

# Validation de la méthode d'approche d'électrophorogramme par une somme de gaussiennes

## Validation of the approach method of electropherogram by a sum of gaussians

par Karim AFDEL\*, E. JOVER\*\*, J. PALMARI\*\*\*, Mostafa BANOUNI\*

\* Laboratoire Instruments et Mesures, Université Ibnou Zohr d'Agadir, Faculté des Sciences, B.P 28/S, Fax (08) 22-01-00, Agadir, Maroc.

\*\* Laboratoire de Neurobiologie des Canaux Ioniques, Université d'Aix Marseille II, Faculté de Médecine de Marseille, Bd Pierre Dramard 13916 Marseille Cedex 20

\*\*\* Laboratoire de Cancérologie Expérimentale, Université d'Aix Marseille II, Faculté de Médecine de Marseille, Bd Pierre Dramard 13916 Marseille Cedex 20

### *résumé et mots clés*

L'estimation quantitative des autoradiogrammes par une somme de gaussiennes est une approche très utilisée. Elle est rendue possible grâce au développement des méthodes densimétriques utilisables après digitalisation de l'image d'autoradiographie du gel. La métrologie apportée par l'analyse et le traitement d'images permet d'éliminer les appréciations subjectives visuelles et humaines. Plusieurs études ont été réalisées sur ce sujet. Notre contribution consiste à valider cet approche par son application à des fragments d'ARN de tailles très proches. Nous montrons qu'il est possible d'évaluer directement par le calcul le niveau d'expression d'ARN dans une configuration d'expérience déterminée, évaluation très difficile à obtenir par d'autres méthodes.

Traitement d'images, autoradiographie, électrophorèse, lissage, estimation.

### *abstract and key words*

Quantitative estimation of one dimensionnal autoradiography film density profiles that are generated in gel electrophoresis of radioactive proteins and nucleic acids is a very used approach. Several studies have been realized on this subject. They operate on the optical density values acquired by CCD cameras. These operations include smoothing of measuring values, background correction, and integration of peak areas after sided fit of half gaussian profile to each band. Our contribution consists to validate this approach by its application to separate RNA fragments of very close size. We show that is possible to use this approach to evaluate directly the relative expression level of each RNA in a determined experiance configuration.

Image processing, autoradiography, electrophoresis, smooting, estimation.

## 1. introduction

Parmi toutes les techniques d'analyse utilisées dans le domaine clinique et en biochimie, l'électrophorèse est celle qui subit actuellement une évolution croissante [1,2]. Elle est rapide et

permet de déterminer avec précision la longueur et la pureté des molécules d'ADN ou d'ARN et par conséquent d'établir des cartes génétiques. La séparation et l'analyse sont basées sur la migration différentielle de particules chargées sous l'influence d'un champ électrique. Au cours d'une électrophorèse, les particules porteuses de charges électriques que l'on veut étudier sont placées en

milieu aqueux en présence d'électrolytes dilués et migrent à l'intérieur d'une phase solide sous forme de gel. L'automatisation de la lecture des gels d'électrophorèse est maintenant rendue possible grâce aux performances des micro-ordinateurs et aux techniques de traitement et d'analyse d'images. Un grand nombre de constructeurs offrent sur le marché des logiciels d'analyse des profils densimétriques dits électrophorogramme. Cependant, ces logiciels sont dans la plupart des cas rigides et peu transparents. Les concepteurs de tels logiciels ne mentionnent pas la technique utilisée pour le calcul des surfaces des pics, ni la méthode de correction de la ligne de base. De plus, ils ne fournissent pas de test de validité. Toutes ces questions soulevées jettent des doutes quant aux résultats fournis par ces logiciels; ce qui exclut leur utilisation dans le domaine de la recherche scientifique. Dans le but d'éclaircir ces questions techniques, nous avons conçu un programme d'estimation des pics par des gaussiennes et nous avons développé un modèle de test adapté pour valider cette méthode d'approche.

## 2. méthode d'analyse des images autoradiographiques

Une fois la migration électrophorétique achevée, le gel est analysé au moyen de plusieurs procédés [3]. L'autoradiographie est l'un des procédés les plus utilisés. Elle consiste à incorporer un radio-isotope au sein de la molécule étudiée [4,5]. Le gel est ensuite placé en contact d'un film radiologique pour obtenir l'autoradiographie qui constitue l'image analogique à analyser. Un exemple d'images analogiques est représenté par la figure 1. De telles images sont généralement composées de plusieurs pistes contenant chacune une série de bandes plus ou moins denses. La densité des bandes contient l'information nécessaire à l'identification des gènes.

La figure 2 présente un profil densimétrique pris le long d'une piste; chaque pic de ce profil correspond à une bande. L'aire sous

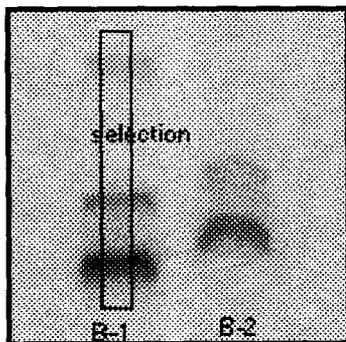


Figure 1. – Partie d'une image d'autoradiographie de gel d'électrophorèse montrant deux pistes B-1 et B-2.

Niveau de gris

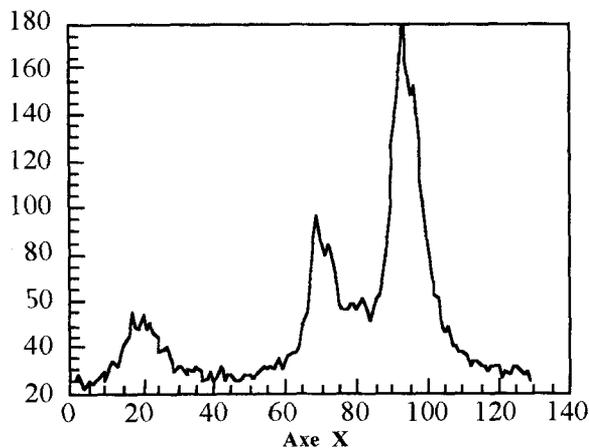


Figure 2. – Profil densimétrique le long d'une ligne de la piste B-1 de la figure 1.

chaque pic est proportionnelle à la densité de la bande correspondante. L'évaluation d'une telle aire n'est pas évidente dans le cas particulier où les pics ne sont pas clairement séparés. De plus, le profil présente souvent une composante continue qui influence négativement sur cette évaluation. La détermination « manuelle » des aires sous les pics dépend ainsi des considérations entreprises par l'opérateur pour séparer correctement les pics et estimer convenablement la composante continue. Il est cependant possible de surmonter ces deux difficultés en procédant à un traitement automatique des images d'autoradiographie. Cette procédure de traitement automatique est composée de trois étapes [1,2,6] :

- le filtrage de l'image pour éliminer les artefacts qui l'entachent,
- la localisation des bandes par la détermination des extremums,
- le calcul d'intensité des bandes.

### 2.1. réduction du bruit et correction de la ligne de base

L'évaluation automatique de l'intensité des bandes est basée essentiellement sur la détermination des maximums et des minimums du profil densimétrique. Cette opération est très sensible aux bruits engendrés par les détecteurs optoélectroniques lors de la numérisation de l'image ou par des impuretés liées à la technique d'électrophorèse. Pour éliminer ce bruit, nous avons calculé la moyenne de plusieurs profils situés sur la partie médiane d'une même piste.

La diffusion des impuretés liées à l'absorption inhomogène du gel entraîne également l'apparition d'une composante continue. Pour éliminer cette composante, nous avons procédé à la méthode de la construction progressive de la ligne de base à partir de segments et en se basant sur le calcul de la pente [2]. La figure 3 présente le résultat de ces deux opérations.

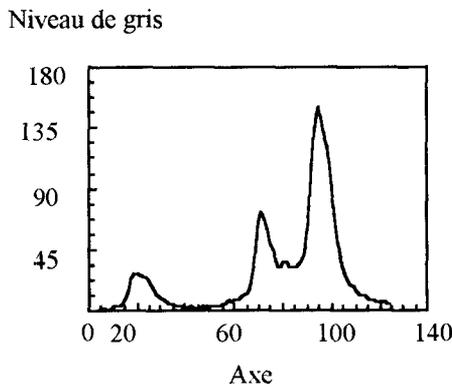


Figure 3. – Moyenne des profils densimétriques délimités par la sélection de la figure 1 et élimination de la composante continue.

### 2.2. estimation du profil densimétrique par une somme de gaussiennes

Les pics des profils densimétriques des gels d'électrophorèse présentent une forme régulière et peuvent être estimés – à priori – par des gaussiennes non normalisées en utilisant la méthode des moindres carrés [2,3,7]. On obtient la valeur du profil estimé, PFL<sub>Estimé</sub>, en un point  $j$  par la somme de ces gaussiennes :

$$PFL_{Estimé}(j) = \sum_{i=1}^p A_i \exp \left[ -\frac{(j - m_i)^2}{2\sigma_i^2} \right] \quad (1)$$

où le paramètre  $P$  représente le nombre des pics du profil densimétrique et les paramètres  $A_i$ ,  $\sigma_i$  et  $m_i$  sont respectivement l'amplitude, l'écart type et la position du maximum du pic  $i$ . Les paramètres  $A_i$  et  $\sigma_i$  permettent d'évaluer l'aire sous le pic  $i$  par l'expression suivante :

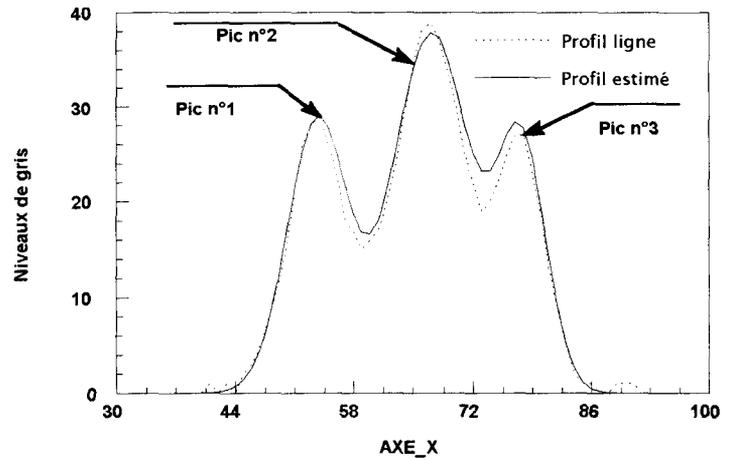
$$S_i = A_i \sigma_i \sqrt{2\pi} \quad (2)$$

### 2.3. recherche des paramètres des gaussiennes

Sur un profil, les pics peuvent être facilement mis en évidence comme ceux présentés par la figure 4-a ci-après. Dans certains cas, un pic peut être « noyé » dans un autre plus intense et constitue ainsi un épaulement; c'est le cas du profil montré sur la figure 4-b.

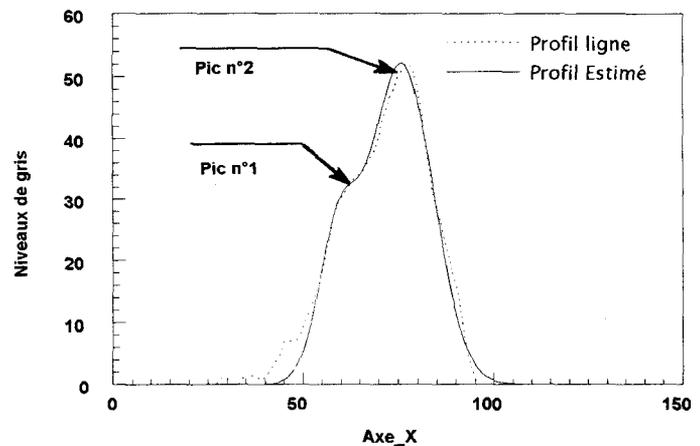
La recherche de pics consiste ainsi, dans une première étape, à localiser les maximums et les épaulements en procédant à un calcul de dérivées première et seconde du profil. Dans une deuxième étape, on évalue les paramètres  $A_1$ ,  $\sigma_1$  et  $m_1$  du premier pic en utilisant la méthode des moindres carrés. Cette évaluation consiste à minimiser la quantité  $\chi_1^2$  donnée par :

$$\chi_1^2 = \sum_{j=m_1-\sigma_1}^{m_1+\sigma_1} \left\{ Y(j) - A_1 \exp \left( -\frac{(j - m_1)^2}{2\sigma_1^2} \right) \right\}^2 \quad (3)$$



	A	m	$\sigma$
Pic n°1	27	51,97	3,69
Pic n°2	36	65,42	4,23
Pic n°3	24	75,94	2,79

Figure 4a. – Profils expérimental et estimé par la méthode d'approche par des gaussiennes.



	A	m	$\sigma$
Pic n°1	24	57	5,78
Pic n°2	50	73	8

Figure 4b. – Profils expérimental et estimé de deux bandes très proches.

traduisant l'écart entre les valeurs expérimentales du profil densimétrique,  $Y(j)$ , et la gaussienne représentant le pic sur un intervalle de largeur  $2\sigma_1$  centré autour de  $m_1$ . Dans le cas d'un épaulement, la minimisation est effectuée dans l'intervalle  $[m_1 - \sigma_1, m_1]$ . Dans une étape suivante, on soustrait la contribution du premier pic représenté par la gaussienne de paramètres  $A_1$ ,  $\sigma_1$  et  $m_1$  ainsi déterminés du profil expérimental. On évalue ensuite les paramètres  $A_2$ ,  $\sigma_2$  et  $m_2$  du second pic dont on soustrait la contribution du profil résultant. On procède ainsi pour le reste des pics.

La figure 4 représente des exemples de profil expérimentaux et estimés ainsi que les valeurs des paramètres des gaussiennes. On peut constater l'accord satisfaisant entre les deux types de profils.

### 3. application à la séparation de fragments d'ADN de tailles très proches

Cette approche a été appliquée à l'analyse des ARN messagers correspondant à trois sous-unités  $\alpha$  de canal sodium potentiel dépendant, nommées RBCI, RBCII et RBCIII [8,9]. Les échantillons d'ADN analysés dans ce travail proviennent d'amplifications RT-PCR réalisées sur des ARN extraits à partir de cinq sources indépendantes : cerveau d'embryon de rat de 16 jours (source A), cerveau de rat adulte (source B), cervelet de rat adulte (source C), neurones embryonnaires 7 jours après la mise en culture (source D) et neurones embryonnaires 10 jours après la mise en culture (source E).

Le signal fluorescent des bandes d'ADN est digitalisé à l'aide d'une camera CCD (système « The Imager », Appligène, Illkirch). L'image est conservée en format TIFF pour analyse ultérieure.

#### 3.1. digestions enzymatiques

La séquence nucléotidique des fragments d'ADN étant connue, il est possible de déterminer quelles enzymes de restriction peuvent hydrolyser spécifiquement un seul des trois fragments (bandes) RBCI, RBCII et RBCIII. Pour la digestion enzymatique du produit de RBCI (614 bp) nous avons utilisé Hind III et Pst I. Pour le produit de RBCIII (625 bp) les enzymes sont Dra I et Hinc II.

Après amplification, plusieurs aliquotes identiques sont séparées. Un échantillon est conservé pour la migration témoin et les autres sont soumis à digestion enzymatique pendant une heure. Avant électrophorèse, tous les échantillons sont équilibrés en volume et, si nécessaire, les concentrations en sels sont aussi rendues équivalentes.

#### 3.2. validation de la méthode d'approche

Notre objectif consiste à établir la contribution relative exprimée en % de chacun des trois fragments RBCI, RBCII et RBCIII au signal total et de comparer ces pourcentages pour les cinq différentes sources d'ARN. Les figures 5 et 6 montrent les images acquises de deux gels, désignés par gel 1 et gel 2, contenant les

5 sources A, B, C, D et E. Pour chaque source et dans les deux gels, la piste témoin 1 contient les trois fragments d'ADN. Dans le gel 1, les enzymes Hind III et Pst I ont été utilisés pour digérer le fragment RBCI et dans le gel 2 ce sont les enzymes Dra I et Hinc II qui ont servi à digérer le fragment RBCIII. Nous donnons également sur les figures 5 et 6 les profils ayant permis l'analyse. Nous pouvons constater sur ces profils que la séparation entre les fragments RBCI et RBCIII est peu marquée.

Les résultats de l'analyse sont présentés dans le tableau I. Nous avons utilisé l'approche d'estimation des pics du profil densimétrique par des gaussiennes pour calculer, pour les deux gels, la contribution relative de chaque pic au signal total dans la piste. Dans la colonne 3, nous avons reporté les contributions des trois fragments RBCI, RBCII et RBCIII de la piste témoin 1 dans les cinq sources. La comparaison de ces contributions obtenues pour les deux gels montre une bonne reproductibilité des résultats. Notons à ce stade que les résultats relatifs à la source B sont similaires à ceux de « Northern Blot » [10,11].

Les résultats reportés dans les colonnes 5 et 6 du tableau 1-a représentent les contributions relatives des fragments RBCI et

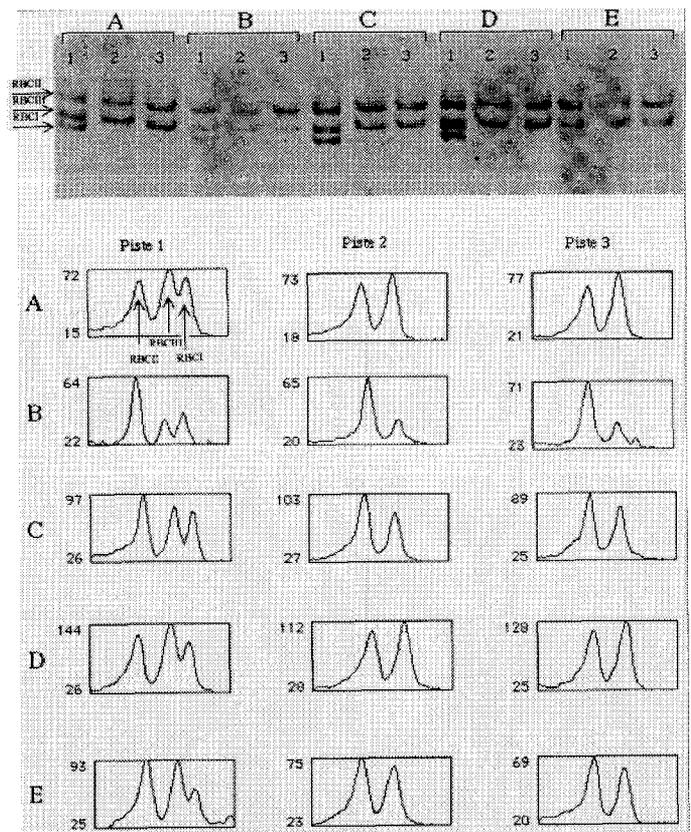


Figure 5. – Image d'autoradiographie du gel 1 contenant 5 source A, B, C, D et E. Pour chaque source la piste contient trois fragments d'ADN. Pour les pistes 2 et 3 on a utilisée respectivement les enzymes Hind III et Pst I pour digérer le fragment RBCI. On a représenté les profils densimétriques correspondant à chaque piste et à chaque source.

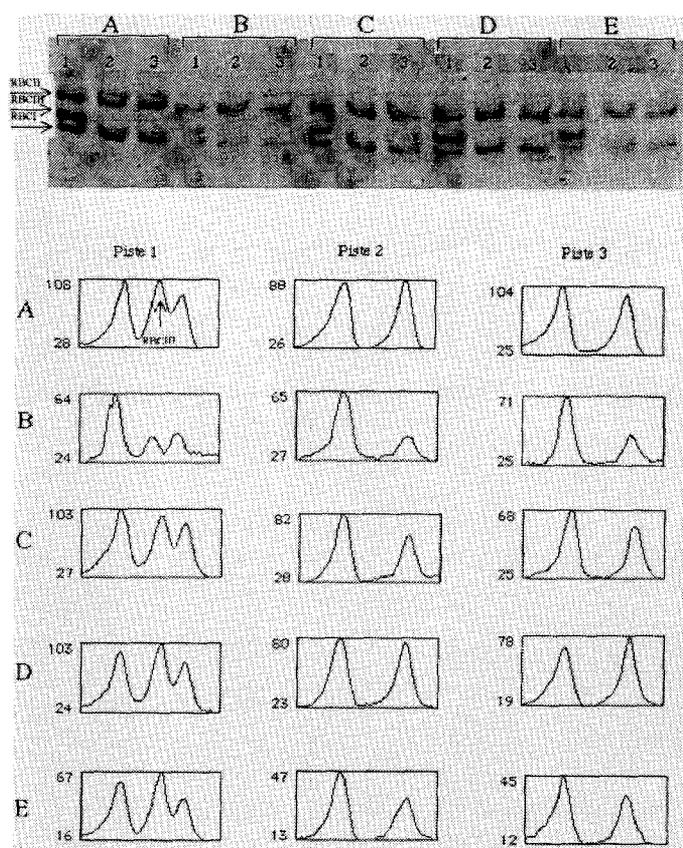


Figure 6. – Image d'autoradiographie du gel 2 contenant 5 source A, B, C, D et E. Pour chaque source la piste contient trois fragments d'ADN. Pour les pistes 2 et 3 on a utilisé respectivement les enzymes Dra I et Hinc II pour digérer le fragment RBCIII. On a représenté les profils densimétriques correspondant à chaque piste et à chaque source.

RBCII dans les pistes 2 et 3. Ces résultats ont été obtenus en utilisant, comme pour la piste 1, l'approche d'estimation par des gaussiennes. Là aussi, on peut noter un accord satisfaisant entre les deux pistes considérées et ceci d'une source à l'autre.

Par ailleurs, en procédant à une simple soustraction mathématique du pic correspondant au fragment RBCI du profil estimé de la piste témoin 1, nous avons obtenu les contributions relatives des deux fragments restants. Les résultats sont reportés dans la colonne 4 du tableau 1-a. En comparant ces valeurs à leurs valeurs homologues dans les colonnes 5 et 6, on peut constater une grande similitude entre elles. La même similitude est constatée pour le gel 2.

## 4. conclusion

A l'aide des résultats qui sont présentés ici, nous montrons que le programme d'estimation des gaussiennes permet d'évaluer directement le niveau relatif d'expression de chaque ARN

Tableau 1. – Contribution des différents fragments RBCI, RBCII et RBCIII aux profils densimétriques. Les colonnes 3, 5 et 6 sont relatives aux pistes 1, 2 et 3 des figures 5 et 6. La colonne 4 contient les contributions des fragments restant après digestion enzymatique; ces contributions sont obtenues, à partir du profil estimé relatif à la piste témoin 1, par simple soustraction des pics correspondant aux fragments digérés.

GEL 1					
Tableau 1a			Soustraction de la bande 3(%)		
source	bande	% relatif	Calcul	Hind III	Pst I
A	RBCII (1)	29	47	47	45
	RBCIII (2)	32	53	53	55
	RBCI (3)	38	–	–	–
B	1	52	72	70	71
	2	20	28	30	39
	3	28	–	–	–
C	1	38	54	55	51
	2	32	46	45	49
	3	30	–	–	–
D	1	35	50	46	51
	2	34,5	50	54	49
	3	30,5	–	–	–
E	1	39	52	54	58
	2	37	48	46	42
	3	24	–	–	–

GEL 2					
Tableau 1b			Soustraction de la bande 2(%)		
source	bande	% relatif	Calcul	Dra I	Hinc I
A	RBCII (1)	26	40	48	51
	RBCIII (2)	36	–	–	–
	RBCI (3)	38	60	52	49
B	1	51	65	68	64
	2	21	–	–	–
	3	28	35	32	36
C	1	36	52	56	55
	2	31	–	–	–
	3	33	48	44	46
D	1	34	54	56	52
	2	36	–	–	–
	3	30	46	44	48
E	1	37	59	58	56
	2	37	–	–	–
	3	26	41	42	44

dans une configuration d'expérience déterminée. Les profils densimétriques obtenus à partir des gels d'électrophorèse montrent que la séparation entre les fragments de RBCI et RBCIII n'est pas suffisante pour pouvoir définir l'apport de chaque bande au signal

total. Grâce au calcul des gaussiennes nous avons pu soustraire de chaque pic l'apport de densité optique provenant de l'autre pic. Comme méthode de confirmation, nous avons comparé les résultats ainsi obtenus avec des mesures densimétriques effectuées sur des échantillons où l'un des fragments d'ADN a été éliminé par une digestion enzymatique spécifique. Le bon accord entre les différents résultats permet de valider la méthode d'approche des profils densimétrique par des gaussiennes. Ainsi, l'association du traitement d'image et de l'analyse des données permet d'accroître le pouvoir de résolution des gels d'électrophorèse.

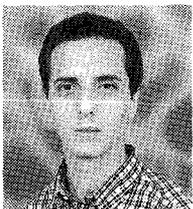
### BIBLIOGRAPHIE

- [1] T. Redman and T. Jacobs, *Electrophoretic Gel Image Analysis Software for de Molecular biology Laboratory*, Biocomputing Technical Report, No. 6, Vol. 10, 1991, page 790-794.
- [2] M. Habin, M. Yakin and H. Kronberg, *Photometric Evaluation of Slab Gels II : Delimitation and Integration of peaks in one dimensionnal Electrophoregrams*. Electrophoresis No. 3, 1989, page 244-254.
- [3] B. Masters, T. Griggs and B. Berde, *High sensitivity quantification of RNA from Gels and Autoradiograms with affordable Optimal Scanning*, Biotechnique, Vol. 12, No. 6, 1992.
- [4] A.T. Andrews, *Electrophoresis*, Oxford University Press, 1981.
- [5] Y. Kitao, M. Miyahara and D. Hiraoka, *Quantitative Determination of Overlapped Proteins in Sodium Sulfate-polyacrylamide Gels Electrophoresis*, Analytical Biochemistry 134, 1983, page 295-302
- [6] K. R. Castelmann, *Digital Image Processing*, Edition Prentice Hall, 1979, page 97-105.
- [7] N. R. Deering and D. Shalloway, *GelMetric : Semi Automated Electrophoretic Mobility Analysis* Cabios, 7, 1, 1991, page 109-110.
- [8] T. Kayano, M. Noda, V. Flockerzi, H. Takahashi and S. Numa, *Primary structure of rat brain sodium channel III deduced from the cDNA sequence* FEBS Lett 225, 1989, page 187-197.
- [9] A. Lara, B. Dargent, F. Julien, F. Courant and E. Jover, *Channel Activators reduce the expression of sodium channel  $\alpha$ -subunit mRNA in developing neurons* Nature 420, 1994, page 140-148.
- [10] M. Noda, T. Ikeda, T. Rayano, H. Takahashi and S. Numa, *Existence of distinct sodium channel messenger RNA in rat brain*, Nature 320, 1986, page 188-192.
- [11] S. Beckh, M. Noda and S. Numa, *Differential Regulation three Sodium Channel messenger RNA species in the rat central nervous system during development*, EMBOJ 8, 1989, page 3611-3616.

Manuscrit reçu le 2 avril 1998.

### LES AUTEURS

Karim AFDEL



Diplômé de doctorat en génie informatique spécialité Analyse et Traitement numérique d'images de l'Université d'Aix Marseille II, professeur assistant à l'Université Ibnou Zohr faculté des Sciences Agadir. Actuellement je m'intéresse aux analyses des images biomédicales et aux techniques de classifications.

E. JOVER

Mostafa BANOUNI



Diplômé de doctorat en physique de l'Université de Montpellier II, des travaux réalisés sur l'application des techniques d'analyse d'images aux images obtenue par photo-émission. Actuellement je m'intéresse aux analyse des images biomédicales et aux techniques de classifications.

J. PALMARI

LOUIS - JEAN  
avenue d'Embrun, 05003 GAP cedex  
Tél. : 04.92.53.17.00  
Dépôt légal : 517 — Juin 1999  
Imprimé en France