

# **Aplicación del Modelo de Reconocimiento Resonante al estudio de las proteínas quinasas.**

M. Sautié, R. Hong, J. Falcón, J.L. Hernández Cáceres

## **RESUMEN**

Se realiza un estudio exhaustivo, utilizando un elevado número de proteínas, del modelo de reconocimiento resonante propuesto por Cosic y col [1] para el análisis de secuencias aminoacídicas de una proteína, a partir de la serie generada mediante la sustitución de aminoácidos por sus respectivos potenciales de interacción ión-electrón. Se estudian los espectros de potencias y los espectros cruzados de varios grupos de proteínas emparentadas, con el objetivo de poder apreciar la conservación de la frecuencia característica planteada en el modelo. Conclusiones: i) El valor de la frecuencia característica de la superfamilia de las proteínas quinasas encontrado por Cosic no coincide con el valor mucho más representativo calculado en el presente trabajo ii) En la aplicación del Modelo de Reconocimiento Resonante al grupo de las proteínas quinasas no se cumple al menos uno de sus aspectos básicos ; a saber, el criterio de discriminación de picos significativos o la correspondencia biunívoca entre función biológica y una única frecuencia característica.

**Palabras clave:** *modelo de reconocimiento resonante, espectro de potencias, espectro cruzado.*

## **A NOTE ABOUT THE RESONANT RECOGNITION MODEL.**

### **ABSTRACT**

A critical revision of the resonant recognition model stated by Cosic et al [1] is made using a high number of related proteins. The model attempts to analyze the aminoacid sequences in proteins by substituting each aminoacid by the corresponding ion-electron pseudopotential. In order to appreciate the conservation of the characteristic frequency for any biological function stated in the model, we study the power spectrum and the cross-spectrum of several groups of related proteins. Conclusion: i) The characteristic frequency reported by Cosic et al [1] for the family of the protein kinases is incorrect, a representative value of this frequency was found for this family, ii)At least one of the fundamental postulates of the resonant recognition model is incorrect: the discrimination criteria for significant peaks or the correspondence between a biological function an a single characteristic frequency.

### **INTRODUCCIÓN**

En los últimos años las técnicas de la biología molecular han puesto a nuestra disposición un gran caudal de información genética en forma de largas secuencias, es por eso cada vez más importante el problema de extraer información útil a partir de estos datos. Una aproximación interesante es la propuesta hecha por Cosic y col en lo que ellos han dado en llamar Modelo de Reconocimiento Resonante (MRR)[1].

Tradicionalmente se ha considerado que con el análisis estructural de toda la proteína, y en especial con la determinación de la configuración, naturaleza y localización del sitio activo, se tenía información suficiente para determinar su función biológica. Por ello, la determinación de la estructura tridimensional activa y de la función biológica han compartido, en principio, los mismos instrumentos de análisis y predicción con todas sus limitaciones, desestimándose parámetros físico-químicos cruciales que podrían dar cuenta de un fenómeno tan importante para la función biológica como es la selectividad [1-4].

No fue hasta la década del 80 que Cosic y col utilizaron técnicas clásicas de procesamiento de señales digitalizadas como la transformada discreta de Fourier (TDF) y diferentes escalas de parámetros físico-químicos para la conversión de la secuencia primaria de la proteína en una secuencia numérica, encontrando que con solo uno de los múltiples parámetros analizados, el pseudopotencial de interacción ión-electrón, se obtiene una correlación óptima entre la función biológica y la frecuencia de amplitud máxima del espectro cruzado de varias proteínas que comparten la misma actividad biológica [1].

A los sorprendentes descubrimientos de que a cada frecuencia característica le corresponde una y solo una función biológica y de que el sustrato comparte esta misma frecuencia pero de fase opuesta (sea este de naturaleza proteica o de ADN), Cosic y col le han encontrado un amplio campo de aplicación como herramientas de predicción de puntos calientes (hot spots) en varias moléculas importantes, destacándose la interleukina2 [5] y el enhancer SV40 [6]; estos autores lograron predecir péptidos que mimetizan la inmunogenicidad del HIV [7] y la región de reconocimiento al receptor del FGF [8]. Estos hallazgos les han servido de base para postular su modelo, cuyo planteamiento esencial es que la función biológica de una proteína es esencialmente 'la transferencia de energía resonante de una macromolécula a otra' [9].

Basados en el MRR Cosic y col [1] reportaron las frecuencias características de numerosos grupos funcionales, aunque en muchos casos contaban con muy pocas secuencias aminoacídicas por grupo. Para las proteínas quinasas en particular, ellos encontraron una frecuencia de 0.42969 con solo 8 proteínas quinasas. [1] Dada la poca representatividad de estos resultados y la considerable envergadura de la superfamilia de las quinasas (que abarcan más de 50 subfamilias ampliamente distribuidas por todos los reinos)[3] nos propusimos, en principio, repetir los experimentos de Cosic utilizando una muestra más grande y heterogénea de proteínas quinasas con el fin de estudiar la validez del MRR para este grupo funcional con una muestra verdaderamente representativa y obtener además, una frecuencia característica más confiable.

## **METODOLOGÍA**

El procedimiento utilizado en el presente trabajo se basó esencialmente en el MRR; primeramente la estructura primaria se convierte en una serie numérica asignándole a cada aminoácido su valor correspondiente de pseudopotencial de interacción ión-electrón, luego a esta serie se le determina el espectro utilizando la transformada discreta de Fourier; debido a que la distancia promedio entre los aminoácidos es de alrededor de 3.8 Å, se asume que los puntos sucesivos de la señal están separados a una misma distancia a la que para mayor comodidad se le asigna un valor unitario [1,7]. Para obtener los componentes de frecuencia común se calcula el espectro cruzado de varias señales. Se utilizaron alrededor de 360 proteínas quinasas pertenecientes a 38 grupos diferentes cuyas secuencias están publicadas en (<http://www.sdsc.edu/kinases/>). Estos grupos están definidos según el

esquema de clasificación de Hanks y Quinn [10], basado en la similitud de la estructura primaria de los dominios catalíticos. Primeramente se calculó el espectro cruzado por grupo (cada grupo tiene de 2 a 30 proteínas), luego, al no contar con la capacidad de cálculo y de memoria operativa necesaria para obtener el espectro cruzado total que incluyera los 38 grupos, se dividieron arbitrariamente en tres grandes conjuntos y a partir de cada uno se obtuvo un espectro consenso, el primero incluyó los 13 grupos siguientes: CaMK I, AGC I, AGC II, AGC III, PTK XIII, AGC IV, AGC V, PTK XV, AGC VI, AGC VII, PTK XVI, AGC VIII, AGC other ; el segundo, los 7 grupos siguientes: PTK XIV, PTK XIX, CaMK II, CMGC I, CMGC II, CMGC III y CMGC IV y el tercero estos 17 grupos: PTK XXI, OPK VI, CaMK other, OPK II, OPK IX, OPK other, CMGC V, OPK VIII, OPK X, CMGC other, PTK I, PTK XVIII, PTK II, PTK XI, PTK other, PTK IV y OPK XII. Por último procedimos a calcular el cociente señal/ruido (S/R) para los picos de mayor amplitud en el espectro, el cociente se calcula dividiendo la intensidad del pico por la media del espectro; este índice se usa para determinar los picos significativos ( $S/R > 20$ ) o como medida de similitud para comparar las magnitudes de los picos de diferentes espectros [1]. Luego a partir de los 17 grupos del espectro cruzado III se seleccionaron aquellos cuyos espectros multiplicados entre sí contribuyen notablemente a la magnitud del pico de frecuencia 0.0923, con el fin de demostrar si es posible reproducir, con al menos algunos de estos grupos, el pico significativo a esta frecuencia obtenido en los espectros cruzados I o II.

## RESULTADOS

Como puede observarse en la figura 1 y en la Tabla 1, el espectro cruzado de las proteínas correspondientes al conjunto I presenta 2 picos significativos teniendo en cuenta el criterio  $S/R > 20$ ; el pico de mayor amplitud es a su vez el único pico significativo que aparece en la figura 2 que representa el espectro cruzado del conjunto II.

Ninguno de los dos picos descritos coincide con los cinco picos significativos del espectro cruzado del conjunto III (figura 3).

Como puede verse en la Tabla 2, de los cinco picos significativos de la figura 3, solamente uno de ellos está asociado a una de las cinco frecuencias que poseen los mayores valores de intensidad en los conjuntos I y II (frecuencia 0.3689), las otras cuatro frecuencias, a pesar de ser significativas en el espectro cruzado III, son mucho menos importantes y ello se debe a dos razones fundamentales:

- la media del espectro cruzado III es mucho más baja que la media de los espectros cruzados I y II.
- estas cuatro frecuencias presentan valores de amplitud muy bajos en los espectros cruzados I y II, siendo estos últimos valores mucho menores como promedio que lo reportado para las otras frecuencias en la Tabla 1.

Como puede apreciarse en la Tabla 1, ambos espectros comparten las cinco frecuencias con los mayores valores de amplitud, aunque estos valores y los de S/N no son los mismos. En el caso del espectro cruzado III, salvo una de las frecuencias (ver Tabla 2), las restantes no están entre las cinco de mayor amplitud de los espectros cruzados I y II.

En ninguno de los tres espectros cruzados totales que incluyen 13, 8 y 17 grupos de proteínas quinasas respectivamente, se detectó un pico en la frecuencia reportada por Cosic para las proteínas quinasas.

A partir de los 17 grupos involucrados en el espectro cruzado III se seleccionaron solamente aquellos grupos que contribuyen notablemente a la frecuencia 0.0923. A continuación se multiplicaron entre sí los espectros de cada uno de los ocho grupos escogidos. Como se puede ver en la figura 4, el espectro total resultante solo tiene un pico significativo en la frecuencia mencionada y además, ninguna de las otras cinco frecuencias de mayor valor de amplitud de ese espectro está entre las frecuencias de mayor valor de amplitud analizadas en las Tabla 1 y 2.

## DISCUSIÓN

Uno de los principios básicos del modelo de reconocimiento resonante propuesto por Cosic plantea que a cada función proteica corresponde una y solo una frecuencia característica en el espectro cruzado de potencias de la secuencia proteica obtenida a partir de los pseudopotenciales de interacción ión-electrón de sus aminoácidos constituyentes [2]. Basándose en este modelo, Cosic obtuvo las frecuencias características correspondientes a varios grupos funcionales, entre los cuales se reporta, por ejemplo, que el grupo de las proteínas quinasas tiene una frecuencia característica de 0.42969. Sin embargo, dado el bajo número de proteínas quinasas utilizadas en ese estudio ( $n=8$ ), el valor reportado pudiera no resultar muy confiable; de hecho en el presente trabajo, utilizando 360 proteínas quinasas, no se encontró la frecuencia reportada por Cosic entre las frecuencias con mayores valores de amplitud. Otro hecho que salta a la vista es la presencia de varias frecuencias con valores elevados de amplitud en el espectro y en muchos casos, más de una frecuencia con amplitud significativa en un mismo espectro cruzado, esto se ve claramente observando cualquiera de las tablas. La explicación podría estar dada por al menos una de las siguientes razones:

- Podríamos suponer que en realidad existan varias funciones comunes o varias periodicidades estructurales comunes entre las 360 proteínas estudiadas. Sin embargo, creemos que esto sea muy poco probable debido a que esta muestra de proteínas posee una amplia diversidad; entre ellas hay serin-treonin quinasas y tirosin quinasas pertenecientes a 38 grupos muy diferentes entre sí, que se encuentran además, en organismos distantes filogenéticamente (bacterias, levaduras, mamíferos, etc.)
- Por otra parte, es cuestionable el uso de  $S/N > 20$  como criterio para determinar los picos significativos en el espectro cruzado, puesto que no siempre hemos obtenido un solo pico significativo, siendo el caso más evidente el del espectro cruzado III (ver Tabla 2 y figura 3). Esto se podría explicar porque el  $S/N > 20$  podría ser un valor subestimado con respecto al real, de aquí que se estén viendo picos significativos que en realidad no lo sean. Es evidente que un artefacto de este tipo se produciría más fácilmente cuando la dispersión de los valores con respecto a la media espectral es muy baja. También podría ocurrir que cuando el espectro cruzado tenga una media por encima o por debajo de determinado rango, el límite de discriminación entre picos significativos definido por Cosic fuera realmente superior al valor establecido, de manera que quedarían excluidos los picos significativos adicionales tanto en el espectro de media de orden  $10^{33}$ , como en el espectro de orden de  $10^{-6}$ . Como confirmación de lo anteriormente sugerido, los espectros 3 y 4, que poseen medias del orden de  $10^2$  y  $10^3$ , caerían dentro del supuesto rango al tener solo un pico significativo (ver figuras 3 y 4).

En el supuesto caso de que se diera cualquiera de estas dos posibilidades, entonces no necesariamente habría que cuestionarse el postulado básico de la relación biunívoca

frecuencia-función y el mejor candidato para frecuencia característica sería el pico de frecuencia 0.0923 (como se puede ver, un valor muy alejado del reportado por Cosic: 0.42969), al tener el valor máximo de amplitud en los espectros 1, 2 y 4 (incluyen los tres juntos 29 de los 38 grupos analizados) o una amplitud promedio superior a la de las restantes frecuencias. Según este modelo, una sola frecuencia característica para la actividad quinasa permitiría una coordinación de varios eventos moleculares en los que están involucrados aminoácidos de diferentes valores de EIIP en un solo proceso de transferencia de energía resonante que movilizaría grupos fosfatos desde la fuente al sustrato.

Es necesario considerar también la posibilidad de que a una función biológica determinada le pueda corresponder en realidad más de un pico en el espectro de frecuencias, lo cual pondría en entredicho la relación biunívoca entre la función biológica y la frecuencia característica. La explicación de este fenómeno puede encontrarse en el hecho de que la función 'quinasa' no es un todo único e indivisible, sino que está conformada por varios eventos moleculares (unión de la fuente de fosfato a la enzima, autofosforilación, reconocimiento y fosforilación del sustrato, etc.), en los que están involucrados diferentes aminoácidos asociados a sitios estructurales y catalíticos. En este caso quedaría por determinar cuáles de las frecuencias con mayores valores de amplitud están asociadas con eventos de la actividad quinasa.

## CONCLUSIONES

Este trabajo nos ha permitido calcular un valor de frecuencia característica para superfamilia de las quinasas mucho más representativo que el reportado por Cosic y, por otra parte, al obtener en dos de los espectros totales de los cuatro estudiados más de un pico significativo, sentamos un precedente importante para futuros estudios, pues esto podría ser consecuencia de que no es válido al menos uno de los siguientes aspectos del modelo, a saber, el criterio para discriminar los picos significativos de los que no lo son ( $S/N > 20$ ), o el postulado fundamental del MRR de que a cada pico significativo le corresponde una y solo una función biológica.

## REFERENCIAS

- [1] Cosic I. IEEE Trans. Biom. Eng. Vol. 41, No12, Dec, 1994
- [2] Cosic I Biotechnology Vol.13 (1995)236-239
- [3] Doolittle R.F. Annu.Rev.Biochem.(1995),64:287-314
- [4] Brenner S.E.; Chothia C.; Hubbard T.D.P. Curr.Op.Struct.Biol. (1997), 7:369-376
- [5] Cosic I.;Pavlovic M.;Vojisavljevic V. Biochimie 71 (1989) 333-342
- [6] Cosic I.;Nesic D. Eur. J. Biochem. Vol.170 (1988)247-252
- [7] Krsmanovic V; Biquard JM; Sikorska-Walker M; Cosic I; Desgranges C; Trabaud MA; Whitfield JF; Durkin JP; Achour A; Hearn MT J. Pept. Res., 52(5):410-20 1998 Nov
- [8] Cosic I; Drummond AE; Underwood JR; Hearn MT. Mol Cell Biochem, 130(1):1-9 1994 Jan 12
- [9] Ciblis P.;Cosic I. J Theor Biol, 184(3):331-8 1997 Feb 7
- [10] Hanks S Quinn AM Meth. Enz. (1991) 200:38-62

## Figuras y Tablas

	EC I( $\mu=6.9513e+030$ )	EC II( $\mu=3.5108e+024$ )	EC III( $\mu=5.2982e-006$ )
Frecuencia 0.0923	$1.69*10e33$ ♣	$1.05*10e27$ ♣	$1.69*10e-18$
Frecuencia 0.2767	$3.89*10e32$ ♣	$3.5*10e11$	$2.7*10e-13$
Frecuencia 0.1846	$7.12*10e30$	$2.3*10e24$	$8.7*10e-16$
Frecuencia 0.3692	$4.51*10e23$	$1.46*10e19$	$1.76*10e-5$
Frecuencia 0.4615	$4.1*10e23$	$5.45*10e2$	$4.8*10e-21$

♣ valor significativo (S/R>20)

$\mu$  media espectral

EC espectro cruzado

**Tabla 1.** En esta tabla se muestran los valores de amplitud de las cinco frecuencias de mayor S/N pertenecientes al espectro cruzado (EC) de mayor media. Los grupos están formados como sigue: EC I (13 grupos), EC II (8 grupos), EC III (17 grupos).

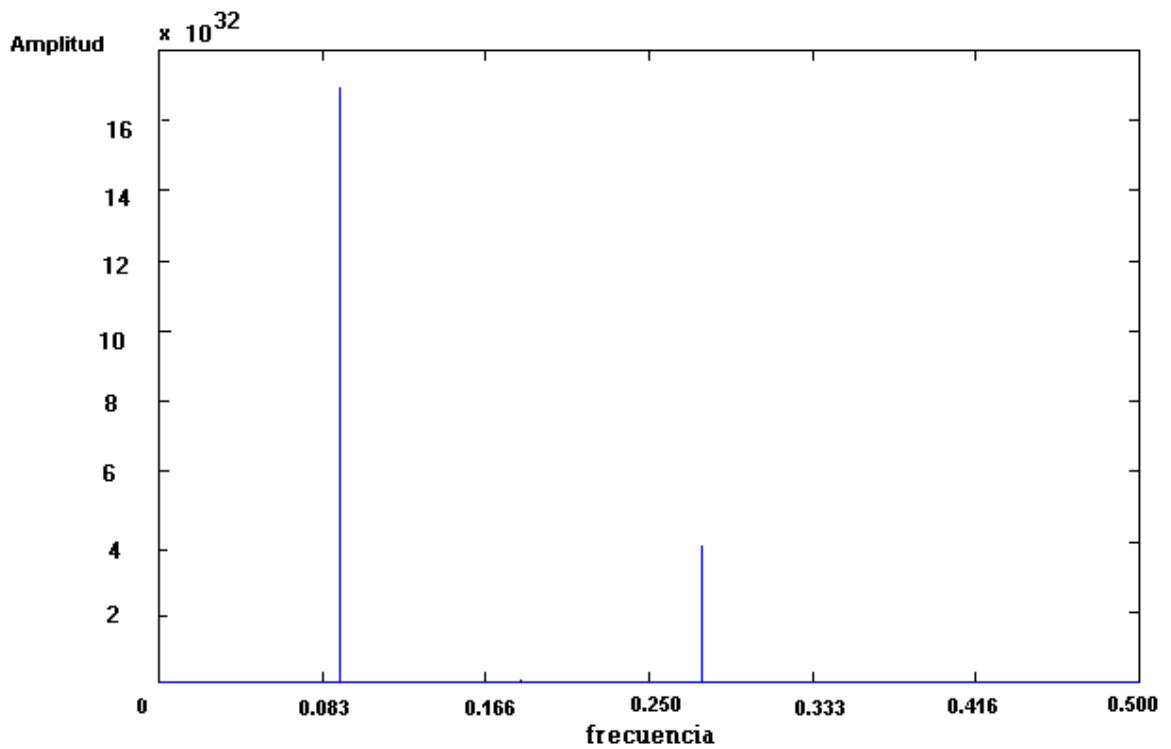
	EC III ( $\mu=5.2982e-006$ )	EC II ( $\mu=3.5108e+024$ )	EC I( $\mu=6.9513e+030$ )
Frecuencia 0.0528	$7.3*10e-4$ ♣	$9*10e-41$	$5.1*10e-13$
Frecuencia 0.0983	$5.25*10e-4$ ♣	$5.2*10e-29$	11.1
Frecuencia 0.1636	$2.8*10e-4$ ♣	$3.4*10e-31$	$3.91*10e-5$
Frecuencia 0.0533	$3.8*10e-5$ ♣	$6*10e-36$	$5*10e-13$
Frecuencia 0.3692	$1.76*10e-5$ ♣	$1.46*10e19$	$4.51*10e23$

♣ valor significativo (S/R>20)

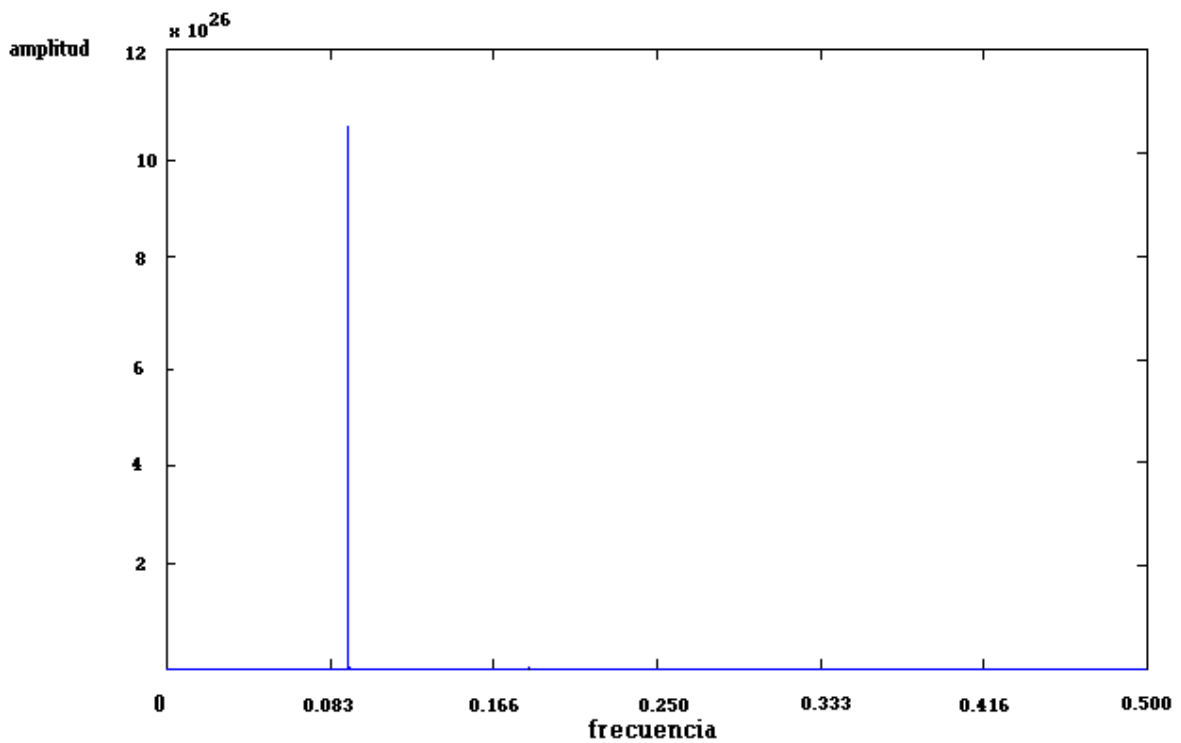
$\mu$  media espectral

EC espectro cruzado

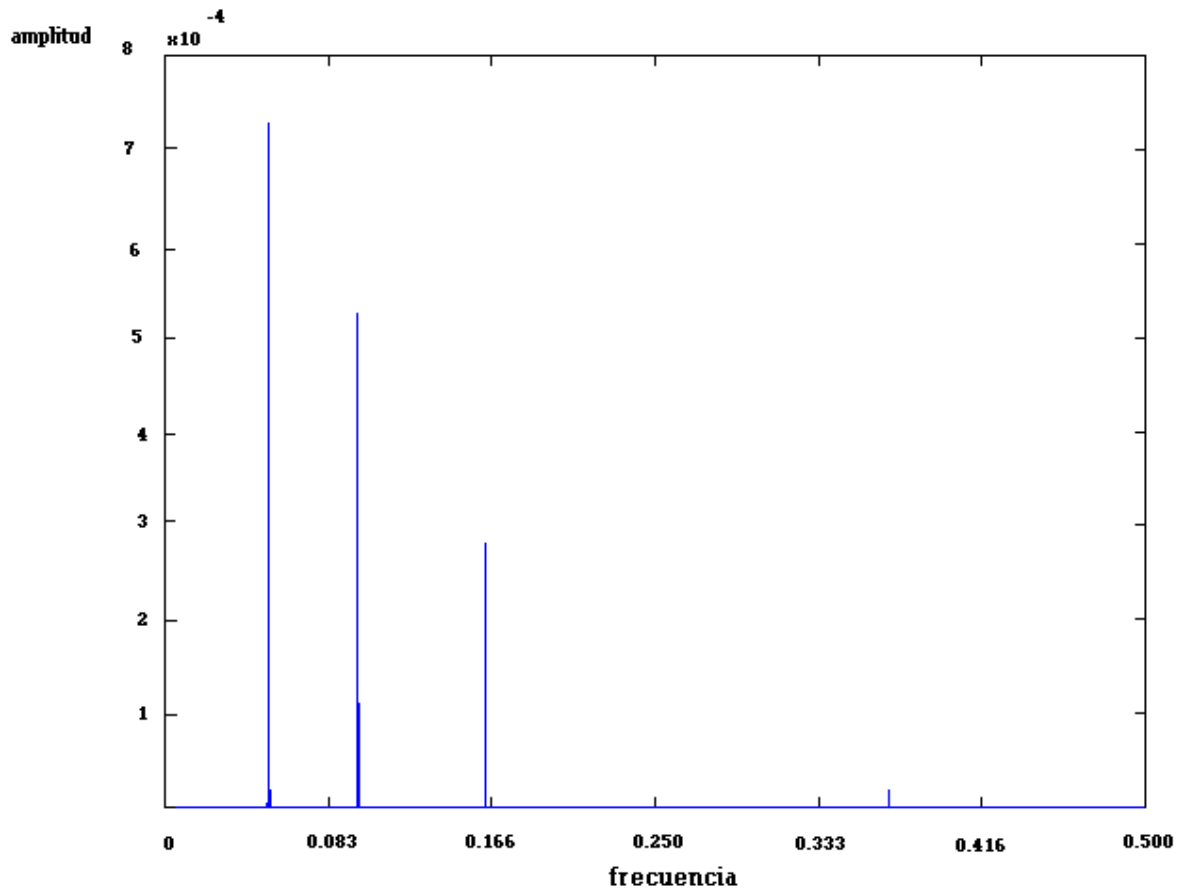
**Tabla 2.** Esta tabla muestra los valores de amplitud correspondientes a las cinco frecuencias significativas del espectro cruzado III (ECIII).



**Figura 1.** Espectro cruzado total correspondiente al conjunto I (13 grupos de proteínas quinasas). Nótese la presencia de dos picos significativos; el primero a frecuencia 0.0923 y el otro a 0.2767. Debido a problemas de escala no se logra apreciar los restantes picos de este espectro; algunos de ellos se reflejan en las Tablas 1 y 2.

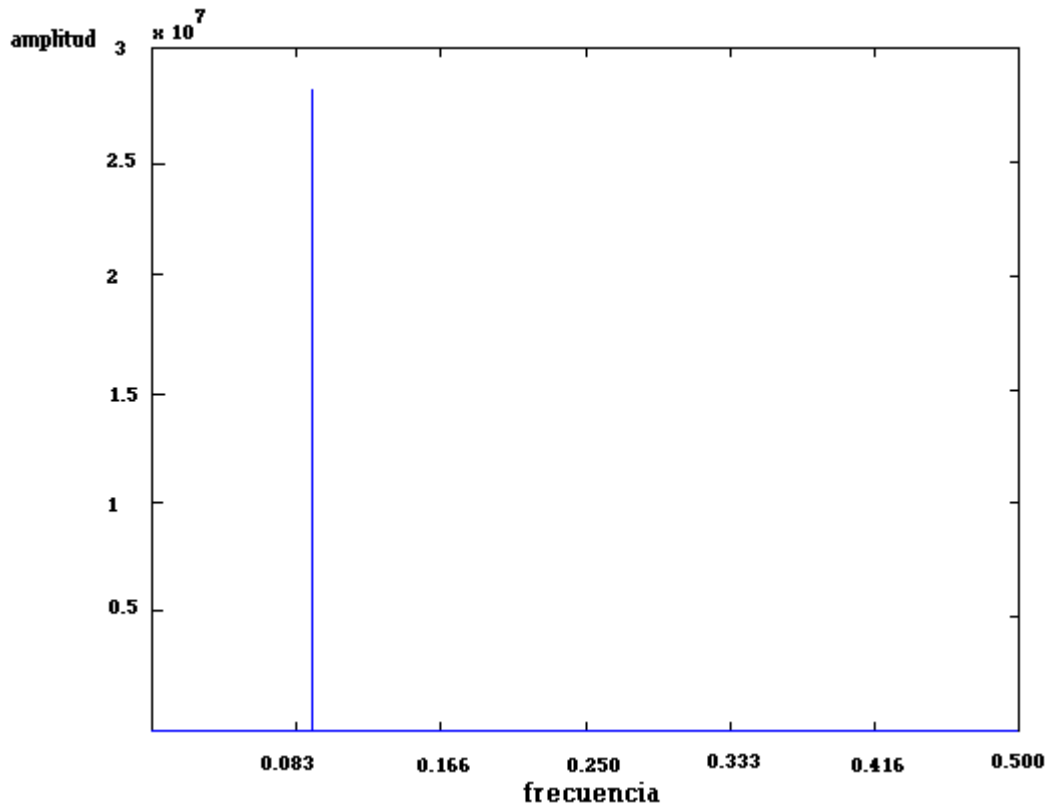


**Figura 2.** Espectro cruzado total II (8 grupos de proteínas quinasas). Nótese la presencia de un pico significativo a frecuencia 0.0923.



**Figura 3.** Espectro cruzado total correspondiente al conjunto III (17 grupos de proteínas quinasas). Nótese la presencia de cinco picos significativos a frecuencias 0.0528, 0.0533, 0.0983, 0.1636, 0.3689. Dada la cercanía de las frecuencias de los dos primeros picos y a la escala usada en esta figura, resulta muy difícil apreciar una separación entre ellos.





**Figura 4.** Espectro cruzado de los ocho grupos que contribuyen positivamente a la magnitud del pico de frecuencia 0.0923, los cuales fueron seleccionados a partir de los 17 grupos del espectro cruzado total III. Estos grupos son, a saber, el OPK VI, CaMK other, CMGC V, OPK VIII, CMGC other, PTK XVIII, PTK IV y OPK XII (según el esquema de clasificación de Hanks y Quinn, 1991)