par chromatographie d'affinité.

L'efficacité des différentes conditions de lavage ou d'élution de protéines marquées a été mesurée par un compteur gamma. L'élution à l'aide de

triéthylamine (pH 11.5) ou de glycine (pH 1.5) s'est avérée efficace. De plus, les conditions de lavage associées ont démontré une efficacité similaire. Nous avons choisi d'utiliser la triéthylamine afin de comparer nos résultats à ceux obtenus par d'autres groupes qui ont utilisé des protocoles similaires (Wilchek *et al.*, 1984).

La séparation des fractions obtenues n'a pas permis d'isoler le complexe. Afin d'améliorer la purification par chromatographie, les conditions de lavages et d'élution ont été légèrement modifiées, mais sans succès. La figure 7 illustre un profil type obtenu par SDS-PAGE. Bien que le gel semble surexposé, une exposition plus courte révèle le même profil.

2.2 Purification de la gp90 par co-immunoprécipitation

Le but étant d'obtenir une quantité suffisante de protéines pour effectuer un séquençage peptidique, il a été envisagé de répéter les expériences d'immunoprécipitation à l'aide d'un protocole standard faisant appel à du sépharose 4B-protéine G pour isoler les complexes. Une deuxième méthode a été utilisée, celle-ci faisant usage de Sepharose 4B auquel a été couplé l'anticorps anti-gp23 (C3C12). Afin de s'assurer de la spécificité de l'immunoprécipitation, toutes les expériences ont été effectuées en parallèle avec un contrôle isotypique (IgG2b de rat, anti-L3T4).

L'immunoprécipitation a été effectuée en parallèle avec des glycoprotéines membranaires de E-5 marquées (^{125}I) et avec une fraction de protéines membranaires marquées (^{125}I). Selon le type de fraction utilisé, les résultats se sont avérés différents. Une protéine formant une bande dominante d'environ 105 kDa a été immunoprécipitée à partir de ces deux fractions. Il a été impossible de déterminer si l'immunoprécipitation de la gp23 pourrait être efficace dans ces conditions. Du moins, elle n'a pas été détectée, et ce, suite à plusieurs expériences effectuées (figure 8).

Il faut également noter, dans plusieurs cas, la présence d'un front de migration dédoublé lors des SDS-PAGE effectuées suite aux immunoprécipitations et à la chromatographie d'affinité. On retrouve souvent un bande d'environ 20 kDa, au front supérieur (figure 7, puits 6) mais celle-ci se retrouve également chez les témoins (figure 8, puits 6 et 8) et semble être constituée de protéines qui migrent avec le front de glycine. De plus, les techniques de purification utilisées semblent favoriser l'immunoprécipitation de protéines de masse moléculaire relative d'environ 100 kDa.

Les méthodes de purification n'ont pas permis d'obtenir assez de matériel pour envisager un microséquençage peptidique. De plus, le peu de spécificité des immunoprécipitations ne permet pas d'attribuer une activité PTK aux protéines purifiées. L'impossibilité de purifier directement la gp90 à laquelle les résultats antérieurs ont associé une activité protéine tyrosine kinase nous a obligé à réorienter le projet. Sachant qu'il est possible de répertorier, par PCR, les transcrits codant pour les protéine kinases, nous avons choisi cette approche qui constitue la deuxième et majeure partie des travaux présentés ici.

Figure 7: Séparation par SDS-PAGE de fractions obtenues suite à l'élution de la colonne d'affinité C3C12, anti-gp23. Les fractions ont été obtenues séquentiellement suite à l'élution de la colonne d'affinité puis séparées par SDS-PAGE (12%, non-réduit, 24h d'exposition). À noter: le front de migration est dédoublé à l'interface glycine/chlorure. Des protéines de faibles poids moléculaire migrent avec celui-ci. La majorité des protéines éluées sont d'environ 70-100 kDa.

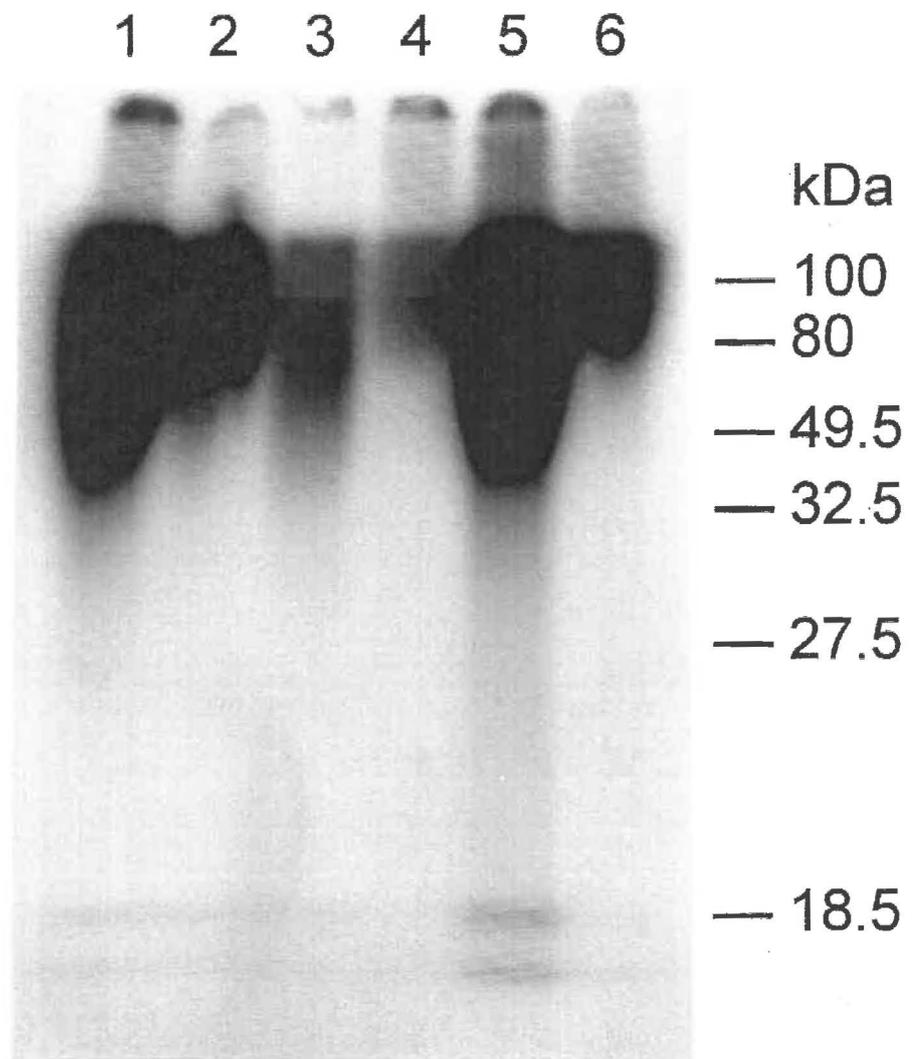
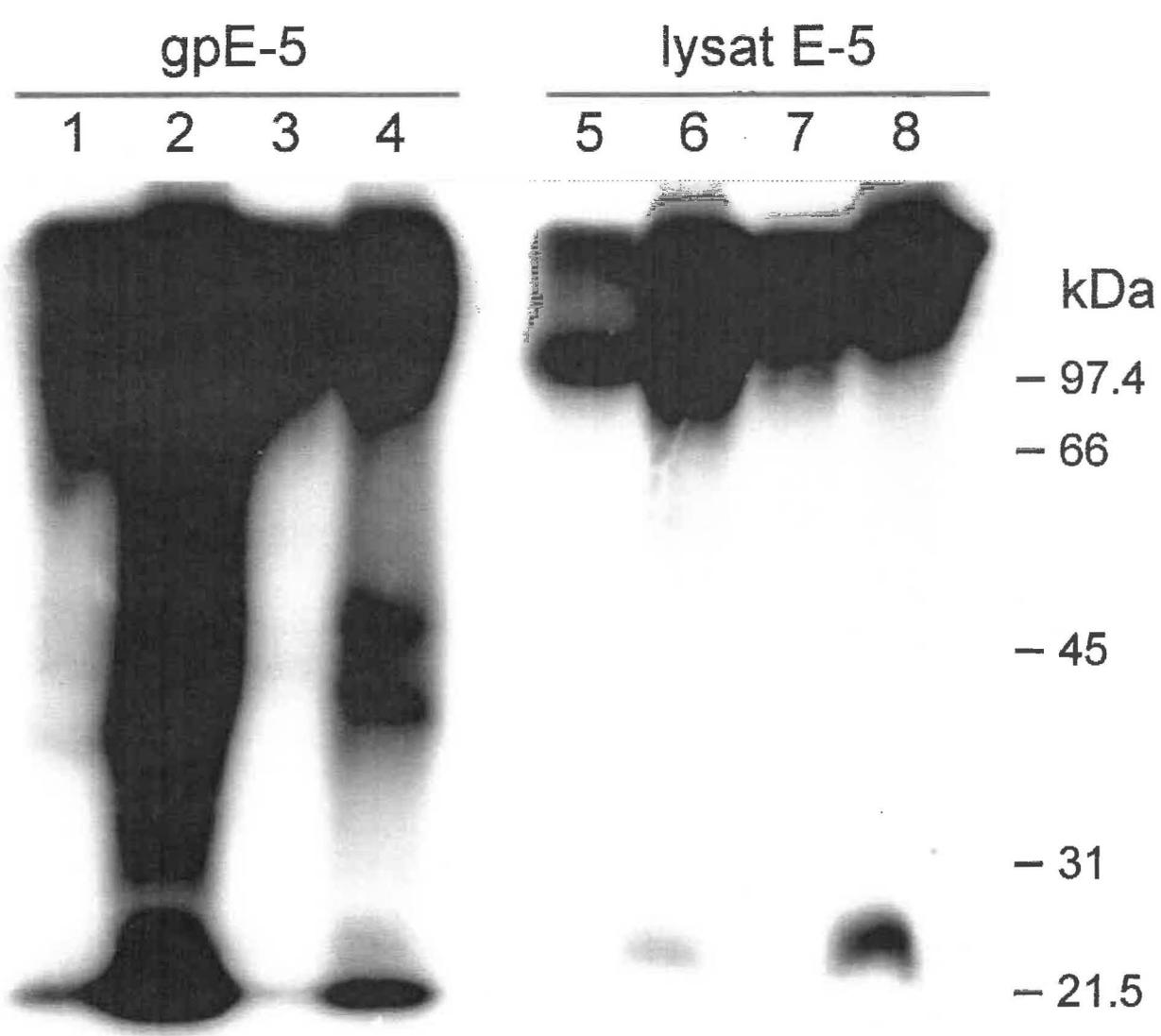


Figure 8: Immunoprécipitation de glycoprotéines membranaires et de fractions de protéines membranaires. Les protéines ont été préalablement marquées à l'iode 125 I. L'immunoprécipitation a été effectuée à partir de glycoprotéines membranaires de cellules E-5 (gauche) ou à partir d'un lysat de cellules E-5 (droite). L'anticorps C3C12, anti-gp23, a été utilisé en conjonction avec de la protéine G-sépharose ou couplé à du sépharose 4B. Un anticorps idiotypique anti-L3T4 a été utilisé comme témoin dans tous les cas. À noter: la présence d'une bande prédominante d'environ 105 kDa purifiée à partir des lysats cellulaires. Bien qu'il est difficile de le constater sur cette reproduction, l'original montre que cette bande n'est pas apparente pour l'immunoprécipitation effectuée à partir de glycoprotéines membranaires. Puits 1,5: mAb C3C12 / Protéine G Sepharose (PGS); 2,6: C3C12 couplé à Sepharose- 4B (S4B); 3,7: mAb anti-L3T4 / PGS; 4,8: anti-L3T4 couplé à S4B.



3. Caractérisation moléculaire des protéine tyrosine kinases des cellules E-5

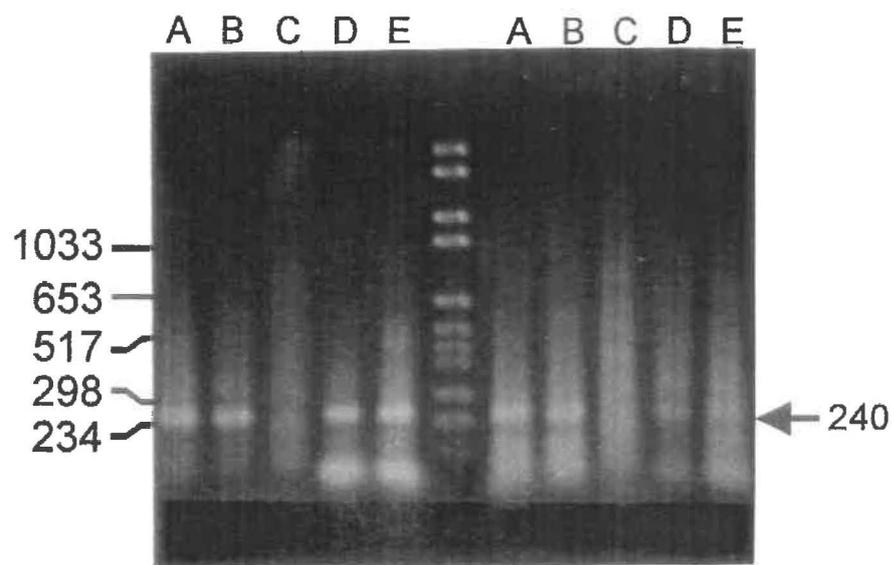
3.1 Amplification des régions VI - IX des domaines catalytiques des protéine tyrosine kinases à partir de cDNA de cellules E-5.

Dans le but de déterminer le profil de transcription des PTKs des cellules E-5, nous avons choisi d'utiliser une variante de la méthode de construction de librairie par PCR, spécifique aux protéine kinases, proposée par Andrew Wilks en 1989. Tel que présenté dans la revue bibliographique, la méthode consiste à purifier le RNA, effectuer la synthèse du cDNA puis finalement, amplifier les régions d'intérêt.

Nous avons utilisé quatre sources différentes de cDNA pour l'amplification par PCR (figure 9). Dans tous les cas, une étape préliminaire à l'amplification consistait à un traitement à la DNase d'un aliquot témoin de cDNA afin de confirmer de l'absence de trace de DNA génomique contaminant le RNA. La synthèse de cDNA à partir d'un oligonucléotide dégénéré spécifique à la région IX avait pour but de favoriser l'amplification des PTKs en pré-sélectionnant le matériel servant de gabarit pour l'amplification. Par ailleurs, nous avons également utilisé l'oligo dT et un analogue (RNA-tag) comme amorce pour la synthèse du cDNA. Ces amorces permettent d'obtenir du cDNA provenant principalement de mRNA (Frohman *et al.*, 1988). Le RNA-tag, élaboré par Peter Tijssen, consiste en un fragment poly-dT auquel a été ajouté une séquence comportant trois sites de restriction. Une amorce, Adapt, a été synthétisée et permet l'amplification par PCR à partir de la région contenant les sites.

L'amplification à partir de cDNA synthétisé avec l'amorce RNA-tag s'est avérée tout aussi efficace que l'oligo dT. Par contre, L'amplification à l'aide de

Figure 9: Amplification des motifs VI-IX des domaines catalytiques de PTKs à partir de différents cDNAs. Deux lots d'amorces ont été utilisés. À gauche, les amorces synthétisées au laboratoire du professeur François Shareck de l'Institut Armand-Frappier et à droite, les amorces provenant du service de synthèse d'oligonucléotides de l'Université de Calgary (lot "b"). Cinq cDNA synthétisés à partir d'amorces différentes ont servi pour l'amplification: A: poly-dT; B: 6-mers alléatoires; C: RNA-tag; D: PTK-c9R; e: PTK-c9Rb (lot "b"). Les amorces utilisées pour l'amplification sont PTK-c6F et PTK-c9R sauf pour le cDNA synthétisé à partir de l'amorce RNA-tag où les amorces PTK-c6F et Adapt ont été utilisées. On remarque que les fragments obtenus sont d'environ 240 nu, soit la taille du fragment estimé plus les amorces. À noter également: l'absence d'amplification avec PTK-c6F et Adapt.



l'amorce PTK-c6F et Adapt illustrée à la figure 9 n'a pas été efficace avec ce cDNA vraisemblablement à cause de l'hétérogénéité des fragments obtenus. Cet essai confirmait l'impossibilité d'amplifier l'extrémité 3' de plusieurs PTKs à l'aide d'une seule amorce dégénérée et de l'amorce Adapt.

Finalement, nous avons utilisé des amorces hexamériques alléatoires (*random hexamers*). Ces amorces sont efficaces pour la synthèse de cDNA à partir de RNA qui comporte des structures secondaires assez importantes qui empêcherait l'obtention de cDNA de pleine longueur.

L'amplification par PCR a généré un fragment d'environ 240 nucléotides (figure 9) correspondant à la taille calculée selon les alignements multiples des PTKs (figure 2) à laquelle la taille des amorces est ajoutée. L'intensité relative des fragments obtenus est la même dans tous les cas, ce qui démontre que chaque cDNA possède une efficacité similaire lorsqu'utilisé comme gabarit pour l'amplification des régions VI-IX des domaines catalytiques de PTKs. Malgré l'absence, à la figure 9, d'un témoin où l'amplification a été effectuée à partir du RNA total, pour chaque lots de RNA préparé, aucun gène apparenté à des PTK n'a été amplifié à partir de trace de DNA génomique. La présence de DNA génomique pourrait, tout au plus, interférer avec l'amplification en diminuant son efficacité. La présence d'introns de taille importante (variable) entre les motifs VII et VIII (au sein de la région amplifiée) permet de distinguer aisément si le fragment amplifié provient de DNA génomique ou de mRNA (Garvin *et al.*, 1988; Hradetzky *et al.*, 1992).

3.2 Clonage des fragments de domaines catalytiques obtenus par PCR

Le clonage dans pBluescript II KS + s'est avéré efficace (nombre de colonies bleues peu important) lorsque le vecteur a préalablement été déphosphorylé à l'aide de la CIP.

Le clonage des fragments a été effectué à l'aide des sites *ClaI* et *BamHI* contenus dans les amorces PTK-c6F et PTK-c9R (figure 10). L'orientation des fragments a permis le séquençage à l'aide de l'amorce RP (*reverse primer*), générant ainsi directement le brin codant (+) suite au séquençage automatique.

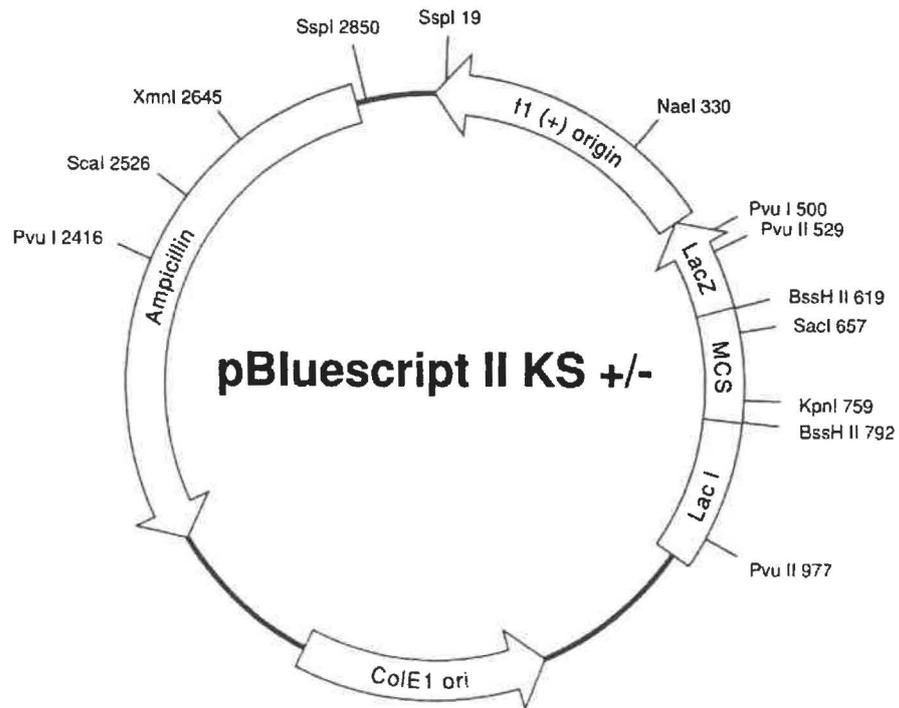
3.3 Criblage par PCR

Afin de déterminer la validité des clones obtenus suite à la transformation de bactéries compétentes XL1-Blue, nous avons procédé au criblage des colonies par PCR. Le criblage peut être effectué à l'aide des amorces ayant servi pour le PCR, d'amorces de séquençage, d'amorces spécifiques à des régions du vecteur ou encore une combinaison de celles-ci.

Les premiers essais ont été effectués à l'aide des amorces de séquençage. Bien que les résultats aient été satisfaisants, la petite taille des fragments clonés compliquait le criblage. La différence entre un fragment de 240 nu et un fragment de 400 nu est faible si l'on désire minimiser le temps de migration en utilisant un gel de 1.2% plutôt qu'un gel de 1.8%. Afin de faciliter le criblage, nous avons choisi l'amorce UPCRMP (*universal PCR mini-prep primer*) située dans le gène *lacZ* (position 1061-1039 dans pBluescript II KS+) retrouvé, entre autre, dans les vecteurs de la famille pUC. Cette amorce remplace l'amorce inverse et permet d'amplifier des fragments d'environ 520 nu en l'absence d'insert et d'environ 680 nu en présence d'un fragment de 210 nu cloné entre *ClaI* et *BamHI*¹. Ces tailles de fragments sont plus faciles à discerner sur des gels de 1.2% (figure 11a) d'autant plus que le criblage a été effectué de routine sur des gels de 15 cm x 10 cm en utilisant deux peignes (figure 11b). Cette méthode est efficace et fiable pour le criblage rapide si des témoins négatifs sont placés aux extrémités du gel.

¹ distance entre les amorces (~ 520) moins la distance entre *ClaI* et *BamHI* (~ 40) plus la taille du fragment (~ 200 nu).

Figure 10: Stratégie utilisée pour le clonage des motifs VI-IX des domaines catalytiques de PTKs. La figure montre la carte de restriction du vecteur pBluescript II KS+ (Stratagene Cloning Systems) utilisé ainsi que les sites principaux du site de clonage multiple (MCS). L'orientation des fragments clonés permet le séquençage de la séquence codante avec l'amorce RP (reverse M13 primer).



Site de clonage multiple (MCS):

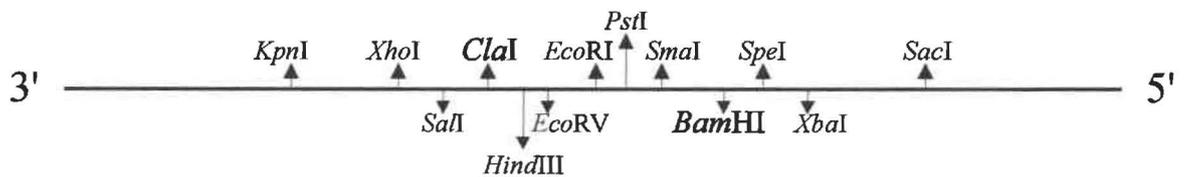
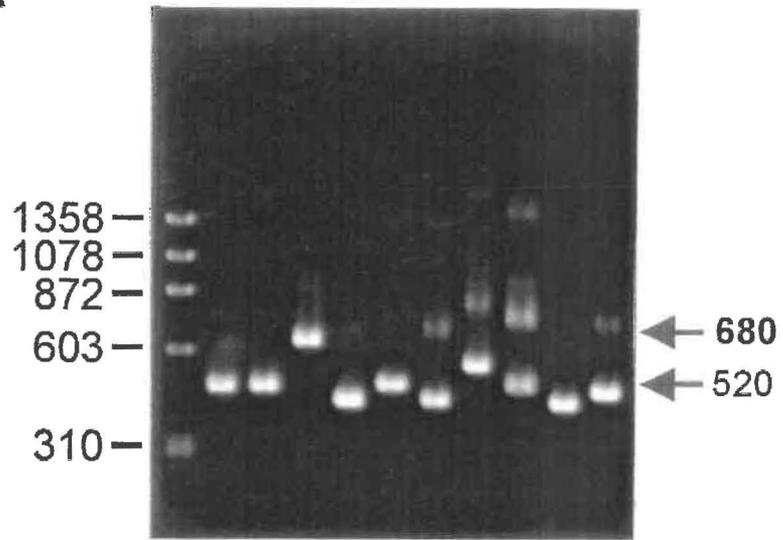


Figure 11: Minipréparations par PCR des clones contenant des fragments de domaines catalytiques. L'amplification a été effectuée telle que décrite dans la section matériel et méthodes, à l'aide des amorces UPCRMP et UP.

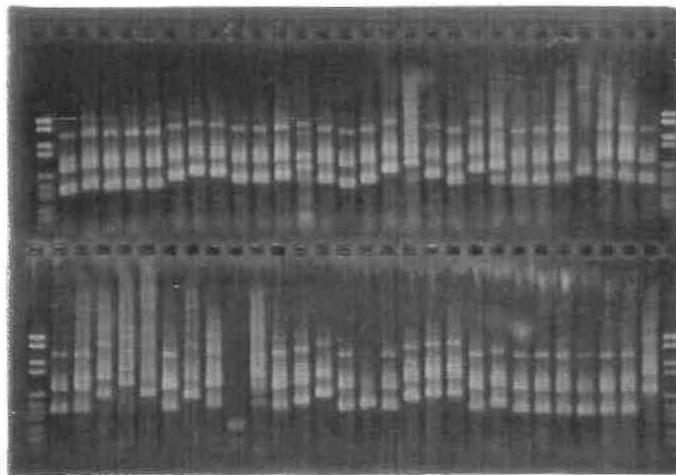
A) La flèche à 680nu indique les amplicons de clones positifs (+) et les amplicons de clones de références bleus migrent à 520.

B) le criblage d'environ 60 clones a été effectué de routine sur un même gel à l'aide de deux rangées de puits. Afin de bien distinguer les clones possédant un fragment cloné, chaque rangée de puits débute et termine par un fragment amplifié à partir d'une colonie bleue, soit un témoin négatif. Ainsi, les faux positifs (de sélection au X-Gal) sont facile à repérer.

A



B



Environ 125 clones contenant des fragments de la taille anticipée ont été choisis de façon aléatoire pour le séquençage.

3.4 Séquençage: établissement du catalogue des mRNAs PTK-spécifiques présents chez les cellules E-5.

Le nombre de protéine kinases connues au mois d'avril 1993 était de 261. Il s'agit des kinases répertoriées dans la banque de données (PKCDD) maintenue par Anne Marie Quinn et Steven Hanks du Salk Institute de San Diego. Ce répertoire ne contient pas la totalité des homologues retrouvés chez toutes les espèces. On peut donc estimer que plus de 300 protéine kinases sont connues. Ce nombre élevé suggèrerait que la méthode utilisée permettrait d'identifier, chez les cellules E-5, plusieurs kinases qui ont déjà été séquencées. Effectivement, nous avons identifié treize protéine kinases connues dans les cellules E-5. Nous estimons que ce nombre pourrait s'accroître si le criblage avait été poursuivi.

Bien que nous ayons tenté de cibler les PTKs, nous avons obtenu deux protéine sérine/thréonine kinases (PSTK) parmi les 13 protéine kinases connues. De plus, le criblage nous a permis d'identifier 3 nouvelles PTKs (tableau V). Durant les premiers mois du projet, nous avons isolé un clone (O21c) codant pour une kinase non répertoriée dans GenBank. Elle est maintenant connue sous le nom de *RYK* (GenBank #M98547; 92-9-13) suite à la publication des travaux de Christopher Hovens du laboratoire de Andrew Wilks (Hovens *et al.*, 1992). Cette PTK de type récepteur est inusitée, principalement à cause de la présence d'une modification au niveau du domaine VII où on retrouve la séquence -DNA- au lieu de la séquence -DFG- retrouvée universellement chez les PTKs (Hanks *et al.*, 1988; Hanks et Quinn, 1991).

Clone	# analogues ¹	Nom de la séquence/protéine
Protéine tyrosine kinases connues		
021c	-	locus <i>Musykina</i> / <i>RYK</i>
067c	-	<i>c-src</i> / pp60 ^{c-src}
071c	4	<i>lyn</i> / pp57 ^{lyn}
077c	-	<i>bek</i> / Bek
032e	-	<i>flg</i> / Fibroblast growth factor receptor
084e	7	PDGF-R a / platelet derived growth factor receptor, type A
099e	2	<i>Jak1</i> (Janus kinase 1) / N.D. ²
101e	2	<i>c-abl</i> / pp145 ^{c-abl}
111e	-	<i>met</i> / MET proto-oncogene PTK receptor
112e	4	<i>Jak2</i> (Janus kinase 2) / N.D.
115e	-	<i>tyro-3</i> / N.D.
Protéine sérine thréonine kinases connues		
073c	-	Branched-chain alpha-keto deH'ase kinase
085c	-	<i>B-Raf</i>
Nouvelles protéines tyrosine kinases		
109c	3	homologue de <i>bek</i> et <i>flg</i> (famille FGF-R)
060e	2	homologue de <i>c-yes</i> (famille <i>src</i>)
113e	-	homologue de <i>Cek-5</i> (famille <i>eph</i>)

1- analogues: les clones incomplets n'ont pas été comptabilisés même si les régions séquencées étaient identiques aux clones complets

2- N.D.: non-déterminé

Tableau V: Liste des protéine kinases identifiées (33) suite au clonage des régions VI-IX des domaines catalytiques.

3.5 Nouvelles protéine tyrosine kinases

3.5.1 Clone 109c -- une protéine tyrosine kinase de type récepteur homologue à *tek*.

Les problèmes de compressions sur les gels de séquençage ont été très fréquents. Le clone 109c en est un exemple. En effet, trois clones homologues ont été obtenus sans que nous ne puissions aligner parfaitement leur séquences. De plus, la traduction de ces clones générerait des codons d'arrêt dans les trois cadres de lecture. Il était, par contre, relativement facile d'identifier les régions conservées des domaines catalytiques pour trouver le bon cadre de lecture.

L'alignement multiple généré par Pile Up (GapWeight = 3.0, GapLengthWeight = 0.2) nous a permis de déterminer la séquence la plus probable du clone 109c, relativement à la séquence des deux autres clones, soit 089c et 111c (figure 12). La séquence nucléotidique déduite de l'alignement (figure 13) ainsi que la traduction de cette séquence ont été utilisées pour la recherche d'homologie à l'aide de l'application FastA (figure 14).

L'homologie maximale obtenue pour la séquence nucléotidique est de 67.0% sur une longueur de 221 nu de chevauchement avec la séquence de *srk1*, une PTK retrouvée chez *Spongilla lacustris*. Contrairement à certains clones analysés, l'homologie de séquence ne résulte pas de la présence d'une région où une concentration de résidus homologues est observée. De plus, il faut noter que le score *init1* le plus élevé a été attribué à *srk1*. Ce score est attribué à chaque clone selon sa région la plus élevée de densité initiale d'homologie. Lors de la juxtaposition, chez un même clone, de toutes les régions ayant initialement un score plus élevé que le seuil minimal prédéterminé d'identité, la présence de brèches (*gaps*) importantes chez *srk1* a affecté son score *initn* ($init1 = initn = 191$). L'indice de similarité initiale le plus élevé, basé sur le score *initn*, a été retrouvé

Figure 12: Alignement des séquences nucléotidiques des clones 109c, 111c et 089c. Cet alignement généré à l'aide du logiciel MACAW permet d'identifier la séquence la plus probable du clone 109c. Les points représentent des espacements dans l'alignement.

	ATCGATATCC	A.CCGAGAT.	TTGG..CTGC	AA..GAAA.T	GTCCTTGTT.
	-----	-----	-----	-----	-----
109c	ATCGATATCC	A.CCGAGAT.	TTGG..CTGC	AA..GAAA.T	GTCCTTGTT.
111c	ATCGATATCC	ACCCGGGAGC	.TGG..CTGC	AA..GAAA.T	GTCCTCGTT.
089c	ATCGATATAC	A.CAGAGAT.	TTGGGCCTGC	AAA.GAAAAT	GT.CTTGTTT
	GGTGAACATA	ATATCTACAA	AGTAGCAGAT	TTTGGACTTG	CAAGAGTTTT
	-----	-----	-----	-----	-----
109c	GGTGAACATA	ATATCTACAA	AGTAGCAGAT	TTTGGACTTG	CAAGAGTTTT
111c	GGTGAACATA	ATATCTACAA	AGTAGCAGAT	TTTGGACTTG	CAAGAGTTTT
089c	GGTGAACATA	ATATCTACAA	AGTAGCAGAT	TTTGGACTTG	CAAGAGTTTT
	TAA.GGTAGA	CAATGAAGAC	ATCTATGAAT	CTAAACACGA	AATAAA..GC
	-----	-----	-----	-----	-----
109c	TAA.GGTAGA	CAATGAAGAC	ATCTATGAAT	CTAAACACGA	AATAAA..GC
111c	TA.GGGTAGA	TAATGAAGAC	ATCTATGAAT	CTAAACACGA	A.T.AA..GC
089c	TAAGGGTAGA	TAATGAAGAC	ATCTATGAAT	CTAAACACG.AATAGC
	.TG.CCAGTG	AAGTGGACTG	CACCCGAAGC	CA.TCGTACT	AATAAA.TCA
	-----	-----	-----	-----	-----
109c	.TGGCCAGTG	AAGTGGACTG	CACCCGAAGC	CATTCGTACT	AATAAATTCA
111c	CT..CCAGTG	AAGTGGACTG	CACCCGAAGC	CA.TCGTACT	AATAA..TCA
089c	.TG.CCAGTG	AAGT.GACTG	CACCCGA.GC	CA.TCGTACT	AATAAA.TCA
	GCATAGTTCT	GATGTGTGGT	CCTATGGGAT	CC	
	-----	-----	-----	---	
109c	GCATAGTTCT	GATGTGTGGT	CCTATGGGAT	CC	
111c	GCATAGTTCT	GA.GTCTGGT	.CTATGGGAT	CC	
089c	GCATAGT.C.	GATGTCT.GT	.CTATCGGAT	CC	

Figure 13: Séquence nucléotidique et motifs conservés du clone 109c. La séquence peptidique est déduite selon la table de correspondance universelle. Les motifs conservés VI, VII, VIII et IX sont indiqués dans l'ordre (encadrés). À noter: la présence du motif VI commun au PTKs n'appartenant pas à la sous-famille *src* conjointement à une forme du motif VIII caractéristique de la sous-famille *src*. La disparité observée entre les séquences provient d'erreurs d'incorporation durant l'amplification (substitutions) ainsi que de la difficulté à séquencer certaines régions riches en structures secondaires.

Figure 14: Recherche d'homologie du clone 109c. Le clone 109c est principalement homologue avec *srk1* et *tek*. La recherche d'homologie a été effectuée avec le logiciel FASTA où *tek* possède le score *initn* le plus élevé et *srk1* possède le score *opt* le plus élevé. Les deux séquences présentent une homologie avec le clone 109c moins élevée que celle que l'on retrouve normalement chez les clones de PTKs analysés.

pour l'alignement du clone 109c avec *tek* suivi de l'alignement avec *flg* (FGF-r). Subséquemment, l'alignement optimal construit à l'aide de l'algorithme Needleman-Wunsch-Sellers (NWS) du logiciel FastA (Needleman et Wunsch, 1970) a démontré que la plus grande identité (score *opt*) est retrouvée entre la clone 109c et la sous-famille du FGF-récepteur (*flg*, *bek*) ainsi que *srk1* et *srk4*.

Les motifs conservés du domaine catalytique étaient présents dans la séquence du clone 109c (figure 13). En effet, on retrouvait le motif VI dont la séquence -IHRDLAARN- reflète l'appartenance aux PTKs autres que celles de la famille *src*. Le motif VII, dans sa forme -DFG- habituelle, ainsi que le résidu tyrosine subséquent étaient retrouvés. Il est à noter que le résidu tyrosine était plus distant (2 acides aminés) du motif -DFG- que chez d'autres PTKs. D'autre part, le motif VIII -KWTAPE- retrouvé chez le clone 109c est souvent représentatif des tyrosine kinases apparentées à la famille *src*.

3.5.2 Clone 060e -- une protéine kinase membre de la sous-famille *src*

Les nombreuses PTKs regroupées dans la sous-famille *src* reflètent les pressions que subissent ces gènes. Ces pressions entraînent des duplications et des mutations. Par contre, il existe d'autres pressions qui suggèrent la nécessité de conserver ces nouveaux gènes sous une forme similaire. L'analyse du clone 060e montrait clairement qu'il appartient à cette sous-famille (figure 15). On retrouvait les motifs -IHRDLRAAN- et -KWTAPE- qui sont associés à *src*. De plus, l'homologie entre les membres de cette famille et le clone 060e était très élevée (figure 16).

La divergence du profil d'homologie entre le clone 060e et les membres de la sous-famille *src* se situait immédiatement en aval du motif VI. Le clone 060e était à toute fin pratique identique à *c-yes*, sauf pour le résidu isoleucine qui remplaçait une valine. Ce changement est peu important d'après la valeur

Figure 15: Séquence nucléotidique et motifs conservés du clone 060e. La séquence peptidique est déduite selon la table de correspondance universelle. Les motifs conservés VI, VII, VIII et IX sont indiqués (encadrés). À noter: la présence la forme des motifs VI et VIII représentatifs des PTKs appartenant à la sous-famille *src*.

Figure 16: Alignement multiple des motifs VI à IX entre le clone 060e et trois membres de la sous-famille *src*. Le clone 060e présente l'homologie la plus élevée avec *c-yes* bien que la présence des deux résidus isoleucine soit caractéristique de *fyn*.

C-SRC vhrdlraanilvg**enlvckv**adfglarliedneyt
C-YES ihrdlraanilvg**enlvcki**adfglarliedneyt
FYN ihrdlrsanilvg**nglicki**adfglarliedneyt
060e ihrdlraanilvg**enlicki**adfglarliedneyt

C-SRC arqgakfpikwtapeaalygrftiksdvwsfgi
C-YES arqgakfpikwtapeaalygrftiksdvwsfgi
FYN arqgakfpikwtapeaalygrftiksdvwsfgi
060e arqgakfpikwtapeaalygrftiksdvwsfgi

E:	GAA GAG	I:	ATT ATC ATA
N:	AAT AAC	V:	GTT GTC GTA GTG
G:	GGN		

LOCUS HUMCYES1, 4517 bp
 DEFINITION Human c-yes-1 mRNA. (ACCESSION M15990)

SCORES Init1: 666 Initn: 666 Opt: 704
 90.5% identity in 211 bp overlap

```

                                10      20      30
Ptk060                        ATCGATATCCACCGAGATCTCCGAGCTGCT
                                | | | | | | | | | | | | | | | | | | | | | |
Humcye GCTGATGGTATGGCATATATTGAAAGAATGAACTATATTCACCGAGATCTTCGGGCTGCT
      1350      1360      1370      1380      1390      1400

                                40      50      60      70      80      90
Ptk060 AATATTCTTGTAGGAGAAAATCTTATATGCAAATAGCAGATTTTGGCTTAGCAAGATTA
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humcye AATATTCTTGTAGGAGAAAATCTTGTGTGCAAATAGCAGACTTTGGTTTAGCAAGGTTA
      1410      1420      1430      1440      1450      1460

                                100     110     120     130     140     150
Ptk060 ATTGAAGACAATGAATACACGGCAAGACAAGGTGCAAAATTTCCAATCAAGTGGACAGCT
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humcye ATTGAAGACAATGAATACACAGCAAGACAAGGTGCAAAATTTCCAATCAAAATGGACAGCT
      1470     1480     1490     1500     1510     1520

                                160     170     180     190     200     210
Ptk060 CCTGAGGCTGCTCTGTATGGTCGATTTACAATAAAGTCAGACGTCTGGTCCTTTGGGATCC
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humcye CCTGAAGCTGCACCTGTATGGTCGGTTTACAATAAAGTCGTATGTCTGGTCATTTGGAATTC
      1530     1540     1550     1560     1570     1580
  
```

ID KYES HUMAN, 543 AA. (AC P07947)
 DE YES PROTO-ONCOGENE TYROSINE KINASE (EC 2.7.1.112) (P61-cYES)

SCORES Init1: 324 Initn: 324 Opt: 324
 98.5% identity in 68 aa overlap

```

                                10      20      30      40      50      60      70
060e IDIHRDLRAANILVGENLICKIADFGLARLIEDNEYTARQGAKFPIKWTAPEAALYGRFTIKSDVWSFGI
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
yes  NYIHRDLRAANILVGENLVCKIADFGLARLIEDNEYTARQGAKFPIKWTAPEAALYGRFTIKSDVWSFGI
      400      410      420      430      440      450      460
  
```

Figure 17: Aligment de la séquence nucléotidique et peptidique déduite du clone 060e avec c-yes. L'aligment a été effectué suite à la recherche d'homologie avec le logiciel FASTA. À noter: alors que plusieurs différences sont observées au niveau de la séquence nucléotidique, la seule différence entre la séquence peptidique déduite de 060e et celle de c-yes provient de la substitution d'un résidu valine (c-yes) pour un résidu isoleucine (060e).

(V → I = 4) associée à ce remplacement dans la matrice PAM250. Selon la séquence nucléotidique, ce changement serait attribué au remplacement du codon GTG → ATA (figure 17).

3.5.3 Clone 113e -- une protéine kinase membre de la famille *eph*

La séquence nucléotidique traduite de 113e présentait tous les motifs communs aux PTKs (figure 18). Le clone 113e présentait une homologie à la sous-famille *eph* dans laquelle on retrouve maintenant 5 membres principaux, soit *eph*, *elk*, *eck* et *eeek* (figure 19; Lindberg et Hunter, 1990; Maru *et al.*, 1988; Chan et Watt, 1991; Lhoták *et al.*, 1991) et un nouveau membre, *sek* (MMsek: AC X65138, Gilardi-Hebenstreit *et al.*, 1992) que l'on croyait présent uniquement chez la souris alors qu'au début de mai 1993, un homologue aviaire (*Cek-8*) a été identifié (AC S57168; Sajjadi et Pasquale, 1993). Le motif VI -IHRDLRAAN-, tel que retrouvé dans le clone 109c, est caractéristique des PTKs autres que *src*. Le motif VII était conservé mais le motif VIII présentait une légère altération commune à la sous-famille *eph*. Le motif -KWTAPE- était altéré en -RWTAPE-. Dans leur ensemble, tant la séquence nucléotidique que peptidique (déduite) présentaient une identité maximale avec *Cek-5* (figure 20) bien qu'elles partageaient des identités propres soit à *Cek-5*, *elk*, *eeek* ou *eck* (figures 21, 22).

La signature d'identité propre à *Cek-5* et 113e consistait en un résidu phénylalanine (F) conservé en position 29 (de 113e). Les autres caractéristiques propres au clone 113e étaient les suivantes:

- 1) l'absence d'un résidu (L) entre les acides aminés en position 16 et 17;
- 2) un motif unique -SDP- (35), -FDG- chez *eph*;
- 3) un motif -SSL- (41) retrouvé chez *Elk*; partiellement chez *Cek-5* (-SAL-);
- 4) un résidu isoleucine (49) présent chez tous sauf *elk*;
- 5) un résidu sérine (16) retrouvé seulement chez 113e, *eck* et *elk*.

La variabilité rencontrée chez cette sous-famille n'était pas différente de celle de la sous-famille *src*. Nous avons pu tirer les mêmes conclusions de la comparaison entre 113e et *Cek-5* en ce qui concerne le pourcentage d'identité des séquences nucléotidique et peptidique. En effet, nous avons observé une dégénérescence élevée à la troisième position de 36.2% (21/58) des codons. La région des amorces (8 aa en 5' et 7 aa en 3') n'a pas été compté afin de ne pas biaiser les résultats. Ces résultats montrent clairement la pression sélective exercée sur certains résidus alors qu'il sont situés dans des régions hautement mutables.

Figure 18: Séquence nucléotidique et motifs conservés du clone 113e. La séquence peptidique est déduite selon la table de correspondance universelle. Les motifs conservés VI, VII, VIII et IX sont indiqués dans l'ordre (encadrés). À noter: la présence du motif VIII caractéristique des PTKs de type récepteur dont les membres de la sous-famille *eph* font partie.

Figure 19: Alignement multiple des motifs VI à IX entre le clone 113e et des membres de la sous-famille *eph*. L'alignement multiple a été généré avec le logiciel Pile Up. Aligné aux principaux membres de la sous-famille *eph*, le clone 113e présente une homologie élevée avec *elk*. Les motifs conservés VI à IX du domaine catalytique sont indiqués.

	121				160
Elk	MNYVHRDLAA	RNILVNSNLV	CKVSDFGLSR	YLQDDTSDPT	
Eck	MNYVHRDLAA	RNILVNSNLV	CKVSDFGLSR	VLEDD.PEAT	
Eek	LGYIHRDLAA	RNILVDGRLV	CKVSDFGLSR	ALEDD.PEAA	
Eph	HNYVHRDLAA	RNILVNQNLV	CKVSDFGLTR	LLDD..FDGT	
113e	...IHSDLAA	RNILVNST.V	CKVSDFGLSR	FLEENSSDPT	

VI

VII

	161				200
Elk	YTSSLGGKIP	VRWTAPEAIA	YRKFTSASDV	WSYGIVMWEV	
Eck	YTTS.GGKIP	IRWTAPEAIS	YRKFTSASDV	WSFGIVMWEV	
Eek	YTTA.GGKIP	IRWTAPEAIA	FRTFSSASDV	WSFGVVMWEV	
Eph	YETQ.GGKIP	IRWTAPEAIA	HRIFTTASDV	WSFGIVMWEV	
113e	YTSSLGGKIP	IRWTAPEAIA	FRKFTSASDV	WSYGI.....	

VIII

IX

Figure 21 a,b,c: Recherche d'homologie de la séquence nucléotidique du clone 113e avec les principaux membres de la sous-famille *eph*. La recherche a été effectuée à l'aide du logiciel FastA. La présence de brèches dans l'alignement est presque sans conséquence pour les séquences peptidiques, le cadre de lecture étant toujours respecté. De plus, ces brèches sont situées à des endroits différents selon les PTKs de cette sous-famille.

LOCUS HUMERK 1225 bp ss-mRNA PRI 25-MAR-1993
 DEFINITION Human mRNA for large erk kinase AC. D14717
 SCORES Init1: 390 Initn: 447 Opt: 502
 76.0% identity in 221 bp overlap

```

                                10      20      30
113e.D                          ATCGATATCCACAGCGACTTGGCTGCTAGG
                                | | | | | | | | | | | | | | | | | |
Humerk GCAGCTGGCATGAAGTACCTGGCAGACATGAACTATGTTACCGTGACCTGGCTGCCCCG
      280      290      300      310      320      330

      40      50      60      70      80
113e.D AACATCTTGGTCAACAGTA---CGGTCTGCAAGGTGTCCGACTTTGGCCTCTCCAGATTC
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humerk AACATCCTCGTCAACAGCAACCTGGTCTGCAAGGTGTCCGACTTTGGGCTCTCACGCTTT
      340      350      360      370      380      390

      90      100     110     120     130     140
113e.D TTGGAGGAGAACTCCTCTGATCCCACCTACACAAGTTCCTGGGAGGCAAGATCCCATC
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humerk CTAGAGGACGATACCTCAGACCCACCTACACCAGTGCCCTGGGCGGAAAGATCCCATC
      400      410     420     430     440     450

      150     160     170     180     190     200
113e.D CGATGGACCGCCCCTGAGGCCATTGCCTTCAGGAAGTTCACCTCTGCCAGTGACGTCTGG
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humerk CGCTGGACAGCCCCGGAAGCCATCCAGTACCGGAAGTTCACCTCGGCCAGTGATGTGTGG
      460     470     480     490     500     510
  
```

LOCUS HUMECK 3921 bp ss-mRNA PRI 13-DEC-1990
 DEFINITION Human protein tyrosine kinase mRNA AC. M59371 M36395
 SCORES Init1: 218 Initn: 387 Opt: 487
 76.5% identity in 221 bp overlap

```

                                10      20      30
113e.D                          ATCGATATCCACAGCGACTTGGCTGCTAGG
                                | | | | | | | | | | | | | | | | | |
Humeck GCAGCTGGCATGAAGTACCTGGCCAACATGAACTATGTGCACCGTGACCTGGCTGCCCCG
      2290     2300     2310     2320     2330     2340

      40      50      60      70      80
113e.D AACATCTTGGTCAACAGTA---CGGTCTGCAAGGTGTCCGACTTTGGCCTCTCCAGATTC
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humeck AACATCCTCGTCAACAGCAACCTGGTCTGCAAGGTGTCTGACTTTGGCCTGTCCCGCGTG
      2350     2360     2370     2380     2390     2400

      90      100     110     120     130     140
113e.D TTGGAGGAGAACTCCTCTGATCCCACCTACACAAGTTCCTGGGAGGCAAGATCCCATC
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humeck CTGGAGGACGACCCC---GAGGCCACCTACACCA---CCAGTGGCGGCAAGATCCCATC
      2410     2420     2430     2440     2450

      150     160     170     180     190     200
113e.D CGATGGACCGCCCCTGAGGCCATTGCCTTCAGGAAGTTCACCTCTGCCAGTGACGTCTGG
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Humeck CGCTGGACCGCCCCGAGGCCATTTCTACCGGAAGTTCACCTCTGCCAGCGACGTGTGG
      2460     2470     2480     2490     2500     2510
  
```

Figure 21a: Recherche d'identité nucléotidique entre le clone 113e et les membres de la sous-famille *eph*.
 Identité du clone 113e avec *erk* et *eck*.

LOCUS HUMTKR 3370 bp ss-mRNA PRI 15-JUN-1989
 DEFINITION Human tyrosine kinase receptor (eph) AC M18391
 SCORES Init1: 220 Initn: 343 Opt: 411
 72.8% identity in 217 bp overlap

```

                                10      20      30
113e.D                          ATCGATATCCACAGCGACTTGGCTGCTAGGAACA
Humtkr CTGGCATGAACTACCTCAGTAATCACAATTATGTCACCGGGACCTGGCTGCCAGAAACA
          2330      2340      2350      2360      2370      2380

          40      50      60      70      80      90
113e.D TCTTGGTCAA-CAGTAC--GGTCTGCAAGGTGTCCGACTTTGGCCTCTCCAGATTCTTGG
Humtkr TCTTGGTGAATCAAAACCTGTGCTGCAAGGTGTCTGACTTTGGCCTGACTCGCCTCCTGG
          2390      2400      2410      2420      2430      2440

          100     110     120     130     140     150
113e.D AGGAGAACTCCTCTGATCCCACCTACACAAGTTCCTGGGAGGCAAGATTCCCATCCGAT
Humtkr ATGA-----CTTTGATGGCACATACGAAA---CCCAGGGAGGAAAGATCCCTATCCGTT
          2450     2460     2470     2480     2490

          160     170     180     190     200     210
113e.D GGACCGCCCTGAGGCCATTGCCTTCAGGAAGTTCACCTCTGCCAGTGACGTCTGGTCTCT
Humtkr GGACAGCCCTGAAGCCATTGCCCATCGGATCTTACCACAGCCAGCGATGTGTGGAGCT
          2500     2510     2520     2530     2540     2550
  
```

LOCUS RNEEKR 1119 bp RNA ROD 11-NOV-1991
 DEFINITION Rat eek mRNA for eph and elk-related kinase AC X59290
 SCORES Init1: 204 Initn: 292 Opt: 404
 71.8% identity in 216 bp overlap

```

                                10      20      30
113e.D                          ATCGATATCCACAGCGACTTGGCTGCTAGGAA
Rneekr CGCTGGCATGCGCTACCTCTCGGACCTGGGCTACATTACCGTGACCTGGCCGCCCGCAA
          340      350      360      370      380      390

          40      50      60      70      80
113e.D CATCTTGGTCAA---CAGTACGGTCTGCAAGGTGTCCGACTTTGGCCTCTCCAGATTCTT
Rneekr CATCCTGGTAGATGGCCGTCTTGTGTGCAAGGTGTGCAAGGTGTGCAAGGTGTGCAAGGTGTG
          400      410      420      430      440      450

          90      100     110     120     130     140
113e.D GGAGGAGAACTCCTCTGATCCCACCTACACAAGTTCCTGGGAGGCAAGATTCCCATCCG
Rneekr GGAGG---ACGACCCTGAGGCCCGCCTACACCACTGC---GGGAGGGAAGATCCCTATCCG
          460      470     480     490     500

          150     160     170     180     190     200
113e.D ATGGACCGCCCTGAGGCCATTGCCTTCAGGAAGTTCACCTCTGCCAGTGACGTCTGGTC
Rneekr ATGGACAGCACCAGAGGCCATTGCCTTCCGCACTTTCTCCTCAGCCAGTGATGTGTGGAG
          510     520     530     540     550     560
  
```

Figure 21c: Recherche d'identité nucléotidique entre le clone 113e et les membres de la sous-famille *eph*.
 Identité du clone 113e avec *eph* et *eek*

DISCUSSION

1. Stratégie d'amplification

1.1 Avant-propos

Les méthodes utilisées pour purifier directement la PTK responsable des événements de phosphorylation observés, suite au contact avec des thymocytes, dans les cellules E-5 se sont avérées inefficaces. Suite à ces tentatives, le projet a été ré-orienté vers une caractérisation moléculaire des cellules. Nous avons donc établi le profil de transcription des gènes codant pour des PTKs chez les cellules E-5. Quatorze domaines catalytiques appartenant à des PTKs différentes ont été amplifiés par PCR. De celles-ci, trois domaines catalytiques de PTKs jamais identifiés jusqu'ici ont été séquencés. Parmi les onze PTKs déjà identifiées, une PTK de type récepteur était encore inconnue au moment du clonage du fragment d'amplification par notre laboratoire. La taille estimée de cette protéine et quelques-unes de ses propriétés suggèrent une ressemblance à la gp90. D'autre part, une nouvelle PTK que nous avons identifiée est associée à la sous-famille *src*, une autre à la sous-famille *eph* et une dernière présente une homologie au récepteur du facteur de croissance des fibroblastes (FGF-R: *fibroblast growth factor receptor*).

Nous discuterons chacune des nouvelles kinases ainsi certaines PTK connues dont la présence dans les cellules E-5 représente un intérêt particulier. De plus, nous discuterons de la divergence observée chez certaines nouvelles kinases face aux autres membres de leurs sous-familles respectives.

1.2 Amplification par PCR

1.2.1 Choix des amorces

Afin de répertorier les transcrits de PTKs, nous avons utilisé une modification de la méthode de librairie par PCR (Wilks *et al.*, 1989). Le clonage de quatorze domaines catalytiques appartenant à des PTKs, soit onze connues et trois nouvelles, démontre que les modifications apportées ne diminuent pas l'efficacité de cette méthode. Le nombre de PTKs obtenues suggère même une amélioration significative du processus d'amplification comparativement au nombre de PTKs obtenues à l'aide de cette méthode chez divers types cellulaires (Wilks, 1990). Les modifications apportées aux amorces ainsi qu'aux conditions de PCR seront discutées en regard des résultats obtenus par d'autres groupes.

Nos résultats confirment que la méthode développée par Wilks est spécifique aux PTKs quoique deux domaines catalytiques (régions VI-IX) de PSTKs connues aient également été clonés. L'amplification non-spécifique de deux PSTKs est minime compte tenu des nombreuses PSTKs présentes dans les cellules épithéliales (Hunter, 1991). Cette spécificité reflète l'efficacité de l'amplification sélective des PTKs. La sélectivité a été imposée par l'élaboration d'amorces selon un alignement multiple de PTKs tel que discuté précédemment.

Les amorces utilisées possédaient un niveau moyen de dégénérescence (PTKc6F = 32; PTKc9R = 16). Les travaux de Wilks démontrent clairement qu'une dégénérescence trop élevée réduit l'efficacité des amorces et qu'il existe un équilibre entre la dégénérescence et la longueur de ces amorces (Wilks *et al.*, 1989; Wilks 1991). Une revue par Compton (1991) suggère que de meilleurs résultats sont obtenus si la dégénérescence est maintenue aux environs de 128 à 256 alors qu'un maximum de dégénérescence équivalent à 512 fois est visé pour un 20 à 30-mer.

Les amorces utilisées étant dérivées d'alignements multiples de PTKs (Wilks, 1989), aucune modification n'a été apportée à la région d'appariement aux motifs de PTKs. De plus, la région 3' des amorces décrites par Wilks possédait une extrémité parfaitement complémentaire et elle n'a pas été modifiée. Ceci est critique pour favoriser l'appariement lors de l'amplification par PCR. Par contre, nous avons apporté quelques modifications à l'extrémité 5' des amorces que nous avons fait synthétiser.

Nous avons ajouté des sites de restrictions à l'extrémité 5' qui diffèrent de ceux originalement inclus par Wilks. Il est important de noter que ces sites n'ont jamais été retrouvés dans la région amplifiée (motifs VI-IX) des tyrosines kinases connues, ce qui empêcherait le clonage d'un fragment complet. Afin de favoriser l'étape subséquente de digestion enzymatique, des nucléotides de stabilisation ont été ajoutés en amont des sites de restriction (Kaufman et Evans, 1990). Ces nucléotides allongent les amorces et favorisent la stabilité des duplex durant l'amplification. L'amorce PTK-c6F possède deux sites de restriction en 5'. Le premier, *ClaI*, a été utilisé pour le clonage des amplicons, alors que le deuxième, *EcoRV*, a été inséré afin de permettre le sous-clonage, en phase, dans le vecteur d'expression procaryotique pMAL (Guan *et al.*, 1987). L'amorce PTK-c9R ne possède qu'un site *BamHI* et a été utilisée pour le clonage des amplicons mais peut être utilisée également pour le sous-clonage dans pMAL.

Nos résultats d'amplification par PCR démontrent que les nouvelles amorces sont spécifiques malgré la difficulté d'établir une comparaison avec les amorces originales décrites par Wilks, celles-ci n'ayant pas été utilisées en parallèle. Dans l'éventualité où un spectre de PTKs plus large serait désiré, il serait possible de modifier les amorces pour inclure plus de dégénérescence et de les allonger de quelques résidus pour pallier à d'éventuels mauvais appariements. Il serait également envisageable d'élaborer des amorces pouvant s'apparier à d'autres régions conservées comme les motifs II ou XI. De plus, il serait possible de cibler

les PTKs possédant un domaine conservé SH2 ou SH3.

1.2.2 Amplification des séquences codant pour les motifs VI-IX des domaines catalytiques de protéine tyrosine kinases.

La présence de sites de restriction et des nucléotides de stabilisation, en plus d'un niveau moyen de dégénérescence, a nécessité une modification des cycles de PCR. La méthode originale d'amplification des domaines catalytiques de PTKs a été décrite avec une température d'élongation de 63°C (Wilks, 1989; Wilks *et al.*, 1991). Les auteurs affirment obtenir de meilleurs résultats à cette température. Ils y observaient une stabilisation des duplex. Nous avons modifié les conditions de PCR afin de favoriser cette stabilisation. La modification consiste à ajouter quatre cycles préliminaires pendant lesquels la température d'appariement est diminuée. Nous avons également apporté d'autres modifications aux cycles de PCR. Certaines conditions ont été adaptées aux amorces dégénérées utilisées.

Le temps de passage entre le plateau de température d'appariement et l'élongation est critique pour favoriser la stabilité des duplex surtout lorsque le PCR est effectué à l'aide d'oligonucléotides dégénérés (Compton 1990). En effet, l'augmentation modérée de la température, suite à l'appariement, permet le début de la réaction de polymérisation à une température sub-optimale qui favorise l'appariement. Cette température est quand même assez élevée pour permettre à l'enzyme d'allonger considérablement le duplex et ainsi augmenter sa stabilité. Il est donc préférable d'augmenter lentement la température plutôt que d'effectuer toute la polymérisation à 63°C, ceci pour permettre à l'enzyme d'opérer à sa température optimale. De plus, la *Taq* polymérase, comme la majorité des enzymes, possède un pH auquel son activité est optimale. Sachant que le tampon d'amplification contient du Tris et que le pH du tampon Tris varie avec la température, il est préférable d'effectuer l'élongation à la température où le pH

optimal sera obtenu. Afin de respecter les conditions optimales de l'enzyme, l'élongation a donc été effectuée à 72°C. De toute façon, le fragment à amplifier est tellement petit (environ 200 nu) que l'enzyme a le temps de synthétiser une copie complète du fragment avant la fin du plateau d'élongation.

Nos travaux ont également démontré l'importance de diminuer la température à un rythme modéré suite à la dénaturation pour parvenir à la température d'appariement. La baisse modérée de la température continue d'imposer une contrainte thermodynamique sur la formation de duplex et favorise ainsi un appariement optimal.

Une fois les quatre cycles préliminaires de PCR terminés, la température d'appariement est augmentée, les temps de passage entre plateaux sont minimisés et les plateaux eux-mêmes sont plus courts (figure 6). Ceci contribue à abaisser le taux d'incorporation fautif de nucléotides par la *Taq* polymérase. Ce taux est minimisé par la réduction des temps d'appariement et d'élongation, par un appariement à température maximale et par l'utilisation d'une concentration équimolaire minimale de dNTP et de MgCl₂ (Eckert et Kunkel, 1990).

L'utilisation de cDNA synthétisé par différentes méthodes n'a pas changé pour autant les résultats d'amplification. Par contre, la synthèse de cDNA à l'aide de PTK-c9R (amorce aval) a permis de biaiser la sélection du mRNA utilisé comme gabarit. La population de cDNA ainsi obtenue est constituée presque uniquement de cDNA PTK-spécifiques.

Lors du PCR, l'amorce PTK-c9R a été utilisée en plus de PTK-c6F (amorce amont). Bien que certains auteurs affirment que des amorces internes sont nécessaires pour obtenir une ré-amplification, cette méthode s'est avérée efficace (Bej *et al.* 1991). Cependant, toute trace de DNA génomique contaminant peut interférer durant l'amplification en diminuant l'efficacité de l'amplification des

cDNAs et en faussant les résultats. Dans le cas des PTKs qui possèdent une organisation génomique très segmentée (Garvin *et al.*, 1988), la taille des amplicons permettrait normalement d'établir leur origine. Par exemple, la séquence génomique de *c-src* dépasse 8 kb alors que celle de *c-abl* dépasse 30 kb (Goff *et al.*, 1980; Takeya et Hanafusa, 1983). Dans le cas de *Drosophila melanogaster*, les introns, généralement plus petits ou même absents (Monson *et al.*, 1982; Snyder *et al.*, 1982), sont présents dans les gènes de PTKs (Sprenger *et al.*, 1989). La localisation génomique des introns des vertébrés diffère de celle de *Drosophila*: un intron est situé entre les motifs VII et VIII chez les vertébrés alors que deux introns sont présents (après le motif VII et avant le motif IX) chez *Drosophila* (Hradetzky *et al.*, 1992; Sprenger *et al.*, 1989). Par contre, il a été démontré que, dans le cas des vertébrés, même si la taille des introns est variable et que leur position diffère entre les vertébrés et *Drosophila*, l'organisation génomique des PTKs apparentées à *src* est seulement variable pour les exons de la portion N-terminale (Rouer *et al.*, 1989).

En résumé, nous avons apporté des modifications mineures aux amorces décrites originalement pour l'amplification des motifs VI-IX des domaines catalytiques des PTKs (Wilks, 1989). Ces modifications ont permis d'amplifier un fragment d'environ 210 nucléotides, représentatif du nombre de nucléotide moyen constituant les motifs VI à IX. Les expériences ultérieures ont permis de cloner 14 domaines différents de PTKs. En plus de onze PTKs déjà répertoriées, cette méthode nous a permis d'identifier trois PTKs jusqu'ici non-répertoriées dans GenBank ou dans une banque de données spécifiques aux domaines catalytiques des protéine kinases (PKCDD, Hanks and Quinn, Salk Institute).

1.3 Adaptations techniques

1.3.1 Augmentation de la sensibilité et de la spécificité de l'amplification par PCR suite à l'utilisation de paraffine comme barrière d'évaporation.

L'utilisation fréquente de la technique de polymérisation en chaîne nous a encouragé à tenter diverses modifications aux protocoles standards. Une de ces modifications a été publiée dans *Molecular and Cellular Probes* (Hébert *et al.*, 1993). Elle consiste en un remplacement de l'huile minérale par de la paraffine comme barrière d'évaporation lors de l'amplification par PCR.

Nous avons démontré que cette modification augmente la sensibilité de l'amplification de 50 à 200 fois (ou même plus) selon les systèmes. De plus, elle permet l'automatisation de la technique "hot-start" (D'aquila *et al.*, 1991). La diminution de l'amplification non-spécifique ainsi que la diminution de la formation des dimers d'amorces ont contribué à améliorer l'amplification des domaines catalytiques des PTKs.

1.3.2 Criblage par PCR

Afin de minimiser le temps entre le clonage des domaines catalytiques et le séquençage des clones susceptibles d'être valides, nous avons eu recours à une variante de la technique de criblage par PCR (Rapley et Walker, 1992).

L'utilisation des amorces dégénérée ayant servi pour le PCR, par opposition à des amorces internes, est déconseillée pour le criblage par PCR à cause la nature de l'amplification. En effet, l'amplification à partir de 1 μ L de bouillon de culture (16 h), ou encore de la colonie même, nécessite l'appariement le plus efficace. De plus, l'utilisation des amorces de PCR est également déconseillée car il est possible que l'intégrité des sites de clonage ne soit pas conservée (suite au clonage). Ceci est souvent reflété par la difficulté à extraire un fragment à l'aide des mêmes enzymes ayant servi pour le clonage. Suite à ces observations, l'amorce UPCRMP (CAGCGAGTCAGTGAGCGAGGAAG) a été conçue. Son utilisation avec 1 μ L de bouillon de culture a permis d'obtenir des résultats fiables. Un autre avantage à utiliser une culture de 16h est la rapidité d'obtention des

résultats par électrophorèse des amplicons. Les résultats pouvaient être obtenus en moins de 90 min une fois l'amplification par PCR terminée. La préparation des plasmides pour le séquençage pouvait donc être effectuée le jour même.

1.3.3 Complications du séquençage: structures secondaires et compressions

L'analyse des clones par séquençage manuel ou automatique à l'aide de la méthode de Sanger a été compliquée par la présence de structures secondaires importantes. Les problèmes de compression sur les gels ainsi que les régions de structures secondaires où la polymérase n'a pu synthétiser efficacement un brin complémentaire sont nombreux (figure 23). Afin de minimiser les problèmes de compressions, nous avons tenté le remplacement du dGTP par le dITP. Cette tentative s'est avérée infructueuse. Le séquençage automatique à l'aide de la trousse AutoRead™ (Pharmacia) s'est avéré plus efficace que le séquençage manuel pour résoudre les compressions grâce à la présence de 7-deaza dGTP (c^7dGTP). Le c^7dGTP permet d'éliminer les compressions qui résultent des interactions Watson-Crick G-C et Hoogsteen G-G (Mizusawa *et al.*, 1986). Par contre, la présence des structures secondaires est difficilement éliminée (figure 24a). Le séquençage à l'aide de l'amorce universelle (dans la direction opposée) a permis, à quelques reprises, de séquencer les régions compressées (figure 24b).

Figure 23: Structure secondaire du clone 085e suggérée par l'algorithme de Zuker et Stiegler ainsi qu'un exemple de compressions observées lors du séquençage manuel par la méthode de Sanger. Une autoradiographie des gels de séquence des clones 084e, 085e et 096e est illustrée. La structure secondaire proposée est basée sur la structure la plus stable possédant un minimum d'énergie libre (ΔG°). Elle n'est probablement pas représentative de la structure complète du mRNA, mais illustre bien la structure thermodynamiquement favorisée que l'on retrouve durant le séquençage. Une structure secondaire différente, mais tout aussi importante, est retrouvée chez tous les clones analysés. De plus, la difficulté à séquencer certains clones peut souvent être reliée à la stabilité élevée de leur structure secondaire (ΔG°). Conditions: Trousse T7 (Pharmacia) avec [α - ^{35}S]dATP sur gel 5%, exposition 16h.

Figure 24: Compressions causées par les structures secondaires lors du séquençage automatique. En (a), le clone 085e séquencé avec l'amorce RP. Les zones de compression sont identifiées par des symboles noirs. À la position 145, une structure secondaire ne peut être interprétée. Il s'agit probablement d'une compression que le logiciel a tenté d'interpréter comme une migration altérée temporairement. En (b), le clone 085e séquencé en direction inverse avec l'amorce UP. La formation des structures secondaires observées au séquençage de l'autre brin n'a pas été favorisée ici.

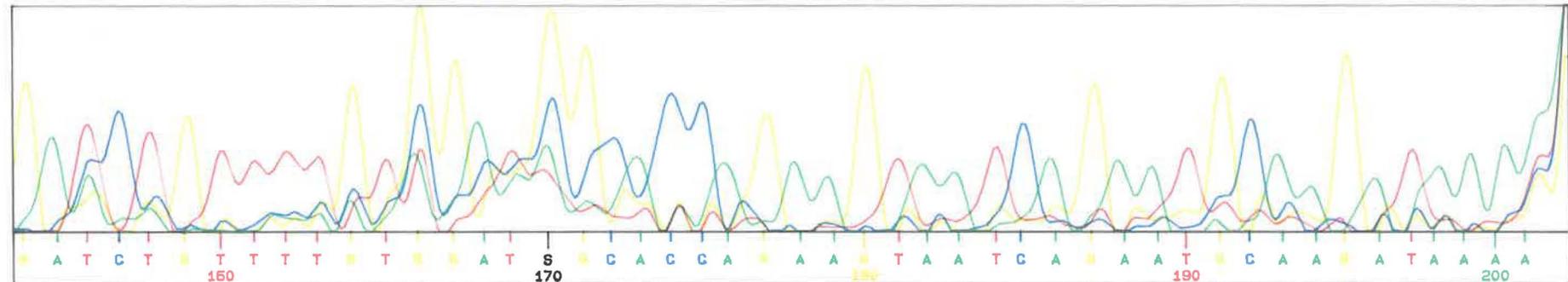
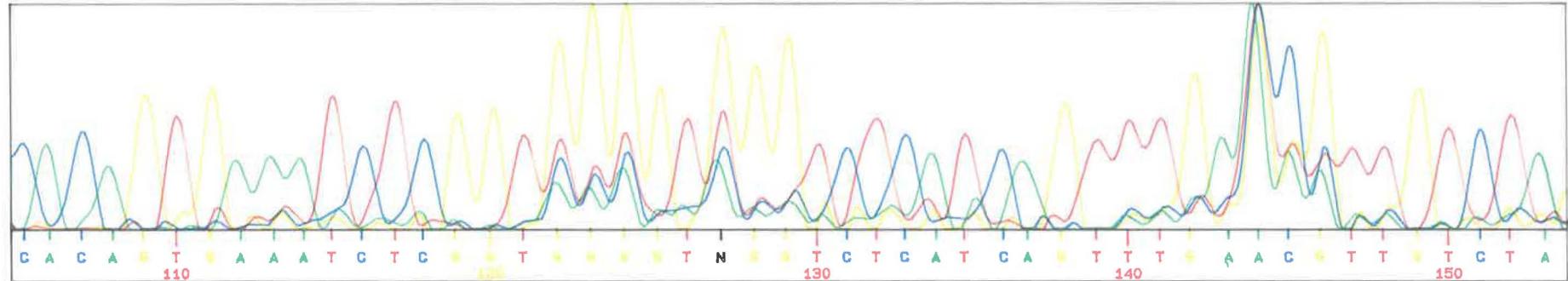
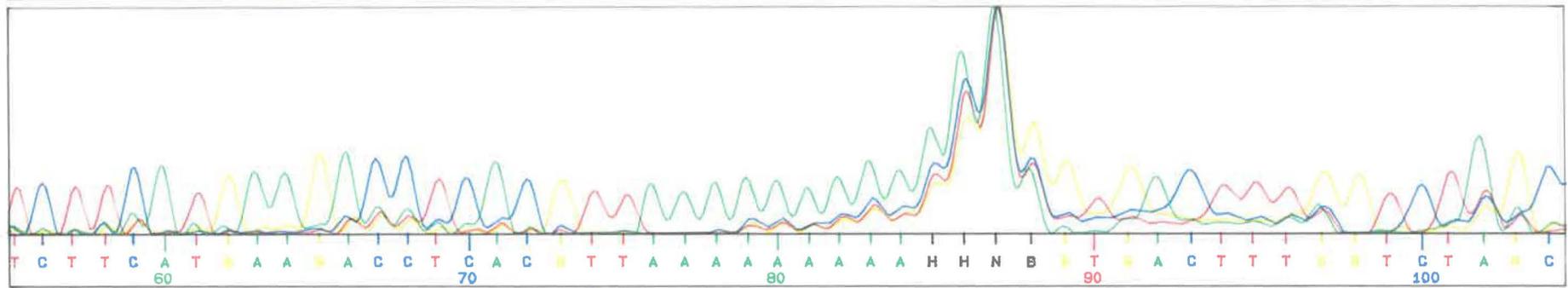
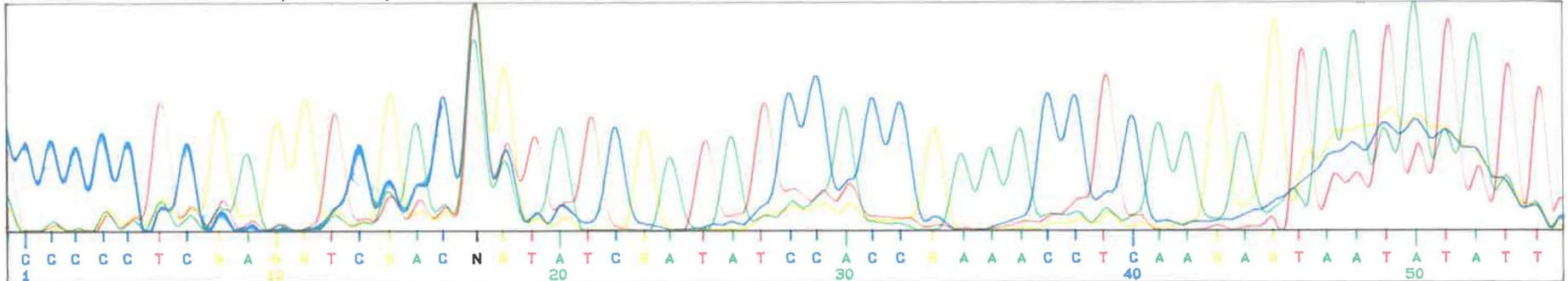
Code de couleurs: A: vert; C: bleu; G: jaune; T: rouge.



File: C:\AMDATA\930503.ALF
Run: PTK VI-IX in pBluescript

Clone: 7, BENOIT #85e RP (P)

Page: 1/1
Time: 01:18 .. 03:09 [h:m] Date: 05-03-93



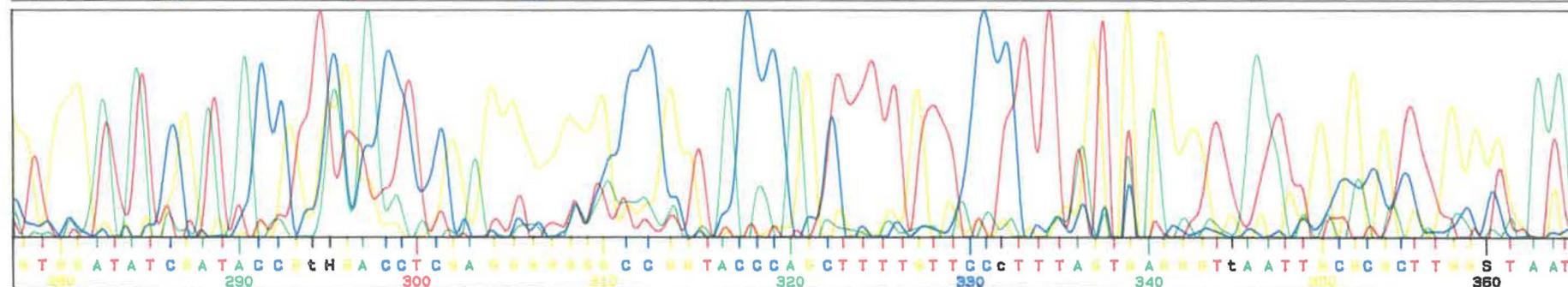
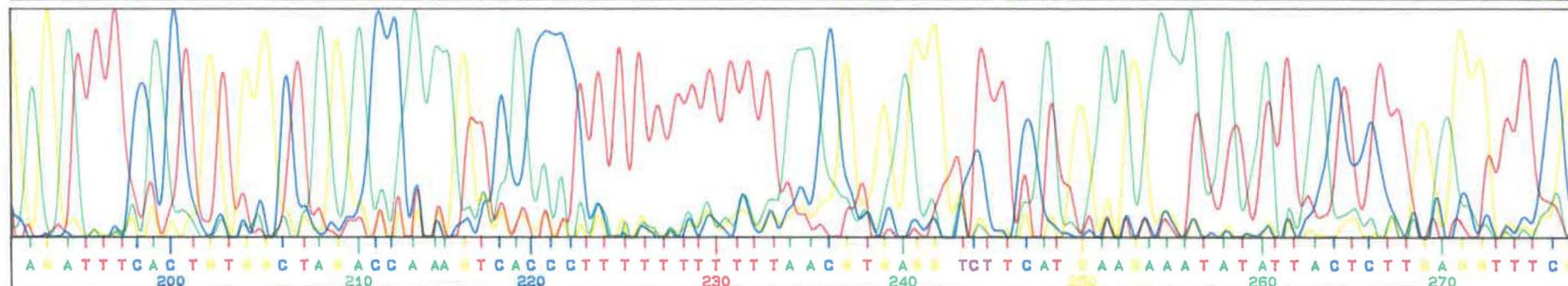
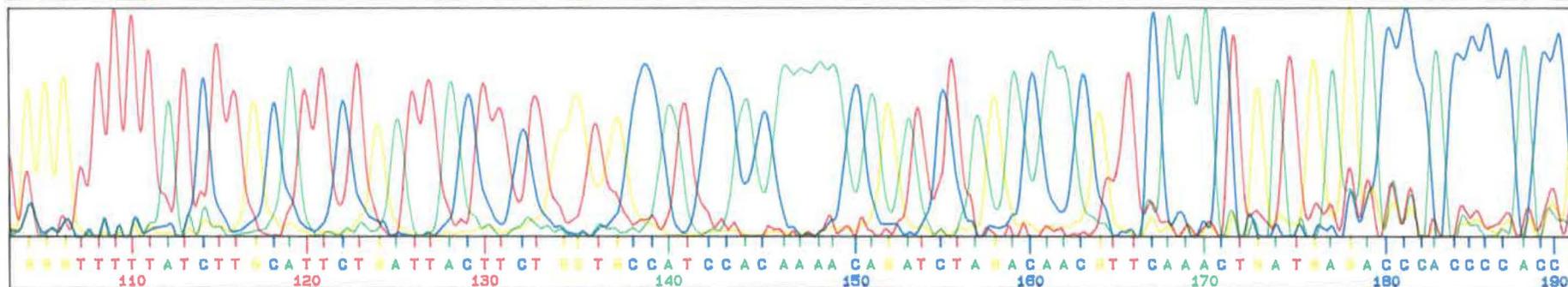
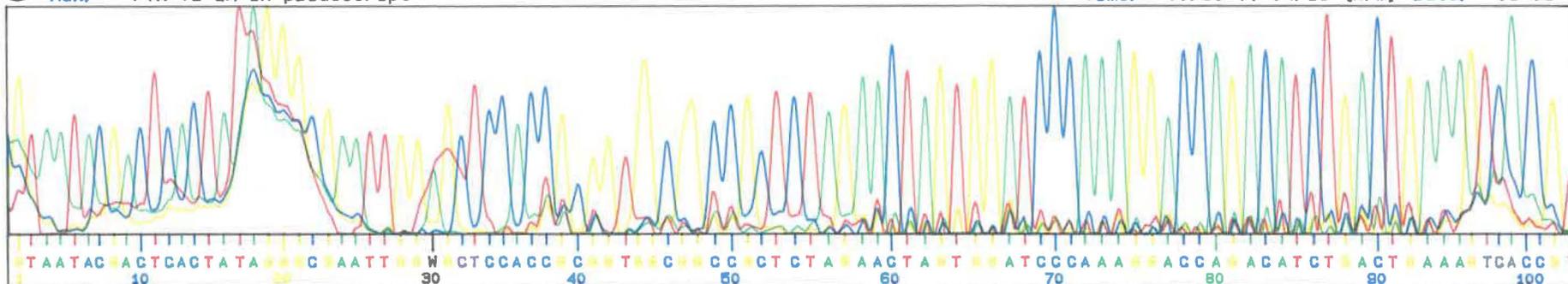
Comment: dATP



File: C:\AMDATA\930503.ALF
Run: PTK VI-IX in pBluescript

Clone: 8, BENOIT #85e UP (P)

Page: 1/1
Time: 00:56 .. 04:15 [h:m] Date: 05-03-93



Comment:

2. Les cellules E-5 transcrivent au moins onze gènes codant pour des PTKs déjà répertoriées.

2.1 Les banques de cDNA par PCR pour l'identification des PTKs

Comme nous l'avons vu précédemment, l'amplification des domaines catalytiques de PTKs a été efficace et sensible. Cette méthode nous a permis d'amplifier onze domaines catalytiques appartenant à des PTKs connues ainsi que trois appartenant à des PTKs jusqu'ici non-répertoriées. La grande diversité de PTKs obtenue démontre clairement la versatilité de cette méthode. Nous avons ainsi cloné des fragments appartenant à des PTKs de type intracellulaire ou de type récepteur. Dans chacun des cas, des membres appartenants à diverses sous-familles ont été obtenus.

L'efficacité de cette méthode n'est plus à démontrer; d'autres groupes l'ont également utilisée avec succès. Chez *Drosophila melanogaster*, Shishido *et al.* (1991) ont identifié sept nouvelles PTKs, soit *dtk1* à *dtk7*, à l'aide de la décrite originale (Wilks 1989).

La présence de quatorze PTK chez une même cellule peut sembler élevée, mais le nombre de rôles associés aux PTKs la justifie (Hunter, 1989). Bien que la méthode ait été sensible, il existe probablement d'autres PTKs connues chez les cellules E-5 que nous n'avons pas identifiées soit parce que le nombre de clones criblés n'est pas assez élevé ou que la méthode utilisée ne permet pas leur amplification. La présente méthode, du moins, permet l'amplification de PTKs appartenant à diverses familles (figure 25).

Deux sources de transcrits sont généralement utilisées par les groupes exploitant la méthode de Wilks. Les premiers ont eu recours aux lignées

Figure 25: Alignement multiple de huit clones correspondant à des PTKs déjà connues. L'alignement a été généré à l'aide du logiciel MACAW. Les séquences ont été utilisées pour la recherche initiale d'homologie. Ces séquences comportent encore certaines ambiguïtés identifiées par la lettre X. Les motifs conservés sont indiqués. À noter: *a priori*, les différentes formes du motif VIII permettent de classer ces PTKs selon les grandes sous-familles.

067c VHRDLRAANILVGENLVCKVADFGLARLIEDNEYTarq--GAKFPIKWTAPEAALYGRFtIKSDVWSFGI- 68
 101e IHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTaha--GAKFPIKWTAPESLAYNKFsIKSDVWAFGV- 68
 071c IHRDLRAANVLVSESLMCKIADFGLARVIEDNEYTare--GAKFPIKWTAPEAINFGCftIKSDVWSFGI- 68
 084e IHRDLAARNVLLAQGKIVKIRDFGLARDIMHDSNYvskg-STFLPVKWMAPESIFDNLYtTLSDVWSYGI- 69
 099e IHRDLAARNVLVESEHQVKIGDFGLTKAIETDKEYYtvkddRDSPVFWYAPECXICKXY-IASDVXXFGVh 70
 112e IHRDLATRNI LVENENRVKIGDFGLTKVLPQDKEYYkvkePGESPIFWYAPESLTESKFswpqcgpms- 70
 111e IHSDLAARNCMLDEKFXVKVXDFGLARDMYDKEYYsvhnktgaklpvkmaleslqtqkfttksdvcvlxg 71
 115e IHRDLAARNCMLAEDMTVCVADFGLSRKIYSGDYyrqgc-ASKLPVKWLALLESADNLYtVHSDVWSFGI- 69

VI

VII

VIII

IX

cellulaires; les autres, soit à du cDNA synthétisé à partir de RNA de source embryonnaire, soit à un homogénat de tissus. Les premiers tentent normalement d'associer la présence d'une kinase à un type cellulaire particulier. Les autres tentent souvent d'utiliser une source qui leur permet d'obtenir soit une grande diversité de kinases, soit une expression tissu-dépendante de celles-ci.

Par exemple, les travaux menant à la découverte de *Jak1* ont été initialement effectués à l'aide de la lignée cellulaire FDC-P1. Cette lignée cellulaire hématopoïétique est dépendante de facteurs de croissance (Wilks, 1989). D'autres groupes ont utilisé une source tissulaire ou cellulaire. La protéine tyrosine kinase *tek* a été isolée à partir de cDNA provenant de coeurs embryonnaires (Dumont *et al.*, 1992) puis *RYK* a été isolée à partir d'une librairie de cDNA provenant de macrophages péritonéaux (Hovens *et al.*, 1992).

Chaque méthode présente des désavantages. L'utilisation de lignées cellulaires immortalisées chimiquement (e.g., la lignée E-5) ou encore naturellement (HeLa) présente un problème potentiel, soit l'expression de PTKs normalement silencieuses dans ce type cellulaire. En effet, certaines PTKs ont été associées à la présence de translocations bien spécifiques (Park *et al.*, 1987). Par contre, l'augmentation du niveau de transcription suite à la transformation peut permettre l'identification de PTKs qui ne pourraient être détectées de façon routinière, malgré la sensibilité de la méthode. En effet, le niveau de transcription influence les résultats obtenus avec cette méthode.

L'utilisation de tissus ne permet pas d'associer l'expression d'une PTK à un type cellulaire précis. De plus, la contamination par des cellules autres que celles désirées, même en nombre restreint (e.g. cellules sanguines), peut générer des résultats difficilement reproductibles et fausser ainsi les interprétations. Chacune de ces méthodes doit être associée à des expériences d'hybridation *in situ* ou d'immunofluorescence afin de valider les résultats obtenus.

Nos résultats confirment que la méthode élaborée par Wilks est efficace étant donné le nombre de domaines catalytiques de PTKs clonés. Par contre, son utilisation est limitée par la source de transcrit de PTKs. Cette méthode, jumelée à des techniques d'hybridation, est probablement la plus sensible pour détecter des transcrits codant pour des nouvelles PTKs.

2.2 RYK: une PTK de type récepteur qui présente des idiosyncrasies dans trois motifs du domaine catalytique.

Parmi les PTKs connues identifiées chez les cellules E-5, la présence de *RYK* présente un intérêt particulier à cause de la ressemblance, *a priori*, à la gp90 co-immunoprécipitée à l'aide de l'anticorps C3C12. Nous verrons plus loin en quoi consistent les similarités.

Tel que mentionné, *RYK* a récemment été caractérisée par le groupe de Wilks (Hovens *et al.*, 1992). Afin de mieux comprendre les similarités que *RYK* peut avoir avec la gp90, les faits saillants des travaux de Hovens sont énumérés ici (Hovens *et al.*, 1992). La protéine codée est une PTK de type récepteur possédant un domaine transmembranaire très hydrophobe ainsi qu'une région extracellulaire. Ce domaine extracellulaire est très court (183 aa), comparativement à celui du PDGF-R (621 aa) ou du EGF-R (524 aa). Son domaine extracellulaire possède cinq sites potentiels de glycosylation de type N. La masse moléculaire estimée de cette protéine est de 64 kDa. Par contre, la même protéine immunoprécipitée de cellules NIH 3T3 puis séparée par SDS-PAGE possède une masse moléculaire relative de 90 kDa. De plus, cette PTK immunoprécipitée peut s'autophosphoryler. La résistance à l'hydroxide de potassium (1 M) suggère que la phosphorylation est localisée sur des résidus tyrosines (Kamps, 1991). Cependant, l'analyse des résidus phosphorylés n'a démontré aucune phosphorylation de tyrosines, seulement la présence de phosphosérine et de phosphothréonine. D'autres expériences de phosphorylation *in vitro* ont mis en

doute ces résultats: une protéine de fusion exprimée chez *E. coli*, purifiée par immunoprécipitation, n'a démontré aucune autophosphorylation ou activité kinase alors que normalement, les kinase traduites *in vitro* possèdent une activité (Ramer *et al.*, 1991). Ils estiment qu'un artéfact d'immunoprécipitation pouvait être en cause. Bien que les auteurs ne suggèrent pas d'explication, la phosphorylation observée est probablement produite par la co-immunoprécipitation d'une PSTK régulatrice de l'activité de *RYK*. Ceci met en doute l'observation que la phosphorylation sur des résidus sérine ou thréonine résulterait de l'activité de *RYK*. De plus, le(s) substrat(s) de cette PTK n'est pas connu. Il reste à déterminer si *RYK* possède réellement une activité tyrosine kinase et si l'autophosphorylation observée est physiologiquement significative.

Le groupe de Wilks ne discute malheureusement pas des autres protéines présentes lors de l'immunoprécipitation. On peut également apercevoir des protéines de 45 kDa, 65 kDa et 120 kDa immunoprécipitées en plus de la protéine de 90 kDa. Elles sont toutes détectées grâce à leur phosphorylation lors de l'essai kinase *in vitro*. La méthode qu'ils ont utilisée ne permet pas de déterminer si d'autres protéines non-phosphorylées sont également immunoprécipitées. L'intérêt majeur de ces travaux résulte de l'analogie entre la taille de protéines immunoprécipitées dans notre laboratoire à l'aide de l'anticorps C3C12 et celles immunoprécipitées par le groupe de Wilks. Bien qu'on ne puisse associer, *a priori*, ces résultats, il sera pertinent de vérifier le profil d'immunoprécipitation obtenu avec l'anticorps anti-*RYK* à partir de lysats de cellules E-5.

D'un point de vue moléculaire, *RYK* présente quelques particularités. La présence d'une altération du motif VII (-DNA- au lieu de -DFG-) du domaine catalytique de *RYK* est très inhabituelle. En effet, seulement une autre PTK (*klg*) possède une altération au motif VII (-ALS-; Chou *et al.*, 1991). D'autres idiosyncrasies au niveau des motifs I et II laissent croire à une structure tridimensionnelle altérée pour la liaison de l'ATP où les motifs I, II et VII seraient

possiblement impliqués (Hanks *et al.*, 1988).

En résumé, nous avons cloné le transcrite partiel de *RYK* à partir des cellules E-5. Cette PTK présente des analogies à la gp90 co-immunoprécipitée par l'anticorps C3C12. De plus, *RYK* possède des caractéristiques propres qui font d'elle une PTK unique.

2.3 La présence de certaines PTKs peut être liée à la transformation cellulaire.

Les cellules E-5 étant des cellules transformées, il est probable de retrouver des transcrits de PTKs souvent associés à la transformation cellulaire. En effet, *abl* et *MET* sont souvent associées à la transformation cellulaire où l'on observe une dérégulation de leur activation. Toutes deux sont présentes dans les cellules E-5.

Il a été démontré que la plupart des lignées cellulaires accumulent des translocations au cours de leur existence. La présence de *abl* y est associée (Davis *et al.*, 1985). Le modèle de leucémie myéloïde chronique est associé à une translocation spécifique: la présence du chromosome de Philadelphie (De Klein *et al.*, 1982) Cette translocation entre les chromosomes 9 et 22 engendre une fusion du locus *bcr* (c22) au gène du proto-oncogène *abl* (c9). Un transcrite hybride conduisant à une protéine hybride en résulte (Chan *et al.*, 1987). Bien que nous n'ayons pas tenté de détecter la présence de cette translocation, l'utilisation d'une cellule transformée augmente la probabilité de la retrouver.

D'autre part, une des premières caractérisations du proto-oncogène *MET* a été effectuée à partir de cellules HOS (MNNG-HOS) traitées au MNNG (Park *et al.*, 1987). L'activation de *MET* dans les cellules HOS provient également d'une translocation, du locus *TPR* (*translocated promoter region*) du chromosome 1 au locus *MET* du chromosome 7. Le mécanisme ressemble à celui menant à la formation de la fusion *BCR-ABL* (de Klein *et al.*, 1982). Il est à noter que *MET* se

retrouve d'une façon prédominante chez les cellules épithéliales ou fibroblastiques mais rarement chez les cellules d'origine hématopoïétique (Park *et al.*, 1987). Cette PTK de type récepteur agit probablement comme récepteur d'un facteur de croissance ou de différenciation spécifique à ces types cellulaires.

2.4 PTKs de type récepteur agissant comme récepteurs de facteurs de croissance

Les cellules E-5, tout comme plusieurs types de cellules cultivées *in vitro*, nécessitent des facteurs de croissances. Plusieurs récepteurs de facteurs de croissance possèdent une activité PTK. En plus de la présence de *MET*, la présence de *bek* et *flg*, toutes deux membres de la famille du FGF-récepteur, suggère le besoin des cellules E-5 pour des facteurs de croissance. Les cellules E-5 possèdent également le transcrite codant pour le PDGF-R de type a.

Outre la présence de kinases de type récepteur, on retrouve également deux PTKs de classe 3 caractérisées récemment, soit *Jak 1* et *Jak 2* (Wilks *et al.*, 1991; Harpur *et al.*, 1992). Ces PTKs de classe 3 possèdent un deuxième domaine catalytique dont la fonction est encore inconnue.

2.5 La présence de *tyro3* suggère un lien entre le système nerveux et le système immunitaire.

La présence de *tyro3* chez les cellules E-5 soulève une question fondamentale. L'expression de *tyro1-6* durant la neurogénèse ainsi que l'expression localisée au niveau du cerveau adulte ont mené l'équipe de Lai et Lemke à formuler l'hypothèse que ces PTKs possèdent un rôle dans le développement du système nerveux (Lai et Lemke, 1991). La présence de *tyro3* chez la cellule E-5 peut être causée par une transcription dérégulée due à la transformation cellulaire car on associe normalement la transformation à une dé-

différenciation cellulaire. Par contre, si l'expression de *tyro3* s'avère normale pour les cellules épithéliales du thymus murin, sa présence ne fait qu'ajouter aux indications liant certains mécanismes communs aux systèmes nerveux et immunitaire, dont la mémoire immunologique.

Il est clair que la signalisation cellule-cellule est un événement central dans la régulation du système immunitaire. Cette régulation s'effectue également via des facteurs solubles (peptide/protéines) et leur(s) récepteurs propres. Les modèles récents voulant expliquer la mémoire ("psychologique") sont basés à la fois sur les interactions peptides-récepteurs et des mécanismes de régulation hormonale. Il est donc possible d'établir des analogies fonctionnelles entre le système immunitaire et le système nerveux (Homo-Delarche et Dardenne, 1993) La présence de PTKs de type récepteur communes à ces deux systèmes ne ferait que renforcer ces hypothèses (Yamamori et Sarai, 1992).

3. Les cellules E-5 transcrivent au moins trois gènes codant pour des PTKs non-répertoriées.

3.1 Le processus décisionnel permettant la découverte de nouvelles PTKs.

Le processus analytique menant à la conclusion qu'un clone possède un fragment appartenant à un domaine catalytique d'une nouvelle PTK est complexe. Aucun logiciel existant ne permet une analyse résultant en un verdict sans équivoque. La recherche d'homologie primaire a été effectuée à l'aide du logiciel FastA. Par la suite, le clone a été analysé par alignements multiples, en fonction de la sous-famille à laquelle on peut l'associer *a priori*. La décision finale est basée sur des interprétations résultant de la connaissance des algorithmes régissant chaque traitement. Ceci implique que l'utilisateur connaisse et accepte les limites de chaque algorithme utilisé. D'ailleurs, Russell Doolittle (1990) l'a bien exprimé:

"Thus, although the computer is a wonderful helpmate for the sequence searcher or comparer, biochemists and molecular biologists must guard against the blind acceptance of any algorithmic output; given the choice, think like a biologist and not like a statistician."

Le pourcentage d'identité entre deux séquences nucléotidiques ou protéiques ne reflète pas toujours l'homologie structurale des protéines, surtout s'il s'agit de fragments. Par contre, la stratégie d'amplification utilisée ici impose une spécificité aux PTKs. Les fragments possèdent une très grande homologie étant donné la présence de motifs conservés, vestiges des pressions conservatrices régissant l'évolution des PTKs (Hanks *et al.*, 1988). Paradoxalement, plus le nombre de motifs conservés est grand, plus le pourcentage d'identité peut diminuer à cause des régions non-conservées entre ces motifs. Par contre, les scores reflétant l'homologie peuvent augmenter. Ces scores sont attribués à partir de la matrice de remplacement d'acides aminés

PAM250 (revue par George *et al.*, 1990). La probabilité que des acides aminés de même type se retrouvent aux mêmes endroits est grande, contribuant ainsi à augmenter les scores d'homologie. Cette probabilité découle de l'importance de l'environnement des motifs conservés au sein des domaines catalytiques. Ainsi, le pourcentage d'identité obtenu avec le logiciel FastA dépend de la longueur du fragment étudié. Il serait facile de déterminer manuellement un pourcentage d'homologie en comparant chaque paire d'acide aminés de deux protéines à condition qu'elles soient étroitement reliées. Par contre, la présence d'insertions ou de délétions empêchera ce type d'analyse.

La recherche d'homologie avec le programme FastA réussit à pallier à la limite imposée par les insertions et les délétions. Bien que la relation d'identité entre deux séquences soit plus évidente lorsque la comparaison de celles-ci est effectuée à l'aide des séquences peptidiques, les difficultés résultant de la présence de structures secondaires lors du séquençage ont compliqué la traduction des séquences. La difficulté à obtenir des cadres de lecture ouverts nous a forcé à effectuer la majorité des premières recherches d'identité à l'aide des séquences nucléotidiques. Les résultats préliminaires d'identité nous ont permis de reviser les régions problématiques de séquençage afin de résoudre certaines compressions et de déterminer si la séquence modifiée pouvait être traduite. La connaissance des motifs retrouvés normalement aide à suivre le bon cadre de lecture tout au long de la séquence. De plus, ces résultats préliminaires permettent de juger de la pertinence de poursuivre l'analyse avec les séquences traduites.

La dernière étape d'analyse, pour s'assurer de l'homologie de séquence, a été effectuée selon les mêmes algorithmes du logiciel FastA incorporés dans le logiciel PROSIS (HIBIO). Chacun des clones a été comparé aux membres de la banque de données des domaines catalytiques des kinases (PKCDD). Cette banque de données constitue une source souvent plus fiable que GenBank/PIR car la

majorité des nouvelles kinases s'y retrouvent avant de faire apparition dans GenBank.

L'appartenance à une sous-famille de PTKs a été déterminée par alignements multiples à l'aide de deux logiciels. Le premier, Pile Up, fait partie, tout comme FastA, de la série d'applications constituant le logiciel d'analyse de séquences de GCG (Sequence Analysis Software Package, version 7.2, Genetics Computer Group) développé pour les systèmes VAX. Le deuxième, MACAW, l'acronyme de *Multiple Alignment and Construction Analysis Workbench* (version 1.06, Schuler *et al.*, 1991), a été développé pour le système d'exploitation Windows™. Quant à leur fonctionnement, le logiciel Pile Up utilise une progression d'alignements pairés alors que MACAW utilise une méthode statistique basée sur un système interactif (donc potentiellement subjectif). MACAW permet de localiser, d'analyser et de valider la signification statistique de régions locales d'identité au sein de plusieurs séquences par opposition à une identité globale entre deux séquences.

Le principal avantage de l'alignement multiple sur les alignements pairés réside dans le fait qu'un motif d'homologie retrouvé entre deux séquences peut ne pas être statistiquement significatif; le même motif retrouvé au sein de plusieurs séquences peut le devenir. Ce type d'analyse permet donc un alignement sensible aux similarités pairées qui sont faibles ou statistiquement non-significatives. La flexibilité d'utilisation, ainsi que la possibilité de modifier la majorité des paramètres, font de MACAW un des plus puissants logiciels d'alignements multiples qui existent, à condition de comprendre les algorithmes utilisés ainsi que leurs limites.

Le processus décisionnel a été simplifié par les alignements multiples entre les clones obtenus et les principaux représentants des sous-familles de PTKs. Il est parfois difficile de déterminer si un clone est nouveau ou s'il s'agit tout

simplement d'un clone hybride. On définit un clone hybride comme un clone résultant de l'hybridation d'un fragment partiellement allongé durant le PCR avec une autre séquence que celle qui lui est propre. L'alignement multiple, à l'aide de MACAW, entre un clone douteux et des membres de différentes sous-familles de PTKs permet de mettre en évidence rapidement des identités locales. Si de multiples régions d'identité locale sont mutuellement exclusives à deux PTKs, le clone étudié est probablement un hybride. Cette possibilité est d'autant plus réelle quand l'homologie est retrouvée dans différentes régions ou domaines non-conservés qui constituent la signature d'une PTK particulière.

L'alignement multiple permet immédiatement de calculer le degré d'identité partagé entre un clone et deux PTKs différentes au niveau des régions non-homologues, c'est-à-dire, situées entre les motifs conservés. Deux des clones hybrides obtenus se sont avérés plus ardues à séquencer d'où la difficulté de trouver le cadre de lecture ouvert. Par exemple, l'analyse du clone 024e a uniquement été effectuée avec sa séquence nucléotidique. La recherche d'homologie ainsi que les alignements multiples obtenus de MACAW ont démontrés que le clone 024e est un hybride (figure 26) de FGF-R de rat et d'un clone (W3.13) issu des premiers travaux de Wilks lors de la description de la méthode de construction de librairie de PTKs par PCR (Wilks *et al.*, 1989).

Les clones qui présentent des fragments de nouvelles PTKs ont tous été comparés, par alignements multiples, à divers représentants des sous-familles de PTK afin de s'assurer de leur authenticité.

3.2 Une nouvelle PTK de type récepteur: le clone 109c est homologue aux membres de la famille du FGF-R.

Le pourcentage d'identité entre la séquence nucléotidique du clone 109c et les membres de la famille du FGF-R est relativement peu élevé lorsque comparé

LOCUS RATFGR4 2648 bp ss-mRNA ROD 30-NOV-1992
 DEFINITION Rat fibroblast growth factor receptor subtype 4 (FGFR4)
 mRNA, complete cds.
 ACCESSION M91599
 SOURCE Rattus rattus male adult lung cDNA to mRNA. . . .

SCORES Init1: 179 Initn: 316 Opt: 183
 79.5% identity in 83 bp overlap

```

                                10      20      30
24e.Dn                           ATCGATATCCACCGGGACCTGGCTGCCAGGAACGTG
                                |||
Ratfgr GGCATGCAGTATCTGGAGTCTCGGAAGTGCATCCACCGGGACCTGGCTGCCCGAAACGTG
      1430      1440      1450      1460      1470      1480

      40      50      60      70      80      90
24e.Dn CTGGTG---GGTGATGACCTCACCTGCAAGGTAGCTGATTTTGGCCTGGCCAGACTGCTC
      |||
Ratfgr CTGGTGACCGAGGATGACGTGA--TG-AAGATCGCTGACTTTGGTCTGGCCCGTGGTGTG
      1490      1500      1510      1520      1530      1540

      100     110     120     130     140     150
24e.Dn AAGGTCAGTGAGGCTCTGGGCTCCTAATCCCTGCATAGAGGAAAATACTACTCTGAACTC
Ratfgr CACCACATCGACTACTATAAGAAAACCAGCAATGGCCGCCTGCCAGTCAAGTGGATGGCT
      1550     1560     1570     1580     1590     1600
  
```

LOCUS MUSPTKG 208 bp ss-mRNA ROD 15-SEP-1990
 DEFINITION Mouse protein-tyrosine kinase (PTK) partial cds,
 clone W3.13.
 ACCESSION M33427
 SOURCE Mouse haemopoietic cell line WEH1-3B D+, cDNA to mRNA

SCORES Init1: 192 Initn: 317 Opt: 199
 66.9% identity in 118 bp overlap

```

      150     160     170     180     190     200
24e.Dn TCTGAACTCCCAACCATTATACCTCATCCACAGGATGATGTCTACTCCCAAGCAGTGGC
                                |||
Musptk GTATCTGATTTTGGAAATGGCCAGGTACGTTCTGGATGATCAGTACACAAGTTCTTCTGGC
      60      70      80      90      100     110

      210     220     230     240     250     260
24e.Dn TCCAAGATCCCTGTCAAGTGGACGGCACCTGAGGCTGCTAATTACCGTGTCTTTCCCAA
      |||
Musptk GCCAAGTTCCTGTGAAGTGGTGTCCCCAGAAGAGTTTAATTACAGCCGCTTTAGCAGC
      120     130     140     150     160     170

      270     280     290
24e.Dn AAGTCAGACGTGTGGTCTATGGGATCC
      |||
Musptk AAGTCAGACGTGTGGTCTATGGAATTCC
      180     190     200
  
```

Figure 26: Recherche d'identité du clone 024e à l'aide du logiciel FASTA. Le clone 024e est un hybride entre FGF-R de rat et le W3.13. À noter: l'homologie de la première moitié de la séquence avec une PTK et l'autre moitié avec une autre.

à l'homologie obtenue entre le clone 060e et la sous-famille *src*. Par contre, une densité d'homologie élevée présente aux domaines conservés permet d'associer ce clone à la sous-famille FGF-R.

L'association entre le clone 109c et la famille FGF-R est compliquée par la présence d'un motif normalement associé à la sous-famille *src* (motif VIII) conjointement à une forme du motif VI caractéristique des membres des autres sous-familles de PTKs excluant la sous-famille *src*. La séquence ne pourrait être rejetée d'emblée à cause de la présence de ces deux motifs qui semblent, *a priori*, incompatibles. En effet, plusieurs PTKs récentes possèdent certaines idiosyncrasies propres soit à des nouvelles sous-familles, soit à des membres d'origine lointaine.

Bien que le clone 109c ne semble pas être issu de la fusion de deux séquences durant le PCR pour former un clone hybride, il demeure le clone le moins attrayant des trois nouvelles séquences rapportées ici. La présence de trois clones presque identiques permet toutefois de croire à son existence réelle. Par contre, ces clones ayant été obtenus à partir du même clonage, il est probable que les trois originent de la même séquence amplifiée. L'existence d'une protéine renfermant les caractéristiques décrites plus haut reste à démontrer.

3.3 Un nouveau membre de la sous-famille *src*: le clone 060e présente une homologie élevée à *c-yes*.

Nous nous sommes interrogés sur la probabilité d'obtenir un clone si peu différent des membres de la sous-famille *src* qui pourrait avoir une quelconque capacité à coder pour une protéine différente. Un alignement multiple démontre la très grande homologie, au niveau du domaine catalytique, retrouvée chez les membres de la sous-famille *src* (figure 27). Le peu de divergence observé suggère une origine ancestrale unique (Hanks *et al.*, 1988). La taille du fragment obtenu

Figure 27: Alignement multiple des domaines catalytiques des principaux membres de la famille *src*. L'alignement a été effectué à l'aide du logiciel Pile Up. Les motifs conservés ombragés sont identifiés selon la nomenclature adoptée par Hanks (Hanks *et al.*, 1988). À noter: la seule divergence observée au niveau des motifs conservés se situe aux motifs I, II et IX. Les régions entre les motifs sont distinctes mais possèdent certains motifs uniques à cette sous-famille.

ainsi que sa localisation dans le domaine catalytique ne permet pas d'observer de divergence, du moins pour cette famille.

La fonction de chacune de ces protéines est dictée par sa région N-terminale (Timson Gauen *et al.*, 1992). Bien que les modes de régulation soit similaires, ils diffèrent d'une protéine à l'autre. Compte tenu que les interactions seront impliquées potentiellement dans cette régulation, la région N-terminale permet de distinguer clairement les PTK les unes des autres. Compte tenu de la taille du fragment que l'on possède ainsi que de l'alignement des deux séquences peptidiques, il semblerait difficile d'affirmer avec certitude que le clone O60e et *c-yes* sont différents. Après avoir analysé en détail les séquences nucléotidiques de *c-yes* et du clone O60e, plusieurs indications démontrent qu'il s'agit de séquences provenant de gènes distincts. La première indication concerne le remplacement d'un résidu valine pour un résidu isoleucine. Tel que mentionné, ce remplacement résulte de la substitution du codon GTG par un codon ATA. Intuitivement, le remplacement aurait semblé résulter d'une mutation ponctuelle GTA → ATA. Lors de la consultation de la séquence de *c-yes*, nous avons cherché à expliquer ce remplacement de GTG → ATA.

La substitution (transition) de deux nucléotides de ce codon peut être associée à une erreur de réplication causée par un mauvais appariement. D'autre part, nous avons rejeté l'hypothèse que ces modifications soient des erreurs d'incorporation de la *Taq* polymérase. Il serait simple d'expliquer les changements inattendus de séquence par l'incorporation d'erreurs dues à la *Taq* polymérase. En effet, il a été démontré que cet enzyme incorpore un nucléotide erroné à tous les 5 000 nucléotides (Saiki *et al.*, 1988) et qu'elle cause un changement de cadre de lecture à tous les 41 000 nucléotides (Tindall et Kunkell, 1988). En termes plus faciles à interpréter, on peut estimer qu'après une amplification de 25 cycles, pour chaque nucléotide d'une séquence quelconque, une molécule sur 400 présentera une substitution par rapport à la séquence originale (Saiki *et al.*, 1988).

Bien que ces erreurs soient cumulées au cours des cycles de réplifications, nous n'avons jamais obtenu un taux d'erreurs d'incorporation aussi élevé. La comparaison des séquences de gènes connus amplifiés par notre méthode avec les séquences publiées démontre qu'elles sont à toute fin identiques. Nous estimons que la *Taq* polymérase n'est donc pas responsable de toutes les substitutions observées. Par contre, nous observons des structures secondaires importantes chez les clones de PTKs. Leur influence sur la fidélité de la *Taq* polymérase n'est pas décrite dans la littérature.

Si la probabilité que la transition d'un résidu valine en isoleucine soit causée par une erreur d'incorporation est faible, il existe donc un mécanisme évolutif qui en est responsable. Le nombre de PTKs possédant une activité catalytique propre qui originent d'un même gène ancestral suggère qu'il existe une très grande pression sélective (conservatrice) qui s'assure de garder une conformation tridimensionnelle bien spécifique du domaine catalytique. Le remplacement d'un acide aminé ne doit pas altérer considérablement la capacité d'une PTK d'interagir avec son substrat. Ainsi, au cours des années, la multiplication de ces changements mineurs entraîne l'apparition d'une nouvelle spécificité. Pour un changement unique, la transition se fait évidemment via une mutation ponctuelle entraînée, la plupart du temps, par un mauvais appariement lors de la réplication.

La présence de deux nucléotides différents peut s'expliquer par des mutations ponctuelles apparaissant séquentiellement. On considère que le codon initialement présent à cette position était GTG. Ce codon est présent chez *src*, reconnu comme le gène ancestral de la famille. Pour ce codon, le remplacement du troisième nucléotide par un adénine entraîne un changement sans conséquence, GTA codant également pour l'acide aminé valine. Par la suite, une deuxième mutation, celle-ci en première position, entraîne le remplacement par l'isoleucine. Ce schéma est validé par l'existence de la forme intermédiaire de *c-yes* chez *Xenopus laevis* (AC X14377) où on retrouve un codon GTA à cette position.

La séquence nucléotidique de *c-yes* (humain) serait la plus homologue au clone 060e selon l'analyse FastA. On assume que les séquences nucléotidiques de *c-yes* humain et de *c-yes* murin sont identiques, la séquence murine n'étant pas répertoriée dans GenBank. L'alignement de séquences nucléotidiques de *c-yes* et du clone 060e révèle un pourcentage d'identité de 90.5% pour une région de 211 nucléotides de chevauchement alors qu'il est de 98.5% d'identité pour une région de 68 acides aminés. La différence entre le nombre de nucléotides et d'acides aminés comptabilisé dans la région d'homologie résulte de la séquence imposée par les amorces. La différence d'identité entre la séquence peptidique et nucléotidique (excluant le changement mentionné précédemment) est uniquement causée par la dégénérescence (*wobble*) observée en troisième position de 16 codons. Il y a donc 24% des codons (16/[68-1]) où une telle dégénérescence est observée (figure 17). Il existe probablement un mécanisme entraînant un taux de substitution très élevé chez ces gènes.

En regard des résultats présentés, le clone 060e est donc un nouveau membre de la famille *src* qui présente un domaine catalytique très analogue à celui de *c-yes*. La différence observée entre les séquences nucléotidiques supporte nos résultats.

3.4 Un nouveau membre de la sous-famille *eph*

3.4.1 Les membres de la sous-famille *eph* sont exprimés spécifiquement dans certains tissus.

Compte tenu de l'homologie partagée entre la séquence nucléotidique du clone 113e et chacun des principaux membres de la sous-famille *eph* (en moyenne 72% pour 212 nu) ainsi que de l'homologie retrouvée avec son plus proche homologue, *Cek-5*, nous considérons le clone 113e comme nouveau membre de

cette sous-famille.

La sous-famille *eph* est constituée de PTKs de type récepteur possédant une expression tissulaire spécifique. Par hybridation, on retrouve *eph* au niveau du foie et des poumons (Maruet *al.*, 1988); *eck* chez les cellules épithéliales de reins et de poumons (Lindberg et Hunter, 1990); *elk* au niveau du cerveau (expression principale) et des testicules (Lhoták et al, 1991); *sek* est exprimée uniquement au niveau du rhombencéphale en voie de développement (Gilardi-Hebenstreit *et al.*, 1992); *eek* est exprimée au niveau du cerveau; et *eck* dans les poumons (Chan et Watt, 1991). L'expression localisée de ces protéines kinases a été confirmée soit par immunofluorescence, par hybridation *in situ*, soit par buvardage de type *Northern* ou *Southern* dans le cas de criblage de banques de cDNA spécifiques à des tissus.

Structuralement, les membres de la sous-famille *eph* sont caractérisés par la présence d'une région riche en résidus cystéine dans le domaine extracellulaire et par une absence d'insertion dans le domaine catalytique (Maru *et al.*, 1988) contrairement à plusieurs autres PTK de type récepteur (e.g. PDGF-R). L'absence de l'insertion définit cette nouvelle sous-famille, soit la quatrième.

L'expression localisée des membres de la famille *eph* ainsi que la présence d'un domaine intramembranaire hydrophobe ajoute à l'intérêt que représente cette nouvelle kinase. Il serait pertinent d'effectuer des expériences d'hybridation *in situ* afin de déterminer si son expression *in vivo* dans le thymus est significative et si elle est localisée uniquement dans cet organe. De plus, la présence d'un domaine hydrophobe suggère son association à la membrane plasmique et par ailleurs, ses sites de glycosylation potentiels permettent de la comparer aux glycoprotéines composant le complexe gp23/45/90.

3.4.2 Particularités moléculaires du clone 113e

Le clone 113e présente, tout comme le clone 060e, des caractéristiques particulières au niveau de sa séquence nucléotidique. Malgré le niveau élevé d'homologie protéique que présente ce clone avec les membres de la sous-famille *eph*, il existe des différences surprenantes au niveau de la séquence nucléotidique. Plusieurs acides aminés localisés hors des motifs caractéristiques des PTKs sont conservés entre les membres de cette sous-famille. Par contre, la séquence nucléotidique pour ces résidus conservés diverge énormément.

Afin de mieux apprécier cette variation, examinons l'alignement entre le clone 113e et *Cek-5* (figure 20). Pour s'assurer de ne pas biaiser la comparaison, nous faisons abstraction des résidus provenant des amorces dégénérées durant le PCR, celles-ci imposant des contraintes de séquence pouvant résulter en quelques appariements non-représentatifs de la séquence originale. Sur un total de 73 acides aminés, 15 sont donc éliminés de nos calculs. Sur les 58 résidus internes aux amorces, 21 possèdent de la dégénérescence en troisième position du codon (*wobble*), soit environ 36% des résidus.

La dégénérescence s'étend jusqu'aux motifs conservés des PTK. En effet, *erk* et *EEK* ne possèdent pas les mêmes codons que les autres membres pour les trois acides aminés constituant le motif -DFG-.

Tous les membres de cette sous-famille possèdent un motif VIII partagé uniquement avec les homologues du récepteur à insuline. À cette position, la substitution d'un résidu lysine pour un résidu arginine semble provenir, *a priori*, d'une mutation ponctuelle transformant ainsi le codon AAa/g en AGa/g. La séquence nucléotidique du clone 113e montre qu'il s'agirait plutôt d'un changement du codon AAG en CGA. En effet, les substitutions observées dans ce motif démontrent qu'il n'existe pas d'explication simple. Quatre codons sont

retrouvés à cette position chez les membres de cette sous-famille: AGA, CAG, CGC et CGG.

Ce motif est exposé à des pressions causant des mutations. Par contre, les pressions conservatrices, similaires à celles que l'on retrouve chez la famille *src*, sont plus importantes. Le motif VIII et les résidus l'entourant sont strictement conservés sauf pour *elk* qui présente un résidu valine en position 788 alors que les autres, incluant le clone 113e, possèdent un résidu isoleucine. Bien que la séquence peptidique de cette région soit conservée, la situation est très différente pour la séquence nucléotidique. Sur douze résidus composant la région étudiée, en moyenne 50% des résidus ont une variation silencieuse en position 3' du codon (*wobble*). Cette variation s'étend de 3 résidus sur 12 pour *eck* jusqu'à 7 résidus sur 12 pour *elk* et *sek*. Les substitutions apparaissent sur plusieurs résidus. En fait, d'une façon globale, 10 résidus sur 12 ont de telles substitutions. Seuls les résidus lysine et tryptophane ne démontrent pas de changements. L'importance du résidu tryptophane (codon TGG), a été démontrée par alignement multiple où on le retrouve chez toutes les PTKs. Le nombre de substitutions rencontrées pour chaque résidus est indiqué à la figure 28. Nous avons utilisé seulement les cinq alignement de la figure 29.

région entourant le motif VIII du clone 113e

	140		150		160							
GGC	AAG	ATT	CCC	ATC	CGA	TGG	ACC	GCC	CCT	GAG	GCC	
G	K	I	P	I	R	W	T	A	P	E	A	
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
2	0	4	3	(1)	3	0	3	2	4	1	1	

nombre de fois que le changement est observé

Figure 28: Substitutions silencieuses au motif VIII du clone 113e. Le nombre de substitutions silencieuses observées en position 3' des codons de la région entourant le motif VIII est indiqué. Quatre membres de la famille *eph* sont alignés avec le clone 113e.

```

Erk Humain
      140      150      160
113e.D GGCAAGATTCCCATCCGATGGACCGCCCCTGAGGCC
      || ||||| ||||| ||||| ||||| || |||
Humerk GGAAAGATCCCCATCCGCTGGACAGCCCCGGAAAGCC
      450      460      470
6/12

Eck Humain
      140      150      160
113e.D GGCAAGATTCCCATCCGATGGACCGCCCCTGAGGCC
      ||||| ||||| ||||| ||||| |||||
Humeck GGCAAGATCCCCATCCGCTGGACCGCCCCGGAGGCC
      2450     2460     2470
3/12

Elk Rat
      140      150      160
113e.D GGCAAGATTCCCATCCGATGGACCGCCCCTGAGGCC
      || ||||| || | ||||| || || |||||
Ratelk GGGAAGATACCTGTTAGATGGACAGCTCCAGAGGCC
      2720     2730     2740     2750
7/12

Sek Murin (Cek-8 Aviaire)
      140      150      160
113e.D GGCAAGATTCCCATCCGATGGACCGCCCCTGAGGCC
      ||||| ||||| ||||| || || || ||
Mmsek GGCAAGATTCCCTATCCGGTGGACTGCGCCAGAAGCA
      2410     2420     2430     2440
7/12

Eek Rat
      140      150      160
113e.D GGCAAGATTCCCATCCGATGGACCGCCCCTGAGGCC
      || ||||| || ||||| || || |||||
Rneekr GGGAAGATCCCTATCCGATGGACAGCACCAGAGGCC
      500      510      520
5/12

```

Figure 29: Alignements des motifs VIII (-RWTAPE-) entre le clone 113e et les principaux membres de la sous-famille *eph*. Les alignements ont été obtenus à l'aide du logiciel FASTA. Les résidus substitués n'entraînant pas de changement d'acide aminé (*wobble*) sont indiqués en caractères gras. Le seul acide aminé modifié se trouve chez *elk* où on retrouve le motif -KIPVRWTAPE-. Ce gène possède également une substitution silencieuse en position 1 (résidu 2732). Le nombre d'acides aminés où l'on retrouve une substitution en position 3' des codons est indiqué à droite.

4. Résumé et conclusion générale

Le nombre de protéine kinases augmente sans cesse. Il y a environ 15 ans, on ne comptait qu'une douzaine de protéine kinases caractérisées. Lors de la dernière parution de la banque de données renfermant les domaines catalytiques des protéine kinases (version avril 1993), 13 nouvelles kinases étaient répertoriées. La version antérieure datait du mois de janvier 1993 où 23 nouvelles kinases apparaissaient. Ce nombre élevé suggère l'existence d'une pression exercée sur ces gènes qui entraîne une duplication de ceux-ci. En effet, l'étude phylogénétique des protéine kinases démontre clairement l'existence d'un ou plusieurs gènes ancestraux (Hanks *et al.*, 1988).

Les résultats obtenus, particulièrement pour les clones 060e et 113e, suggèrent que de très grandes pressions sont exercées sur ces gènes, les forçant à incorporer des substitutions au fil des années. Une théorie expliquant l'apparition spontanée de substitutions a été élaborée originalement par Watson et Crick puis modifiée par Tobal et Fresco. La théorie repose sur les équilibres tautomériques, spécialement sur la forme énol de la thymine qui peut ainsi s'apparier avec une guanine. Les dernières constatations ont démontré que la tautomérisation céto-énol, amino-imino ou anti-syn (rotation du nucléotide autour du lien glycosidique de sa position normale, anti, à une position peu fréquente, syn) permettent à une double hélice d'accomoder une complémentarité non-conventionnelle sans distorsion majeures (e.g. A/C imino, G/T énol, A imino/C, G énol/T, A imino/A syn, etc). Il est à noter que dans de telles conditions, les liaisons purine-purine sont possibles alors que les liaisons pyrimidine-pyrimidine ne le sont pas. La fréquence calculée de ces appariements fautifs est de 1 par 10^4 - 10^5 pour les formes imino/énol et de 1 par 10-20 pour la forme syn. L'environnement immédiat des nucléotides ainsi que les mécanismes de validation d'appariements naturels retrouvés chez les enzymes influencent ces résultats. Par contre, ces mécanismes

n'expliquent probablement qu'une partie des substitutions retrouvées chez les PTKs puisqu'il en existe un nombre très élevé entre les membres d'une même sous-famille.

Par ailleurs, le nombre élevé de duplications de gènes entraîne sûrement l'apparition de pseudogènes. Ceux-ci peuvent être transcrits et donner lieu à des résultats difficilement interprétables dans le cas de la méthode d'amplification utilisée pour nos travaux. Lors du séquençage que nous avons effectué, plusieurs clones possédaient une région VI-IX partielle de taille légèrement plus élevée que celle attendue. Plusieurs de ces fragments ont démontré une identité avec des séquences n'ayant aucun lien avec des protéines kinases. À quelques reprises, ces clones ont été alignés avec des régions intergéniques humaines répertoriées dans GenBank. Ces résultats suggèrent que la méthode utilisée permet le clonage fortuit de pseudogènes de PTKs transcrits tout comme pour la β -actine (Menon *et al.*, 1991).

De plus, les pseudogènes peuvent avoir un rôle dans la régulation de la transcription des gènes de PTKs. Une analogie est retrouvée chez les oligodendrocytes déficients en myéline. En effet, la juxtaposition de deux locus MBP (*Myelin basic protein*) en tandem où le locus en amont possède une inversion majeure de sa région 3' entraîne la production de hnRNA antisense. Le promoteur en amont du tandem n'est pas altéré et la transcription débute au premier locus du tandem produisant ainsi une quantité abondante de transcrit antisense (Fremeau et Popko, 1990). En regard de ces observations, l'utilisation d'amorces inverses pour l'amplification des PTKs permettrait de déceler la présence de transcrits antisense.

En résumé, nos travaux ont démontré la présence de trois nouvelles PTKs chez les cellules E-5. De plus, les cellules E-5 possèdent des transcrits codant pour onze PTKs connues. Des trois nouvelles PTKs, l'existence de celle

correspondant au clone 109c est la moins certaine à cause de la présence simultanée de deux motifs normalement retrouvés chez des sous-familles distinctes. D'autre part, nous proposons que le clone 060e code pour une PTK appartenant à la sous-famille *src*. La divergence de la séquence nucléotidique entre 060e et son plus proche homologue, *c-yes*, suggère que 060e n'est pas une isoforme. Finalement, nos travaux indiquent que le clone 113e provient d'un domaine catalytique d'une PTK apparentée à la sous-famille *eph*. La séquence de 113e n'a jamais été répertoriée jusqu'ici. De plus, l'intérêt des PTKs de cette sous-famille provient de leur expression localisée. Il reste à déterminer si le transcrit de cette nouvelle PTK est spécifiquement localisé dans les cellules épithéliales du thymus.

Parmi les onze PTKs connues retrouvées chez la E-5, la présence de *RYK* permet d'établir certaines corrélations avec la gp90 associée au complexe gp23/45/90. Les caractéristiques biochimiques des deux protéines suggèrent qu'elles peuvent être apparentées. Il sera prioritaire de vérifier si le gène *RYK* est transcrit et traduit *in vivo*, dans le thymus.

Bien que nos résultats suggèrent un rôle important pour les PTKs des cellules épithéliales thymiques, des expériences d'hybridation *in situ* permettront de vérifier leur présence dans un thymus normal. Ces travaux permettront d'établir l'importance relative des différentes PTKs des cellules E-5 dans un contexte thymique et ultérieurement, d'associer une fonction à leur présence.

REMERCIEMENTS

Le présent travail n'aurait pu être accompli sans le support de mes directeurs de recherche, MM. Edouard Potworowski et Peter Tijssen. Je suis très reconnaissant pour la patience, le dévouement et la confiance qu'ils ont démontrés au cours des deux dernières années. J'aimerais également remercier mon collègue et ami, Jean Bergeron, sans qui je n'aurais pu apprivoiser la biologie moléculaire.

Plusieurs personnes ont contribué à la réalisation de ce projet. J'aimerais remercier Pierre Chrétien qui a su m'initier aux plaisir du séquençage automatique ainsi qu'à Joanne Roger qui a effectué une partie des travaux de séquençage. J'aimerais également remercier MM. Pravin Patel et François Shareck pour l'utilisation du logiciel GCG ainsi que pour leur encouragement. De plus, je remercie Anne Marie Quinn et Steven Hanks sans qui le travail d'analyse aurait été encore plus laborieux.

Finalement, un gros merci à Claire Beauchemin pour son aide technique et principalement pour son support quotidien indispensable à la réussite d'un étudiant.

BIBLIOGRAPHIE

- ADKINS, B., C. MUELLER, C.Y. OKADA, R.A. REICHERT, I.L. WEISSMAN et G.J. SPANGRUDE 1987. Early events in T-cell maturation. *Ann. Rev. Immunol.* 5: 325-365.
- ALTING-MEESE, M.A. et J.M. SHORT 1989. pBluescript II: gene mapping vectors. *Nucl. Acids Res.* 17: 9494.
- ANDERSON, N.G., J.L. MALLER, N.K. TONKS et T.W. STURGILL 1990. Requirement for integration of signals from two distinct phosphorylation pathways for activation of MAP kinase. *Nature* 343: 651-653.
- ANDREWS, P. et R.L. BOYD 1985. The murine thymic nurse cell: an isolated thymic microenvironment. *Eur. J. Immunol.* 15: 36-42.
- ANDREWS, P., K. SHORTMAN, R. SCOLLAY, E.F. POTWOROWSKI, A.M. KRUISBEEK, G. GOLDSTEIN, N. TRAININ, et J.F. BACH 1985. Thymus hormones do not induce proliferative ability or cytolytic function in PNA⁺ cortical thymocytes. *Cell. Immunol.* 9:455-466.
- ARASE, H., N. FUKUSHI, S. HATAKEYAMA, K. OGASAWARA, K. IWABUCHI, C. IWABUCHI, I. NEGISHI, R.A. GOOD et K. ONOE 1990. Sequential analysis of the thymocyte differentiation in fully allogeneic bone marrow chimera in mice. II. Further characterization of the CD4⁺ or CD8⁺ single positive thymocytes. *Immunobiol.* 180: 167- 183.
- ARDAVIN, C., L. WU, C.L. LI et K. SHORTMAN 1993. Thymic dendritic cells and T cells develop simultaneously in the thymus from a common precursor population. *Nature* 362: 761-763.
- AUERBACH, R. 1961. Morphogenetic interactions in the development of the mouse thymus. *Dev. Biol.* 2:271-284.
- BARCLAY, N.A. et G. MAYRHOFER 1981. Bone marrow origin of Ia-positive cells in the medulla of rat thymus. *J. Exp. Med.* 153:1666-1671.
- BARKER, W.C. et M.O. DAYHOFF 1982. Viral *src* gene products are related to the catalytic chain of mammalian cAMP-dependant protein kinase. *Proc. Natl. Acad. Sci. USA* 79: 2836-2839.
- BEJ, A.K., M.H. MAHBUBANI et R.M. ATLAS 1991. Amplification of nucleic acids by polymerase chain reaction (PCR) and other methods and their applications. *Crit. Rev. Biochem. & Mol. Biol.* 26: 301-334.
-

BENOIST, C. et D. MATHIS 1989. Positive selection of the T cell repertoire: where and when does it occur? *Cell* **58**: 1027-1033.

BERGMAN, M., T. MUSTELIN, C. OETKEN, J. PARTANEN, N.A. FLINT, K.E. AMREIN, M. AUTERO, P. BURN et K. ALITALO 1992. The human p50^{csk} tyrosine kinase phosphorylates p56^{lck} at tyr-505 and down regulates its catalytic activity. *EMBO J.* **11**: 2919-2924.

BERTICS, P.J. et G.N. GILL 1985. Self-phosphorylation enhances the protein-tyrosine kinase activity of the epidermal growth factor receptor. *J. Biol. Chem.* **260**: 14642-14647.

BLACKMAN, M., J. KAPPLER et P. MARRACK 1990. The role of the T-cell receptor in positive and negative selection of developing T-cells. *Science* **248**: 1335-1341.

BIRNBOIM, H.C. et J. DOLY 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Res.* **7**: 1513-1523.

BLUESTONE, J.A., D. PARDOLL, S.O. SHARROW et B.J. FOWLKES 1987. Characterization of murine thymocytes with CD3-associated T-cell receptor structures. *Nature* **326**: 82-84.

BOSMA, G.C., R.P. CUSTER et M.J. BOSMA 1983. Severe combined immunodeficiency in the mouse. *Nature* **301**: 527-530.

BRENNER, S. 1987. Phosphotransferase sequence homology. *Nature* **329**:21.

BRYANT, D. et J.T. PARSONS 1983. Site-directed point mutation in the *src* gene of rous sarcoma virus results in an inactive *src* gene product. *J. Virol.* **45**: 1211-1216.

BUSS, J.E. AND SEFTON, B.M. 1985. Myristic acid, a rare fatty acid, is the lipid attached to the transforming protein of Rous sarcoma virus and its cellular homolog. *J. Virol.* **53**:7-12.

CANTLEY, L.C., K.R. AUGER, C. CARPENTER, B. DUCKWORTH, A. GRAZIANI, R. KAPPELLER et S. SOLTOFF 1991. Oncogenes and signal transduction. *Cell* **64**: 281-302.

CARLOW, D.A., N.S. VAN OERS, S.J. TEH et H.S. TEH 1992. Deletion of antigen-specific immature thymocytes by dendritic cells requires LFA-1/ICAM interactions. *J. Immunol.* **148**: 1595-1603.

CHAN, J., et V.M. WATT 1991. *eeb* and *erk*, new members of the subclass of receptor protein-tyrosine kinases. *Oncogene* 6: 1057-1061.

CHAN, L.C., K.K. KARHI, S.I. RAYTER, N. HEISTERKAMP, S. ERIDANI, R. POWLES, S.D. LAWLER, J. GROFFEN, J.G. FOULKES, M.F. GREAVES et L.M. WIEDEMANN 1987. A novel *abl* protein expressed in Philadelphia chromosome positive acute lymphoblastic leukemia. *Nature* 325: 635-637.

CHIRGWIN, J. M., A. E. PRZYBYLA, R. J. MACDONALD et W. J. RUTTER 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*. 24:5294-5299.

CHOTHIA, C., et A.M. LESK 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J.* 5: 823-826.

CHOU, P.Y., et G.D. FASMAN 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol, & Related Areas Mol. Biol.* 47: 45-148.

CHOU, Y.H., et M.J. HAYMAN 1991. Characterization of a member of the immunoglobulin gene superfamily that possibly represents an additional class of growth factor receptor. *Proc. Natl. Acad. Sci. USA* 88: 4897-4901.

CICCHETTI, P., B.J, MAYER, G, THEIL et D, BALTIMORE 1992. Identification of a protein that binds to the SH3 region of *abl* and is similar to Bcr and GAP-rho. *Science* 257: 803-806.

CLARK, S.G., M.J. STERN et H.R. HORVITZ 1992. *C. elegans* cell- signalling gene *sem-5* encodes a protein with SH2 and SH3 domains. *Nature* 356: 340-344.

COCHET, C., G.N. GILL, J. MEISENHEDER, J.A. COOPER et T. HUNTER 1984. C-kinase phosphorylates the epidermal growth factor receptor and reduces its epidermal growth factor-stimulated protein kinase activity. *J. Biol. Chem.* 261: 350-355.

COHEN, J.J. 1990. Programmed cell death in the immune system. *Advances in Immunology*, vol 50, 55-85.

COHEN, S., H USHIRO, C STOSCHECK et M CHINKERS 1982. A native 170,000 epidermal growth factor receptor-kinase complex from shed plasma membrane vesicles. *J. Biol. Chem.* 257: 1523-1531.

COMPTON, T. 1990. "Degenerate primers for DNA amplification", dans Innis, M.A., D.H. Gelfand, J.J. Sninsky et T.J. White (éd.): PCR Protocols: A Guide to Methods and Applications. Academic Press, 39-45.

COOPER, N.A., K.L. GOULD, C.A. CARTWRIGHT et T. HUNTER 1986. Tyr527 is phosphorylated in pp60c-src: implications for regulation. *Science* 231: 1431-1433.

COUTURE, C., P.C. PATEL et E.F. POTWOROWSKI 1990. A novel thymic epithelial adhesion molecule. *Eur. J. Immunol.* 20: 2769-2773.

COUTURE, C., G. AMARANTE-MENDES et E.F. POTWOROWSKI 1992. Tyrosine kinase activation in thymic epithelial cells: necessity of thymocyte contact through the gp23/45/90 adhesion complex. *Eur. J. Immunol.* 22: 2579-2585.

CROSS, F.R., E.A. GARBER, D. PELLMAN et H. HANAFUSA 1985. A short sequence in the p60^{src} N terminus is required for p60^{src} Myristylation and membrane association and for cell transformation. *Mol. Cell. Biol.* 4:1834-1842.

D'AQUILA, R.T., L.J. BLECHTEL, J.A. VIDELER, J.J. ERON, P. GORCZYCA et J.C. KAPLAN 1991. Maximizing sensitivity and specificity of PCR by pre-amplification heating. *Nucleic Acids Res.* 19: 3749.

DAVIS, R.L., J.B. KONOPKA et O.N. WITTE 1985. Activation of the c-abl oncogene by viral transduction or chromosomal translocation generates c-abl proteins with similar in vitro kinase properties. *Mol. Cell. Biol.* 5: 204-213.

DE KLEIN, A., A.G. VAN KESSEL, G. GROSVELD, C.R. BARTRAM, A. HAGEMEIJER, D. BOOTSMA, N.K. SPURR, N. HEISTERKAMP, J. GROFFEN et J.R. STEPHENSON 1982. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. *Nature* 300: 765-767.

DE WAAL-MALEFIJT, R, W. LEENE, P.J.M. ROHOLL, J. WORMMEESTER et K.A. HOEBEN 1986. T cell differentiation within thymic nurse cells. *Lab. Invest.* 55: 25-34.

DENNING, S.M., J. KURTZBERG, P.T. LE, D.T. TUCK, K.H. SINGER et B.F. HAYNES 1988. Human thymic epithelial cells directly induce activation of autologous immature thymocytes. *Proc. Natl. Acad. Sci. USA.* 85: 3125-3129.

DENNING, S.M., D.T. TUCK, K.H. SINGER et B.F. HAYNES 1987. Human thymic epithelial cells function as accessory cells for autologous mature thymocyte activation. *J. Immunol.* 138:680-686.

DEVEREUX, J., P. HAEBERLI et O. SMITHIES 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12: 387- 395.

DEXTER, T.M. et J.M. AUSTYN 1974. Proliferation of hematopoietic stem cells in vitro. *Br. J. Haematol.* 28: 525-530.

DOOLITTLE, R.F. 1990. Searching through sequence databases. *Meth. Enzymol.* 183: 99-110.

DORSHKIND, K, G.M. KELLER, R.A. PHILLIPS, R.G. MILLER, G.C. BOSMA, M. O'TOOLE et M.J. BOSMA 1984. Functional status of cells from lymphoid and myeloid tissues in mice with severe combined immunodeficiency disease. *J. Immunol.* 132: 1804-1808.

DUIJVESTIJN, A.M. et E.C.M. HOEFSMIT 1981. Ultrastructure of the rat thymus: the microenvironment of T lymphocyte maturation. *Cell Tissue Res.* 218:279-292.

DUIJVESTIJN, A.M., R. SCHUTTE, Y.G. KOHLER, G. KORN, et E.C.M. HOEFSMIT 1983. Characterization of the population of phagocytic cells in thymic cell suspensions. A morphological and cytochemical study. *Cell Tissue Res.* 231:313-323.

DUMONT, D.J., T.P. YAMAGUCHI, R.A. CONLON, J. ROSSANT et M.L. BREITMAN 1992. *tek*, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in endothelial cells and their presumptive precursors. *Oncogene* 7: 1471-1480.

DUSTIN, M.L., R. ROTHLEIN, A.K. BHAN, C.A. DINARELLO et T.A. SPRINGER 1986. Induction by IL-1 and interferon-gamma: Tissue distribution, biochemistry, and function of a natural adherence molecule (ICAM-1). *J. Immunol.* 137: 245-254.

ECKERT, K.A. et T.A. KUNKEL 1990. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* 18: 3739-3744.

ERIKSON, R.L., M.S. COLLETT, E. ERIKSON et A.F. PURCHIO 1979. Evidence that the avian sarcoma virus transforming gene product is a cyclic AMP-independant protein kinase. *Proc. Natl. Acad. Sci. USA* 76: 6260-6264.

FAIRCHILD, P.J. et J.M. AUSTYN 1990. Thymic dendritic cells: phenotype and function. *Int. Rev. Immunol.* 6: 187-196.

FARR, A.G., A. NELSON, J. TRUEX et S. HOSIER 1991. Epithelial heterogeneity in the murine thymus - A cell surface glycoprotein expressed by subcapsular and medullary epithelium. *J. Histochem. & Cytochem.* 39: 645-653.

FARR, A.G., S. HOSIER, S.C. BRADDY, S.K. ANDERSON, D.J. EISENHARDT, Z.J. YAN, et C.P. ROBLES 1989. Medullary epithelial cell lines from murine thymus constitutively secrete IL-1 and hematopoietic growth factors and express class II antigens in response to recombinant Interferon- γ . *Cell. Immunol.* 119:427-444.

FARR, A.G. et S.C. BRADDY 1989. Patterns of keratin expression in the murine thymus. *Anat. Rec.* 224:374-378.

FENG, D. et R.F. DOOLITTLE 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25: 351-360.

FOWLKES, B.J., R.H. SCHWARTZ et D.M. PARDOLL 1988. Deletion of self-reactive thymocytes occurs at a CD4⁺8⁺ precursor stage. *Nature* 334:620-623.

FREMEAU, R.T.J. et B POPKO 1990. In situ analysis of myelin basic protein gene expression in myelin-deficient oligodendrocytes: antisense hnRNA and readthrough transcription. *Embo J.* 9: 3533-3538.

FROHMAN, M.A., M.K. DUSK et G.R. MARTIN 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85: 8998-9002.

FUJIWARA, H., M. OGATA, Y. MIZUSHIMA, Y. TATSUMI, Y. TAKAI et T. HAMAOKA 1990. Proliferation and differentiation of immature thymocytes induced by a thymic stromal cell clone. *Thymus* 16: 159-172.

GAILLARD, C. et F. STRAUSS 1990. Ethanol precipitation of DNA with linear polyacrylamide as carrier. *Nucleic Acids Res.* 18: 378.

GALY, A.H.M., C.A. DINARELLO, T.S. KUPPER, A. KAMEDA et J.W. HADDEN 1990. Effects of cytokines on human thymic epithelial cells in culture .2. Recombinant IL-1 stimulates thymic epithelial cells to produce IL-6 and GM-CSF. *Cell. Immunol.* 129: 161-175.

GARVIN, A.M., S. PAWAR, J.D. MARTH et R.M. PERLMUTTER 1988. Structure of the murine lck gene and its rearrangement in a murine lymphoma cell line. *Mol. Cell Biol.* 8: 3058-3064.

GEORGE, D.G., W.C. BARKER et L.T. HUNT 1990. Mutation data matrix and its uses. *Meth. Enzymol.* 183: 333-351.

GILARDI-HEBENSTREIT, P., M.A. NIETO, M. FRAIN, M.G. MATTÉI, A. CHESTIER, D.G. WILKINSON et P. CHARNAY 1992. An eph-related receptor protein tyrosine kinase gene segmentally expressed in the developing mouse hindbrain. *Oncogene* 7: 2499-2506.

GOFF, S.P., E. GILBOA, O.N. WHITE et D. BALTIMORE 1980. Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. *Cell*. 22: 777-785.

GUAN, C., P. LI, P.D. RIGGS et H. INOUE 1987. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 67: 21-30.

HAN, J. H., C. STATOWA, et W.J. RUTTER 1987. Isolation of full-length putative rat lysophospholipase cDNA using improved methods for mRNA isolation and cDNA cloning. *Biochemistry*. 26:1617-1625.

HANKS, S.K., A.M. QUINN, et T. HUNTER 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241:42-52.

HANKS, S.K. et A.M. QUINN 1991. Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* 200: 38-62.

HARDIE D.G. 1989. Protein phosphorylation and dephosphorylation. *Current Opinion in Cell Biology* 1: 220-226.

HARPUR, A.G., A.C. ANDRES, A. ZIEMIECKI, R.R. ASTON et A.F. WILKS 1992. JAK2, a third member of the JAK family of protein tyrosine kinases. *Oncogene* 7: 1347-1353.

HÉBERT, B., J. BERGERON, E.F. POTWOROWSKI et P. TIJSSEN 1993. Increased PCR sensitivity by using paraffin wax as a reaction mix overlay. *Mol. Cell. Probes* 7: 249-252.

HELIN, K., T. VELU, P. MARTIN, W.C. VASS, G. ALLEVATO, D.R. LOWY et L. BEGUINOT 1991. The biological activity of the human epidermal growth factor receptor is positively regulated by its c-Terminal tyrosines. *Oncogene* 6: 825-832.

HIRAMINE, C., K. HOJO, M. KOSETO, T. NAKAGAWA et A. MUKASA 1990. Establishment of a murine thymic epithelial cell line capable of inducing both thymic nurse cell formation and thymocyte apoptosis. *Lab. Invest.* 62: 41-54.

HOMO-DELARCHE, F. et M. DARDENNE 1993. The neuroendocrine-immune axis. Springer Sem. Immunopathol. 14: 221-238.

HONNEGGER, A.M., T.J. DULL, D. SZAPARY, A. KOMORIYA, R. KRIS, A. ULLRICH et J. SCHLESSINGER 1988. Kinetic parameters of the protein tyrosine kinase activity of EGR-receptor mutants with individually altered autophosphorylation sites. EMBO J. 7: 3053- 3060.

HONNEGGER, A.M., T.J. DULL, F. BELLOT, E. VAN OBBERGHEN, D. SZAPARY, A. SCHMIDT, A. ULLRICH et J. SCHLESSINGER 1988. Biological activities of EGR-receptor mutants with individually altered autophosphorylation sites. EMBO J. 7: 3045-3052.

HOVENS, C.M., S.A. STACKER, A.C. ANDRES, A.G. HARPUR, A. ZIEMIECKI et A.F. WILKS 1992. *RYK*, a receptor tyrosine kinase-related molecule with unusual kinase domain motifs. Proc. Natl. Acad. Sci. USA 89: 11818-11822.

HRADETZKY, D., K. STREBHARDT et H. RUESAMEN-WAIGMANN. 1992. The genomic locus of the human hemopoietic-specific cell protein tyrosine kinase (PTK)-encoding gene (*HCK*) confirms conservation of exon-intron structure among human PTKs of the *scr* family. Gene 113: 275-280.

HUGO, P. et E.F. POTWOROWSKI 1989. Dynamics of complex formation between thymocytes and thymic medullary epithelial cells. Scand. J. Immunol. 29:399-408.

HUGO, P. et E.F. POTWOROWSKI 1990. Selection of CD4⁺CD8⁺ thymocytes by complex formation with medulla-derived epithelial cells. Cell. Immunol. 126:143-154.

HUGO, P., A. MITTAL et E.F. POTWOROWSKI 1989. Thymic lymphoma cells as a model for complex-formation between thymocytes and thymic medullary epithelial cells. Immunology Letters 22:73-77.

HUGO, P., Y. ST-PIERRE et E.F. POTWOROWSKI 1988. Characterization of the in vitro interaction between thymocytes and a medullary thymic epithelial cell line. Thymus 12: 27-37.

HUGO, P., G.A. WAANDERS, R. SCOLLAY, K. SHORTMAN et R. BOYD 1990. Ontogeny of a novel CD4⁺CD8⁻CD3⁻ thymocyte subpopulation: a comparison with CD4⁺CD8⁺CD3⁻ thymocytes. Int. Immunol. 2:209-218.

HUGO, P., G.A. WAANDERS, R. SCOLLAY, H.T. PETRIE et R.L. BOYD 1991. Characterization of Immature CD4⁺CD8⁻CD3⁻ Thymocytes. Eur. J. Immunol. 21: 835-838.

HUNT, P., D. ROBERTSON, D. WEISS, D. RENNICK, F. LEE et O.N. WITTE 1987. A single bone marrow-derived stromal cell type supports the in vitro growth of early lymphoid and myeloid cells. *Cell* 48: 997-1007.

HUNTER, T. et J.A. COOPER 1985. Protein-tyrosine kinases. *Ann. Rev. Biochem.* 54:897-930.

HUNTER, T. 1987. A thousand and one protein kinases. *Cell* 50: 823-829.

HUNTER, T. 1989. Protein modification: phosphorylation on tyrosine residues. *Cur. Op. Cell Biol.* 1: 1168-1181.

HUNTER, T. 1991. Protein kinase classification. *Methods Enzymol.* 200: 3-37.

ITOH, T., H. DOI, S. CHIN, T. NISHIMURA et S. KASAHARA 1988. Establishment of mouse thymic nurse cell clones from a spontaneous Balb/c thymic tumor. *Eur. J. Immunol.* 18: 821-824.

JACOBSEN, K., K. MIYAKE, P.W. KINCADE et D.G. OSMOND 1992. Highly restricted expression of a stromal cell determinant in mouse bone marrow in vivo. *J. Exp. Med.* 176: 927-935.

JENKINS, M.J., R.H. SCHWARTZ et D.M. PARDOLL 1988. Effects of cyclosporin A on T cell development and clonal deletion. *Science* 241: 1655-1658.

JENKINSON, E.J., R. KINGSTON, C.A. SMITH, G.T. WILLIAMS et J.J. OWEN 1989. Antigen-induced apoptosis in developing T cells: a mechanism for negative selection of the T cell receptor repertoire. *Eur. J. Immunol.* 19: 2175-2177.

KAMPS, M.P. et B.M. SEFTON 1986. Neither arginine nor histidine can carry out the function of lysine-295 in the ATP-binding site of pp60^{src}. *Mol. Cell. Biol.* 6: 751-757.

KAMPS, M.P., J.E. BUSS et B.M. SEFTON 1985. Mutation of NH₂-terminal glycine of p60^{c-src} prevents both myristoylation and morphological transformation. *Proc. Natl. Acad. Sci. USA* 82: 4625- 4628.

KAMPS, M.P. 1991. Determination of phosphoamino acid composition by acid hydrolysis of protein blotted to immobilon. *Methods Enzymol.* 201: 21-27.

KAUFMAN, D.L. et G.A. EVANS 1990. Restriction endonuclease cleavage at the termini of PCR products. *BioTechniques* 9: 304-306. [erratum 9:720].

KAZLAUSKAS, A., C. ELLIS, T. PAWSON et J.A. COOPER 1990. Binding of GAP to activated PDGF receptors. *Science* 247: 1578-1581.

KAZLAUSKAS, A. et J.A. COOPER 1989. Autophosphorylation of the PDGF receptor in the kinase insert region regulates interactions with cell proteins. *Cell* 58: 1121-1132.

KINCADE, P.W., G. LEE, C.E. PIETRANGELI, S. HAYASHI et J.M GIMBLE 1989. Cells and molecules that regulate B lymphopoiesis in bone marrow. *Ann. Rev. Immunol.* 7: 111-143.

KISIELOW, P., H. BLUTHMANN, U.D. STAERZ, M. STEINMETZ et H. VON BOEHMER 1988. Tolerance in T cell receptor transgenic mice involves deletion of nonmature CD4⁺CD8⁺ thymocytes. *Nature* 333:742-746.

KOCH, C.A., D. ANDERSON, M.F. MORAN, C. ELLIS et T. PAWSON 1991. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. *Science* 252: 668-674.

KOSAKA, H., M. OGATA, I. HIKITA, S. MARUO, S. SUGIHARA, H. MATSUBARA, Y. TAKAI, T. HAMAOKA et H. FUJIWARA. 1989. Model for clonal elimination in the thymus. *Proc. Natl. Acad. Sci. USA* 86: 3773- 3777.

KREBS, E.G., D.J. GRAVES et E.H. FISCHER 1959. Factors affecting the activity of muscle phosphorylase b kinase. *J. Biol. Chem.* 234: 2867-2873.

KYEWSKI, B.A. 1987. Seeding of thymic microenvironments defined by distinct thymocyte-stromal cell interactions is developmentally controlled. *J. Exp. Med.* 166: 520-538.

KYPTA, R.M., Y. GOLDBERG, E.T. ULUG et S.A. COUTNEIDGE 1990. Association between the PDGF receptor and members of the src family of tyrosine kinases. *Cell* 62: 481-492.

LAEMMLI, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685.

LAI, C. et G. LEMKE 1991. An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system. *Neuron* 6: 691-704.

LE, P.T., D.T. TUCK, C.A DINARELLO, B.F. HAYNES et S.K. SINGER 1987. Human thymic epithelial cells produce interleukin 1. *J. Immunol.* 138: 2520-2526.

- LE, JM, G. FREDRICKSON, M. POLLACK et J. VILCEK J 1989. Activation of thymocytes and T cells by interleukin-6. *Ann. N.Y. Acad. Sc.* 557: 444-453.
- LEVIN, D., C. HAMMOND, R. RALSTON et J.M. BISHOP 1987. Two yeast genes that encode novel protein kinases. *Proc. Natl. Acad. Sci. USA* 84: 6035-6039.
- LHOTÁK, V., P. GREER, K. LETWIN et T. PAWSON 1991. Characterization of *elk*, a brain-specific receptor tyrosine kinase. *Mol. Cell. Biol.* 11: 2496-2502.
- LINDBERG, R.A. et T. HUNTER 1990. cDNA cloning and characterization of *eck*, an epithelial cell receptor protein-tyrosine kinase in the *eph/elk* family of protein kinases. *Mol. Cell. Biol.* 10: 6316- 6324.
- LIPMAN, D.J. et W.R. PEARSON 1985. Rapid and sensitive protein similarity searches. *Science* 227: 1435-1441.
- LO, D., J. FREEDMAN, S. HESSE, R.L. BRINSTER et L. SHERMAN 1991. Peripheral Tolerance in Transgenic Mice - Tolerance to Class-II MHC and Non-MHC Transgene Antigens. *Immunol. Rev.* 122: 87-102.
- LO, D. et J. SPRENT 1986. Identity of cells that imprint H-2 restricted T- cell specificity in the thymus. *Nature* 319: 672-675.
- MACDONALD, H.R., R.C. BUDD, et R.C. HOWE 1988. A CD3- subset of CD4⁺8⁺ thymocytes: a rapidly cycling intermediate in the generation of CD4⁺8⁺ cells. *Eur. J. Immunol.* 18: 519-523.
- MACDONALD, H.R., H. HANGARTNER, et T. PEDRAZZINI 1988. Intrathymic deletion of self-reactive cells prevented by neonatal anti-CD4 antibody treatment. *Nature* 335:174-176.
- MACK, D., et J.J. SNINSKI 1988. A sensitive method for the identification of uncharacterized viruses related to known virus groups: hepadnavirus model system. *Proc. Natl. Acad. Sci. USA* 85: 6977-6981.
- MARGOLIS, B. 1992. Proteins with SH2 Domains - Transducers in the Tyrosine Kinase Signaling Pathway. *Cell Growth & Differentiation* 3: 73-80.
- MARU, Y., H. HIRAI, M.C. YOSHIDA et F. TAKATU 1988. Evolution, expression, and chromosomal location of a novel receptor tyrosine kinase gene, *eph*. *Mol. Cell. Biol.* 8: 3770-3776.

MATSUBARA, H., H. KOSAKA, S. SOGO, S. MARUO, S. SUGIHARA, Y. ASANO, Y. KOHNO, M. KIMOTO, G. SUZUKI, K. ISHIMURA, H. FUJITA, T. HAMAOKA et H. FUJIWARA 1990. T-Cell clones are killed by a thymic stromal cell monolayer following stimulation of T-cell receptor with antigen and or H-2 molecules on the monolayer. *Int. Immunol.* 2: 755-763.

MATSUDA, M., B.J. MAYER, Y. FUKUI et H. HANAFUSA 1990. Binding of transforming protein p47gag-crk to a broad range of phosphotyrosine-containing proteins. *Science* 248: 1537-1539.

MAYER, B.J., P.K. JACKSON, R.A. VANETTEN et D. BALTIMORE 1992. Point mutations in the abl SH2 domain coordinately impair phosphotyrosine binding in vitro and transforming activity in vivo. *Mol. Cell. Biol.* 12: 609-618.

MAYER, B.J., M. HAMAGUCHI et H. HANAFUSA 1988. A novel viral oncogene with structural similarity to phospholipase C. *Nature* 332: 272-275.

MENON R.S., Y.-F. CHANG, J. ST-CLAIR et R.G. HAM 1991. RT-PCR artifacts from processed pseudogenes. *PCR Methods and Applications* 1: 70-71.

MIZUSAWA, S., S. NISHIMURA et F. SEELA 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. *Nucleic Acids Res.* 14: 1319-1324.

MIZUTANI, S., S.M. WATT, D. ROBERTSON, S. HUSSEIN, L.E. HEALY et A.J.W. FURLEY 1987. Cloning of human thymic subcapsular cortex epithelial cells with T lymphocyte binding sites and hemopoietic growth factor activity. *Proc. Natl. Acad. Sci. USA* 84: 4999-5003.

MONSON, J.M., J. NATZLE, J. FRIEDMAN et B.J. McCARTHY 1982. Expression and novel structure of a collagen gene in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 79: 1761-1765.

MORGAN, D.O., J.M. KAPLAN, J.M. BISHOP et H.E. VARMUS 1989. Mitosis-specific phosphorylation of p60c-src by p34^{cdc-2}-associated protein kinase. *Cell* 57: 775-786.

MURPHY, K.M., A.B. HEIMBERGER et D.Y. LOH 1990. Induction by antigen of intrathymic apoptosis of CD4⁺CD8⁺TCR^{lo} thymocytes in vivo. *Science* 250: 1720-1723.

NAKAYAMA, K. et D.Y. LOH 1992. No requirement for p56lck in the antigen-stimulated clonal deletion of thymocytes. *Science* 257: 94-96.

- NEEDLEMAN, S.B. et C.D. WUNSCH 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48: 443-453
- OSMOND, D.G. 1986. Population dynamics of B lymphocytes. *Immunol. Rev.* 93: 103-124.
- PAPIERNIK, N., H. DOMBRET, S. STEFANOS et J. WIETZERBIN 1986. Control of Ia antigen expression on phagocytic cells of the thymic reticulum by interferon-gamma and prostaglandins. *Eur. J. Immunol.* 16: 296-300.
- PARK, M., M. DEAN, K. KAUL, M.J. BRAUN, M.A. GONDA et G. VANDE WOUDE 1987. Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors. *Proc. Natl. Acad. Sci. USA* 84: 6379-6383.
- PAWSON, T. et G. GISH 1992. SH2 and SH3 domains: from structure to function. *Cell* 71: 359-362.
- PENIT, C. et F. VASSEUR 1989. Cell proliferation and differentiation in the fetal and early postnatal mouse thymus. *J. Immunol.* 142: 3369-3377.
- POTWOROWSKI, E.F., F. TURCOTTE, C. BEAUCHEMIN, P. HUGO et M.G. ZELECHOWSKA 1986. Establishment and characterization of a thymic medullary epithelial cell clone. *In vitro cell. develop. biol.* 22: 557-560.
- POTWOROWSKI, E.F., P. HUGO, et C. COUTURE 1989. Binding of cortical thymocytes to a medullary epithelial cell line: a brief review. *Thymus* 13:237-243.
- POTWOROWSKI, E.F., L. THIBODEAU et M.G. ZELECHOWSKA 1986. Specific adherence of thymocytes to a thymic medullary epithelial cell line. *Immunol. Lett.* 13: 89-94.
- RAMER, S.E., D.G. WINKLER, A. CARRERA, T.M. ROBERTS et C.T. WALSH 1991. Purification and initial characterization of the lymphoid-cell protein-tyrosine kinase p56lck from a baculovirus expression system. *Proc. Natl. Acad. Sci. USA* 88: 6254-6258.
- RAPLEY, R. et M. WALKER 1992. PCR screening of DNA cloned into polylinker-containing vectors with M13 sequencing primers. *BioTechniques* 12: 516-517.
- RAULET, D.H., R.D. GARMAN, H. SAITO, et S. TONEGAWA 1985. Developmental regulation of T cell receptor gene expression. *Nature* 314:103-107.

RODAWAY, A.R.F., M.J.E. STERNBERG et D.L. BENTLEY 1989. Similarity in membrane proteins. *Nature* 342: 624.

RODAWAY, A.R.F., C.G. TEAHAN, C.M. CASIMIR, A.W. SEGAL et D.L. BENTLEY 1990. Characterization of the 47-kilodalton autosomal chronic granulomatous disease protein: tissue-specific expression and transcriptional control by retinoic acid. *Mol Cell Biol* 10: 5388-5396.

ROEHM, N., L. HERRON, J. CAMBIER, D. DIGUISTO, K. HASKINS, J. KAPPLER et P. MARRACK 1984. The major histocompatibility complex-restricted antigen receptor on T cells: Distribution on thymus and peripheral T cells. *Cell* 38: 577-584.

ROUER, E., T. VAN HUYNH, S.L. DE SOUZA, M.-C. LANG, S. FISCHER et R. BENAROUS. 1989. Structure of the human *lck* gene: Differences in genomic organisation within *src*-related genes affect only N-terminal exons. *Gene* 84: 105-113.

SAJJADI, F.G. et E.B. PASQUALE 1993. Five novel avian *eph*-related tyrosine kinases are differentially expressed. *Oncogene* 8: 1807-1813.

SALACINSKI, P.R.P., C. MCLEAN, J.E.C. SYKES, V.V. CLEMENT-JONES et P.J. LOWRY 1981. Iodination of proteins, glycoproteins, and peptides using a solid-phase oxidizing agent, 1,3,4,6-tetrachloro-3 α ,6 α -diphenyl glycoluril (Iodogen). *Anal. Biochem.* 117:136-146.

SAMBROOK, J., E.F. FRITSH et T. MANIATIS 1989. *Molecular Cloning: a laboratory manual*, second edition. Cold Spring Harbor Laboratory. 3 vol.

SANGER, F., S. NICKLEN, et A.R. COULSON, 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci., U.S.A.* 74: 5463-5467.

SCHIEVEN, G., J. THORNER et G.S. MARTIN 1986. Protein-tyrosine kinase activity in *Saccharomyces cerevisiae*. *Science* 231: 390- 393.

SCHULER, G.D., S.F. ALTSCHUL et D.J. LIPMAN 1991. A workbench for multiple alignment construction and analysis. *Proteins: Structure, Function, and Genetics* 9: 180-190.

SCOLLAY, R., A. WILSON et K. SHORTMAN 1984. Thymus cell migration: analysis of thymus emigrants with markers that distinguish medullary thymocytes from peripheral T cells. *J. Immunol* 132: 1089-1094.

SCOLLAY, R.G., E.C. BUTCHER et I.L. WEISSMAN 1980. Thymus cell migration: quantitative aspects of cellular traffic from the thymus to the periphery in mice. *Eur. J. Immunol.* 10: 210-218.

SCOTT, B., H. BLUTHMANN et H.S. TEH 1989. The generation of mature T cells requires interaction of the $\alpha\beta$ T cell receptor with major histocompatibility antigens. *Nature* 338: 591-593.

SHI, Y.F., R.P. BISSONNETTE, N. PARFREY, M. SZALAY, R.T. KUBO et D.R. GREEN May 15, 1991. *In vivo* administration of monoclonal antibodies to the CD3 T-cell receptor complex induces cell death (Apoptosis) in immature thymocytes. *J. Immunol.* 146: 3340-3346.

SHISHIDO, E., Y. EMORI et K. SAIGO 1991. Identification of seven novel protein-tyrosine kinase genes of *Drosophila* by the polymerase chain reaction. *FEBS letters* 2: 235-238.

SHORES, E.W., W. VAN EWIJK, et A. SINGER 1991. Disorganization and restoration of thymic medullary epithelial cells in T cell receptor-negative scid mice: evidence that receptor-bearing lymphocytes influence maturation of the thymic microenvironment. *Eur. J. Immunol.* 21:1657-1661.

SHORTMAN, K et H. JACKSON 1974. The differentiation of T lymphocytes. I. Proliferation kinetics and interrelationships of subpopulations of mouse thymus cells. *Cell. Immunol.* 12: 230-246.

SMITH, P.K., R.I. KROHN, G.T. HERMANSON, A.K. MALLIA, F.H. GARTNER, M.D. PROVENZANO, E.K. FUJIMOTO, N.M. GOEKE, B.J. OLSON et D.C. KLENK 1985. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150: 76-85. [erratum 163:279]

SMITH, C.A., G.T. WILLIAMS, R. KINGSTON, E.J. JENKINSON et J.J. OWEN 1989. Antibodies to CD3/T-cell receptor complex induce death by apoptosis in immature T cells in thymic cultures. *Nature* 337: 181-184.

SMITH, D.L., C.C. CHEN, B.B. BRUEGGER, S.L. HOLTZ, R.M. HALPERN et R.A. SMITH 1974. Characterization of protein kinases forming acid-labile histone phosphates in Walker-256 carcinosarcoma cells nuclei. *Biochemistry* 13: 3780-3785.

SNYDER, M., M. HUNKAPILLER, D. YUEN, D. SILVERT, J. FRISTROM et N. DAVIDSON 1982. Cuticle protein genes of *Drosophila*: structure, organization and evolution of four clustered genes. *Cell* 29: 1027-1040.

- SPEISER, D.E., E. KOLB, R. SCHNEIDER, H. PIRCHER, H. HENGARTNER, H.R. MACDONALD et R.M. ZINKERNAGEL 1989. Tolerance to Mls_a by clonal deletion of V-beta-6⁺ T-cells in bone marrow and thymus chimeras. *Thymus* 13: 27-33.
- SPRENGER, F., L.M. STEVENS et C. NÜSSLEIN-VOLHARD 1989. The *Drosophila* gene *torso* encodes a putative receptor tyrosine kinase. *Nature* 338: 478-483.
- SPRINGER, T.A., G. GALFRE, D.S. SECHTER et C. MILSTEIN 1979. MAC-1: a macrophage antigen identified by monoclonal antibody. *Eur. J. Immunol.* 9:301-306.
- STERNBERG, M.J.E. et W.R. TAYLOR 1984. Modelling the ATP-binding site of oncogene products, the epidermal growth factor receptor and related proteins. *FEBS Lett.* 175: 387-392.
- SWAT, W., L. IGNATOWICZ, H. VON BOEHMER et P. KISIELOW 1991. Clonal deletion of immature CD4⁺8⁺ thymocytes in suspension culture by extrathymic antigen-presenting cells. *Nature* 351: 150-153.
- TAKEYA, T. et H. HANAFUSA 1983. Structure and sequence of the cellular gene homologous to the RSV *src* gene and the mechanism for generating the transforming virus. *Cell* 32: 881-890.
- TIMSON GAUEN, L.K., A.N.T. KONK, L.E. SAMELSON et A.S. SHAW 1992. p59^{fyn} tyrosine kinase associates with multiple T-cell receptor subunits through its unique amino-terminal domain. *Mol. Cell. Biol.* 12: 5438-5446.
- TINDALL, K.R. et T.A. KUNKEL 1988. Fidelity of DNA synthesis by the *Termus aquaticus* DNA polymerase. *Biochemistry* 27: 6008-6013.
- ULLRICH, A. et J. SCHLESSINGER 1990. Signal transduction by receptors with tyrosine kinase activity. *Cell* 61: 203-212.
- VAN EWIIJK, W., R.V. ROUSE, et I.L. WEISSMAN 1980. Distribution of H-2 microenvironments in the mouse thymus. *J. Histochem. Cytochem.* 28:1089-1099.
- VAN EWIIJK, W. 1984. Immunohistology of lymphoid and non-lymphoid cells in the thymus in relation to T lymphocyte differentiation. *Am. J. Anat.* 170:311-330.
- VAN EWIIJK, W. 1988. Cell surface topography of thymic microenvironments. *Lab. Invest.* 59:579-590.

VAN EWIJK, W. 1991. T-cell differentiation is influenced by thymic microenvironments. *Ann. Rev. Immunol.* 9:591-615.

VAN VLIET, E., E.J. JENKINSON, R. KINGSTON, J.J.T. OWEN, et W. VAN EWIJK 1985. Stromal cell types in the developing thymus of the normal and nude mouse embryo. *Eur. J. Immunol.* 15:675-681.

VAN VLIET, E., M. MELIS et W. VAN EWIJK 1984. Monoclonal antibodies to stromal cell types of the mouse thymus. *Eur. J. Immunol.* 14:524-529.

VEILLETTE, A., J.C. ZUNIGAPFLUCKER, J.B. BOLEN et A.M. KRUISBEEK 1989. Engagement of CD4 and CD8 expressed on immature thymocytes induces activation of intracellular tyrosine phosphorylation pathways. *J. Exp. Med.* 170: 1671-1680.

WALSH, D.A., J.P. PERKINS et E.G. KREBS 1968. An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. *J. Biol. Chem.* 243: 3763-3765.

WEISS, A. 1993. T cell antigen receptor signal transduction: A tale of tails and cytoplasmic protein-tyrosine kinases. *Cell* 73: 209-212.

WEKERLE, H. et U.-P. KETELSON 1980. Thymic nurse cells. Ia-bearing epithelium involved in T-lymphocyte differentiation. *Nature* 283:402-404.

WEKERLE, H., U.-P. KETELSON et M. ERNST 1980. Thymic nurse cells. Lymphoepithelial cell complexes in murine thymuses: Morphological and serological characterization. *J. Exp. Med.* 151: 925-944.

WIERENGA, R.K., et W.G.J. HOL 1983. Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant. *Nature* 302: 842-844.

WILCHEK, M., T. MIRON et J. KOHN 1984. Affinity Chromatography. *Methods Enzymol.* 104: 3-55.

WILKS, A.F. 1990. Structure and function of the protein tyrosine kinases. *Progress in Growth Factor Research* 2: 97-111.

WILKS, A.F., A.G. HARPUR, R.R. KURBAN, S.J. RALPH, G. ZÜRCHER et A. ZIEMIECKI 1991. Two novel protein-tyrosine kinases, each with a second phosphotransferase-related catalytic domain, define a new class of protein kinase. *Mol. Cell. Biol.* 11: 2057-2065.

WILKS, A.F. 1991. Cloning members of protein-tyrosine kinase family using polymerase chain reaction. *Methods Enzymol.* 200: 533-565.

WILKS, A.F., R.R. KURBAN, C.M. HOVENS et S.J. RALPH 1989. The application of the polymerase chain reaction to cloning of members of the protein tyrosine kinase family. *Gene* 85: 67-74.

WU, L., M. ANTICA, G.R. JOHNSON, R. SCOLLAY et K. SHORTMAN 1991. Developmental Potential of the Earliest Precursor Cells from the Adult Mouse Thymus. *J. Exp. Med.* 174: 1617-1627.

YAMAMORI, T. et A. SARAI 1992. Coevolution of cytokine receptor families in the immune and nervous systems. *Neuroscience Res.* 15: 151-161.

APPENDICE I

Increased PCR sensitivity by using paraffin wax as a reaction mix overlay.

**Benoît Hébert^{1,2}, Jean Bergeron²,
Edouard F. Potworowski¹ and Peter Tijssen².**

Molecular and Cellular Probes, 7(3):249-252 (June 1993)

**Centre de recherche en immunologie¹ and Centre de recherche en virologie²,
Institut Armand-Frappier, Laval, Qc, Canada. H7N 4Z3**

**Corresponding Author: Professor Peter Tijssen
Institut Armand-Frappier
Centre de recherche en virologie
531, boulevard des Prairies
Laval, Qc H7N 4Z3**

The use of mineral oil as a reaction mix overlay in conventional PCR may lead to problems. In addition to more difficult handling, traces of mineral oil in amplicon suspensions have been shown to decrease the efficiency of post-PCR manipulations. Commercial alternatives aimed at resolving the problem more than double the cost of an amplification. This is an important drawback when dealing with a large number of clinical samples. The use of inexpensive paraffin wax as a reaction mix overlay eliminated problems associated with the presence of mineral oil, was more practical and safer in handling potentially contaminated clinical samples. Moreover, when used in conjunction with a modified hot start technique, the use of paraffin wax increased the specificity and sensitivity of PCR amplifications over mineral oil in similar hot-start conditions. Using the aforementioned method, the increase in specificity and sensitivity has enabled specific detection of viral DNA in clinical samples which the conventional PCR method failed to detect.

KEYWORDS: PCR, Hot-Start, paraffin wax overlay, PCR sensitivity

A number of protocols have been devised to broaden the applications of the polymerase chain reaction technique (PCR) (1) and to minimize technical problems inherent to this technique. One such problem is the evaporation of the reaction mix during thermal cycling. While the use of mineral oil as a reaction mix overlay prevents evaporation, it generates novel problems related to the presence of traces of mineral oil in post-PCR cloning, electrophoresis and transfers. Commercial alternatives to mineral oil are either unpractical (such as specially designed tubes that do not fit the more common cyclers) or expensive (Ampliwax™ PCR Gem 100 by Cetus Corp.). Ampliwax™, increases specificity over mineral oil, but more than doubles the cost.

We report here two modifications to the standard PCR technique which, together, increase specificity and sensitivity of the PCR. We used an inexpensive alternative to both mineral oil and Ampliwax™ which prevents evaporation of the reaction mix and enables further cloning without the need of chloroform extractions of the amplicon (amplified fragment of nucleic acid) suspension. This method relies on using paraffin wax, available for food preservation, or for tissue embedding (Paraplast Plus, Oxford Labware; containing DMSO and polyisobutylene) as a reaction mix overlay. In conjunction with paraffin wax as a reaction mix overlay, we modified the "hot start" technique (2) to increase sensitivity and specificity of the PCR which resulted in a decreased formation of non-specific amplicons or "primer-dimers".

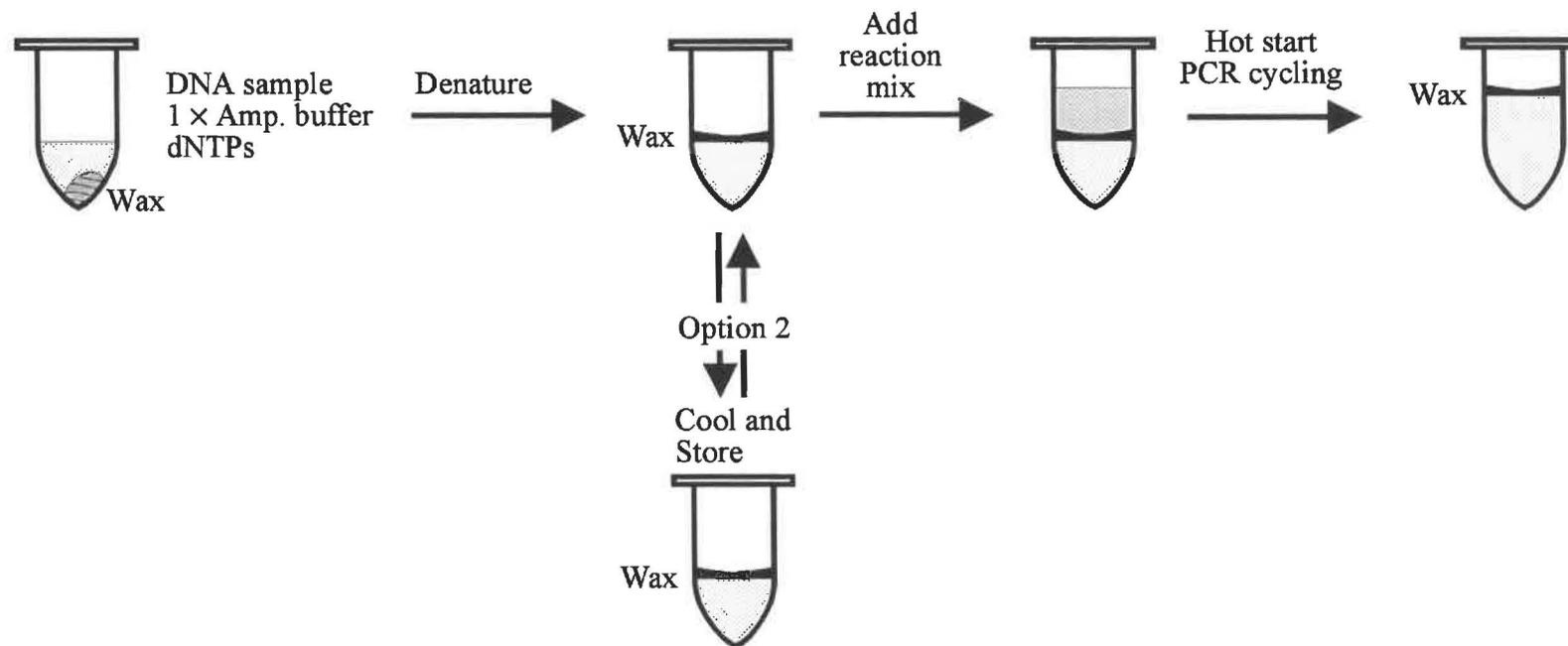
Prior to use, the paraffin wax was melted in a 90°C water bath (paraffin wax is not suited for microwave melting) and filtered at 0.45 μm (Corning) to remove possible particulate contaminants. Molten paraffin wax was distributed into 0.5 mL microfuge tubes. The volume of aliquots depends on the final reaction volume (table I). Tubes were placed in racks and autoclaved.

Final reaction volume (μL)	Paraffin wax (μL)
<25	not recommended
25	15
50	25
100	40
150	45

Table I: Volumes of paraffin wax used vary with final reaction volume.

For the modifications to the hot start technique, two strategies were chosen (figure 1). In the first option, DNA samples, 1X amplification buffer (Bio/Can Scientific; 50 mM KCl, 10 mM Tris-HCl, pH 9.0 at 25°C, 0.1% Triton X-100), 3 mM MgCl₂ (final conc.) and 0.2 mM dNTPs (each, final conc.) were added to the microfuge tubes containing paraffin wax and heated for 5 minutes at 95°C. They were then left at 75°C for 8 to 10 minutes, during which the reaction mix was added to the test tubes. The reaction mix contained 1X amplification buffer, primers (1 pmol/ μL final reaction volume) and Taq thermostable polymerase (0.01 u/ μL final reaction volume; Bio/Can Scientific). It was layered on top of the molten paraffin wax which solidified momentarily. The thermal cycling was thereafter conducted according to standard protocols (1): denaturation 1 min at 94°C, annealing 1 min at 55°C and elongation 2 min at 72°C for 35 cycles followed by a 10 min extension at 72°C. A ramp time of 2 min was imposed between the denaturation and annealing temperatures.

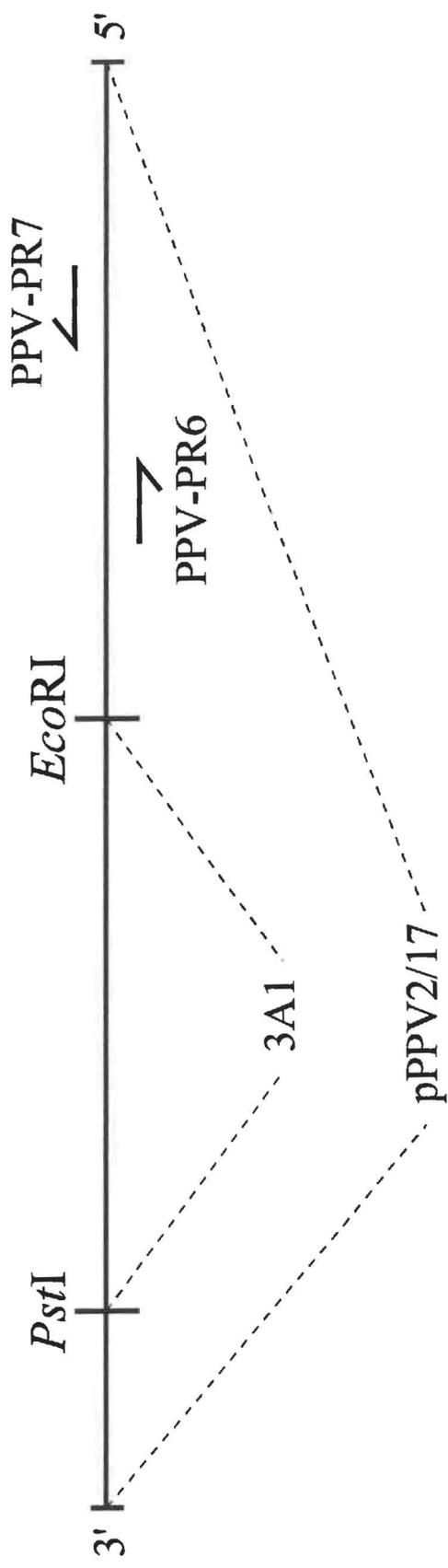
In the second option, DNA samples, dNTPs, MgCl₂ and 1X amplification buffer were added to the tubes together with the paraffin wax and heated 5 min at 95°C (e.g. in a pre-PCR laminar flow hood). After cooling of the paraffin wax, the reaction mix was added to the tubes. Control tubes containing mineral oil were

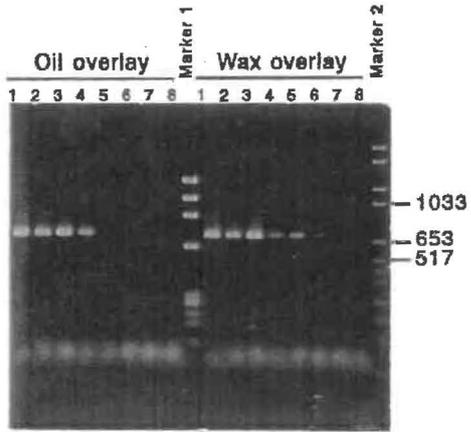


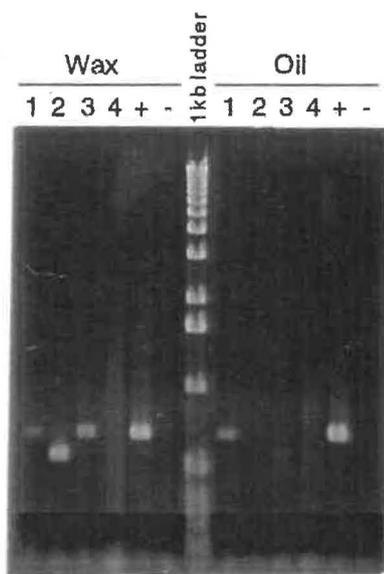
prepared according to the second option, the wax being replaced by an oil overlay. A hot-start was equally performed on the control tubes.

In order to evaluate the aforementioned technique, one pair of primers (PPV-PR6, GTTAATAGTAAACACATGAGAGCTTG and PPV-PR7, CAATACTGCACCTGTATTTCCAAATG) was designed to amplify a segment of the porcine parvovirus (PPV) strain NADL2 (clone p17) and expected to yield a fragment of 687 bp. Another NADL2 clone (3A1), used as negative control, consisted of a *Pst*I - *Eco*RI fragment lacking the sequence complementary to the primers (figure 2). The amplification was specific and the detection, using paraffin wax as a reaction mix overlay, was ~200 better than with mineral oil (figure 3). Thus, we were able to amplify from 0.1 pg of starting material with the paraffin overlay vs 20 pg for the oil overlay. Similar results were obtained in four repeat experiments. Upon the start of cycling, the paraffin wax melted only above 55°C permitting mixing of the reagents and thus prevented mispriming below this temperature. Clinical samples were also tested using this method, and in several cases, the paraffin method detected viruses which failed to amplify in reactions with a mineral oil overlay (figure 4). These DNA samples have been isolated from fetal porcine tissues using the method described by Hirt (3) except for sample 2 which is a crude aliquot of culture supernatant of porcine parvovirus infected PFT cells. This sample had no detectable amplification when an oil overlay was used. The smaller fragment amplified is characteristic of the Kresse pathogenic strain in which a 127bp repeat is absent (4,5).

The above modifications to the overlay offer three advantages: First, the ease of removal of the amplified material without contaminating oil. Successful cloning of the amplified material without chloroform extractions has proven to save valuable time and to minimize loss of amplicon. Second, a clean environment is required during pre-PCR stages to aliquot clinical samples in tubes containing paraffin wax. Sealing of the tubes is done by heating the samples in a dry block







followed by cooling at room temperature; there is increased safety in handling infectious samples. Workers are thus isolated from infectious samples. Third, the PCR reaction mix can be added over the solid paraffin wax at the benchtop which minimizes handling time when a batch of tubes have already been prepared. During thermal cycling, the heavier aqueous phase will pass through the melting wax and mix with the other reagents only at a temperature which is usually above the annealing temperature of the primers. The modified hot start reduced primer oligomerization and the generation of non-specific amplicons by optimizing cycling efficiency in the first cycles of PCR (6). These advantages, added to the low cost of the paraffin wax overlay, amply justify the time required for preparing a batch of tubes.

The two alternate modifications to the reaction mix overlay combined with the hot start technique give similar results. They are currently successfully used in four instances: cloning of a novel protein tyrosyl kinase from murine thymic epithelial cells (7; B.Hébert, manuscript in preparation), PCR-Mini prep screening (J.Bergeron, unpublished), RT-PCR in order to detect mRNA coding for murine cytokines (D. Rochefort, personal communication) and clinical detection of porcine parvovirus (8). In all these applications, results were clearer and in some instances, improved yields have been obtained.

Paraffin waxes, of which three brands have been used, show no noticeable difference among them when used as reaction mix overlay. They are available from local grocery stores in the case of food preservation paraffin wax or from histological supply firms in the case of tissue embedding paraffin wax. They all perform at least as well as an expensive specially designed alternative.

References

- 1- Tijssen, P. (1993) *Hybridization with Nucleic Acid Probes* (2 vol.), Elsevier, Amsterdam. (In Press).
- 2- D'Aquila, R.T., Bechtel, L.J., Videler, J.A., Eron, J.J., Gorczyca, P. and Kaplan, J.C. (1991). Maximizing sensitivity and specificity of PCR by pre-amplification heating. *Nucl. Acids Res.* **19**, 3749.
- 3- Hirt, B. (1967). Selective extraction of polyoma virus DNA from infected cell cultures. *J. Mol. Biol.* **36**, 365-369.
- 4- Bergeron, J., St-Laurent, G., Menezes, J. and Tijssen, P. (1993). Porcine Parvovirus. I. Genomic organization and comparison of isolates. (Submitted)
- 5- Bergeron, J., Menezes, J. and Tijssen, P. (1993). Porcine Parvovirus. II. Mapping of viral transcripts and translation products. (Submitted)
- 6- Mullis, K.B. (1991). The polymerase chain reaction in an anemic mode: How to avoid cold oligodeoxyribonucleotide fusion. *PCR Meth. Appl.* **1**, 1-4.
- 7- Couture, C., Amarante-Mendes, G. and Potworowski, E.F., (1992). Tyrosine kinase activation in thymic epithelial cells: necessity of thymocyte contact through the gp23/45/90 adhesion complex. *European J. Immunol.* **22**, 2579-2585.
- 8- Bergeron, J., Jetté, R., Montpetit, C., Marsolais, G., Descôteaux, J.-P. and Tijssen, P. (1992). Specific Detection of Porcine Parvovirus Strains. *Proceedings of the International Pig Veterinary Society*, vol. 1, p99. The Hague, The Netherlands.

Acknowledgements

B.H. is supported by an Institut Armand-Frappier graduate fellowship. This work is funded in part by grants of the Medical Research Council of Canada to E.F.P. and the National Research Council of Canada to P.T.

APPENDICE II

Appendice II: **LE RÉSEAU INTERNET**

Les néophytes sont fortement encouragés de lire la publication de Brendan P. Kehoe intitulée **ZEN AND THE ART OF THE INTERNET** (Prentice-Hall, août 1992). Ce livre est également disponible gratuitement sur le réseau Internet, via FTP anonyme à l'adresse quartz.rutgers.edu dans le répertoire internet/docs/zen.

1. La banque de données PKCDD

La banque de données des domaines catalytiques de protéine kinases maintenue par Steven Hanks et Anne Marie Quinn est disponible par FTP anonyme à l'adresse salk-sc2.sdsc.edu dans le répertoire principal. Les fichiers à récupérer sont:

pkinases.pep	la banque de données sous format IG (IntelliGenetics)
pkinases.fa	la banque de données sous format FastA
pkinases.index	la liste des protéine kinases répertoriées dans la banque de données. Cette liste inclus les références.

2. Le logiciel MACAW

Le logiciel MACAW est disponible au *National Center for Biotechnology Information* (NCBI) du *National Institute of Health* (NIH; Bethesda). Le logiciel MACAW est disponible par FTP anonyme au NCBI à l'adresse ncbi.nlm.nih.gov dans le répertoire pub/macaw.

3. Le logiciel ReadSeq

Le logiciel ReadSeq développé par Don Gilbert est disponible par FTP anonyme à l'adresse ftp.bio.indiana.edu dans le répertoire molbio/readseq.

