

UCH-FC
Biotecnología
V 152
C. 1



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS
ESCUELA DE PREGRADO

“Generación de un vector de expresión para células CHO en base a regiones del gen Rpl14, junto a elementos de promotores *Core*”

Seminario de Título entregado a la Universidad de Chile en cumplimiento parcial de los requisitos para optar al Título de Ingeniero en Biotecnología Molecular

Luis Alfonso Valenzuela Villa



Director del Seminario de Título: Dr. Juan Carlos Aguillón Gutiérrez.

Co-Directora del Seminario de Título: Dra. María Carmen Molina Sampayo.

Patrocinante Seminario de Título: Dr. Miguel Allende Connelly.

INFORME DE APROBACIÓN SEMINARIO DE TÍTULO

Se informa a la Escuela de Pregrado de la Facultad de Ciencias, de la Universidad de Chile que el Seminario de Título, presentado por el Sr. Luis Alfonso Valenzuela Villa

“Generación de un vector de expresión para células CHO en base a regiones del gen Rpl14, junto a elementos de promotores Core”

Ha sido aprobado por la Comisión de Evaluación, en cumplimiento parcial de los requisitos para optar al Título de Ingeniero en Biotecnología Molecular

Dr. Juan Carlos Aguill[on]
Director Seminario de Título

Dra. Maria Carmen Molina
Co-Director Seminario de Título

Dr. Miguel Allende Connelly
Patrocinante Seminario de Título

Comisión de Evaluación

Dr. Álvaro Glavic Maurer
Presidente Comisión

Dr(a). Claudia Stange Klein
Evaluador

Santiago, Marzo 2016



En un 24 de diciembre de no hace muchos veranos atrás, nació Luis Valenzuela Villa. Tranquilo pero curioso desde siempre, el pequeño Luis entró a la Escuela Nacional de Artes Gráficas, y luego a la Universidad de Chile. Después de un proceso de ir y venir en carreras, finalmente, se decidió por estudiar Ingeniería en Biotecnología Molecular.

En esta fascinante etapa he conocido a seres muy valiosos que actualmente siguen presentes, aportando desde sus experiencias sabiduría a mi vida. En todo este periodo, miles de cosas han cambiado en mi vida, siempre para mejor, siempre rescatando lo valioso de todo. Después de tantos años, claramente ya no soy el mismo pequeño Luis, aunque siga pequeño. Pero la curiosidad y las ganas de seguir creciendo en todos los aspectos de la vida siguen intactos. Gracias a todos los que han aportado un granito de arena para formar quién soy actualmente.

“Sí algo tiene solución,
¿qué necesidad hay de preocuparse?,

Y si no la tiene,
tampoco sirve de nada hacerlo.”

Shantideva.

A mi madre Julia Inés Villa Fuentes.

AGRADECIMIENTOS

Primero que todo agradecer al Centro de InmunoBiotecnología, del Programa de Inmunología, del Instituto de Ciencias Biomédicas de la Facultad de Medicina, de la Universidad de Chile, especialmente al Dr. Juan Carlos Aguillón y la Dra. María Carmen Molina, quienes me dieron la oportunidad de realizar mi unidad de investigación y el presente seminario de título en su laboratorio. Gracias por tan valiosa oportunidad y por su infinita paciencia y comprensión en este largo período.

También quiero agradecer el apoyo de Roberto, quien me enseñó gran parte de todo lo que aprendí durante este proceso, además de ayudarme a sobrellevar una serie de dificultades que fueron surgiendo durante el desarrollo del trabajo. A Matías, Norberto y Lorenzo quienes aportaron con sus útiles sugerencias en las reuniones de trabajo.

Gracias a personas que quizás ya no estén ni sean los mismos de antes, pero que hicieron grata mi estancia en el laboratorio, a Pía, Vero, Carito, Agnes, Daniel, Diego, Octavio, Jaxaira, Javier, Perry, Gabi, Nancy y a todos con quienes compartí un almuerzo, una comida de fiestas patrias o navidad, a todos por quienes desarrollé un cariño.

A mis amigos, Yuly, Ariel, Carito, Naty, José, Sayaka, Karlita, Geri, Jorge, Pachy, Nico, Nacho, que siempre ha estado brindándome energía y vitalidad para seguir adelante en momentos difíciles. A mi familia; a mis hermanos, sobrinas y papá que me apoyaron en esta etapa universitaria. A Raúl por apoyarme estos últimos dos años y siempre incentivarme a terminar el proceso. Y por supuesto, especialmente agradecido estoy de mi madre, que siempre ha estado de algún modo ahí. Gracias a todos por su presencia en mi vida.

ÍNDICE DE CONTENIDOS

RESUMEN.	1
1. INTRODUCCIÓN	3
1.1 Productos Biofarmacéuticos.	3
1.2 La transcripción y el promotor <i>core</i> : el sitio de inicio.	4
1.3 Motivos canónicos de PCs.....	8
1.4 Promotores de genes codificantes de proteínas ribosomales y el motivo 5'TOP.....	10
1.5 Importancia del primer intrón y primer exón en los niveles de transcripción.	12
Planteamiento del Problema.	16
HIPÓTESIS.	17
OBJETIVO GENERAL.	17
OBJETIVOS ESPECÍFICOS.	17
2. MATERIALES Y MÉTODOS.	18
2.1. Materiales.	18
2.1.1 Cepa Bacteriana.	18
2.1.2 Células eucariontes.	18
2.1.3 Plasmidios.	19
2.1.4 Oligonucleótidos.	20
2.1.5 Enzimas.	23
2.1.6 Marcadores de ADN.	23
2.1.7 Otros reactivos y sistemas comerciales.	23

2.1.8 Soluciones y Medios de cultivo.....	24
2.2 Métodos.	25
2.2.1 Generación de vectores “pGL Core 1-8”.	25
2.2.1.2 Electrotransformación de <i>E. coli</i> DH5 α	28
2.2.1.3 PCR Colonia.	28
2.2.1.4 Extracción ADN plasmidial.	29
2.2.2.1 Generación de vector “pGEMT Easy-EIE”.	29
2.2.2.2 Clonamiento de la región EIE en los vectores “pGL core 1-8”.	30
2.2.3 Optimización cantidad PEI/ADN en transfección de células CHO-K1.	31
2.2.4 Transfección células CHO-K1.	32
2.2.5 Co-Transfección vectores portando los genes reporteros Luc.	33
2.2.6 Medición de luminiscencia.	34
2.2.7 Cálculo de la Concentración de ADN por espectrofotometría.	35
2.2.8. Electroforesis en Gel de Agarosa.	35
2.2.9. Reacción de ligación.	35
2.2.10. Extracción de DNA plasmidial de <i>E. coli</i>	36
2.2.11. Análisis estadístico.	36
2.3. Recursos Materiales y financiamiento.	36
3 RESULTADOS.....	37

3.1 Objetivo Específico 1: Generar vectores de expresión que porten promotores <i>core</i> con los elementos MTE, DPE, caja TATA y TOP en distintas combinaciones, y luego incorporar la región EIE del gen Rpl14	37
3.2 Objetivo Específico 2: Evaluar la actividad transcripcional de los vectores generados mediante ensayos reporteros utilizando enzimas luciferasas, luego transfectar transientemente células CHO-K1	47
3.2.1 Optimización cantidad PEI/ADN en transfección de células CHO-K1.	47
3.2.2 Co-Transfección vectores Luc y medición actividad luciferasa.	48
4. DISCUSIÓN.	52
4.1 Optimización cantidades de PEI y ADN.	54
4.2 Actividad luciferasa vectores “pGL Core” y “pGL Core-EIE”.	56
5. CONCLUSIONES.	63
6. BIBLIOGRAFÍA.	63

ÍNDICE DE FIGURAS

Figura 1. Reconocimiento del Promotor Core (PC).	5
Figura 2. Ensamble del complejo de Pre-iniciación (CPI).	6
Figura 3. Figura 3. Actividades relativas de luminiscencia de cultivos celulares de origen animal (HeLa) transfectados con vectores de expresión con PCs sintéticos dirigiendo la expresión del gene reportero.....	10
Figura 4. Vector de expresión construido en base a promotor de gen ribosomal.	12
Figura 5. Enriquecimiento de nucleótidos y dinucleótidos en el primer intron.	13
Figura 6. Largo del primer exón y presencia de TFIID o ARNP1I.	15
Figura 7. Esquema generación vectores "pGL Core".	27
Figura 8. Secuencia de ADN de los 8 promotores core.....	28
Figura 9. PCR usando como ADN molde los vectores "pGL Core1-8" purificados.	38
Figura 10. Alineamiento de secuencias teóricas y experimentales de los promotores core....	39
Figura 11. Chequeo incorporación región EIE en vector pGL TATA.	44
Figura 12. Esquemas vectores "pGL Core-EIE".	45
Figura 13. Optimización cantidades de PEI y ADN.	47
Figura 14. Actividad luciferasa vectores "pGL Core 1-8".	49
Figura 15. Actividad luciferasa vectores "pGL Core 1-8- EIE".	50

ÍNDICE DE TABLAS

Tabla 1. Plasmidios comerciales utilizados en el presente seminario.	19
Tabla 2. Plasmidios generados en el presente seminario de título.	20
Tabla 3. Oligonucleótidos utilizados y/o diseñados en este trabajo.	21
Tabla 4. Cantidades de ADN y PEI (μ g), clasificados en 3 niveles.	32

LISTA DE ABREVIATURAS

AdML	: Promotor mayor tardío de Adenovirus.
ADN	: Ácido desoxirribonucleico.
ADNdh	: ADN de doble hebra.
ANOVA	: Análisis de la varianza.
ARN	: Ácido ribonucleico.
ARNPII	: ARN polimerasa II.
ARNr	: Ácido ribonucleico ribosomal.
BHK	: Riñón de hámster bebé.
CDT	: Dominio C-terminal.
ChIP-seq	: Inmunoprecipitación de cromatina seguido de secuenciación.
CHO	: Ovario de Hamster chino.
CMV	: Citomegalovirus.
CPI	: Complejo de pre-iniciación.
D.O.	: Densidad Óptica.
DMEM	: Medio Eagle modificado de Dulbecco.
DNasaI	: Desoxi-ribonucleasa I.
dNTP	: Desoxinucleótido trifosfato.
DPE	: Elemento de promotor río abajo.
EIE	: Región con Exón 1, Intrón 1 y Exón 2 de gen Rpl14
ENCODE	: Enciclopedia de elementos de ADN.
GFP	: Proteína Fluorescente Verde.
GMP	: Prácticas de buen manufacturado.
HEK	: Riñón de embrión humano.
HSV-TK.	: Promotor del gen Timidina Kinasa del Virus del Herpes Simple.
IBT	: InmunoBioTecnología.
INR	: Iniciador.

IPTG	: Isopropil β -D-1-tiogalactopiranosido.
Kb	: Kilobase.
MTE	: Motivo de 10 elementos.
pb	: Pares de bases.
PBS	: Tampón fosfato salino.
PC	: Promotor "Core".
PCR	: Reacción en cadena de la Polimerasa.
PEI	: Polietilenimina.
PNK	: Kinasa polinucleótidos.
SFB	: Suero fetal bovino.
SIT	: Sitio de inicio de la transcripción.
TAF	: Factor asociado a proteína de unión a caja TATA.
Taq	: <i>Thermus aquaticus</i> .
TBP	: Proteína de unión a caja TATA.
TFII	: Factor de Transcripción de ARN polimerasa II.
TOP	: Tracto de oligopirimidinas.
TRF2	: Factor relacionado a TBP 2.

RESUMEN

Mediante la ingeniería genética se han generado diversas proteínas recombinantes con aplicación terapéutica, conocidas también como Biofármacos. Dentro de éstos, se han desarrollado nuevas formas de anticuerpos terapéuticos, cuyas aplicaciones de mayor éxito están en el tratamiento del cáncer y de enfermedades autoinmunes.

Los anticuerpos recombinantes son expresados principalmente en células de origen animal como la línea celular CHO proveniente de tejido ovárico de hámster chino, utilizando vectores de expresión cuyos genes de interés son comandados por promotores virales como el promotor de Citomegalovirus (CMV). Sin embargo, estos genes recombinantes disminuyen su expresión por silenciamiento luego de sucesivos subcultivos. Para evitar esto, un campo de investigación es el mejoramiento y generación de nuevo vectores de expresión con promotores optimizados.

La elección de promotores de genes endógenos de los sistemas de cultivo que presenten una elevada expresión, junto a modificaciones racionales de los elementos del promotor *core*, constituyen una de las posibles aproximaciones a dicha problemática.

Se ha demostrado que el gen Rpl14 que codifica la proteína ribosomal L14 tiene una elevada expresión en tejido ovárico, además presenta en su promotor un iniciador rico en polipirimidinas denominado 5'TOP (o TCT, o simplemente TOP, de ahora en adelante) típico de genes de proteínas ribosomales, que le brinda características esenciales para el control traduccional, con el propósito de mantener un nivel adecuado de la maquinaria de síntesis proteica. El promotor del gen Rpl14 no contiene elementos de promotor *core*

como caja TATA, INR, DPE ni MTE, los cuales se han visto pueden actuar sinérgicamente en líneas celulares de origen de humano.

Frente a esto, se propuso utilizar como base la región correspondiente al promotor *core* del gen Rpl14 que contiene el motivo TOP, incorporando además elementos como caja TATA, INR, MTE y DPE en diferentes conformaciones, para posteriormente añadir la región que contenía parte del primer exón no traducido, primer intrón y parte del segundo exón no traducido de dicho gen (en adelante, región EIE). Con estos elementos se generaron diferentes promotores con el objetivo de clonarlos en un vector de expresión que portase el gen reportero sin promotor que codifica la enzima luciferasa, y así estudiar el nivel de expresión de dichas construcciones mediante transfecciones transientes en células CHO-K1. De ese modo, se pudo establecer que en esta línea celular existe una complementariedad funcional entre la caja TATA y TOP, TATA e INR, y este último par con los motivos MTE-DPE, no así entre la combinación TATA-TOP y MTE-DPE. Ello implica que algunos motivos que no están naturalmente juntos en promotores, pueden funcionar cooperativamente a pesar de que sean diferentes las proteínas que los unen para reclutar a la maquinaria transcripcional.

1. INTRODUCCIÓN

1.1 Productos Biofarmacéuticos.

Los productos biofarmacéuticos son reactivos clínicos, vacunas y drogas que se elaboran usando técnicas de ingeniería genética para usos terapéuticos, tratamientos preventivos y diagnóstico (Zhu, 2012). Dentro de éstos se encuentran las proteínas recombinantes, incluyendo los anticuerpos monoclonales, cuyas aplicaciones de mayor éxito están en el tratamiento del cáncer y de enfermedades autoinmunes, debido a su elevada especificidad y afinidad.

La producción industrial de anticuerpos monoclonales se lleva a cabo generalmente usando sistemas de cultivo celular de origen mamífero, ya que éstos son capaces de realizar un correcto plegamiento de las proteínas, además de complejas modificaciones post-traduccionales, esenciales para la actividad biológica del producto (Hauser & Wagner, 1997). Los principales sistemas que suplen la industria biofarmacéutica son líneas celulares como las derivadas de células de ovario de hámster chino; CHO (del inglés, *Chinese Hamster Ovary*), de riñón de hámster; BHK (del inglés, *Baby Hamster Kidney*), de hibridomas murinos, y de células humanas como HEK (del inglés, *Human Embryonic Kidney*). De estos sistemas, el más usado es el que emplea las células CHO, aisladas en los años '50 (Zhu, 2012). Estos sistemas son transfectados con vectores de expresión portando los genes de interés, los cuales generalmente están dirigidos por promotores virales como CMV (*Citomegalovirus*) o SV40 (del inglés, *Simian vacuolating Virus 40*). Sin embargo, la transcripción de estos genes recombinantes

disminuye considerablemente luego de sucesivos subcultivos, ya que ocurre un silenciamiento debido a metilaciones de la región promotora viral y/o modificaciones en los residuos aminoacídicos de las histonas en las proximidades a dichos promotores, entre otras causas (Yang y col., 2010).

Es por esto que un campo activo de investigación es la generación de vectores de expresión cuyos promotores sean capaces de mantener una producción sostenible de proteínas recombinantes. Para lograr este objetivo, una de las aproximaciones es la modificación de la secuencia de ADN del promotor, dada su importancia en relación con la regulación de la transcripción.

1.2 La transcripción y el promotor *core*: el sitio de inicio.

La expresión de decenas de miles de genes dentro de la célula está regulada durante el crecimiento, desarrollo y en respuesta a estímulos ambientales. En eucariontes, la transcripción de genes que codifican proteínas es mediada por la enzima ARN polimerasa II (ARNPII). El proceso canónico comienza con la unión de un activador a secuencias de ADN específicas en regiones denominadas potenciadoras o “*enhancers*”, lo cual luego promueve el reclutamiento secuencial de los factores de transcripción generales (TFII D, A, B, E, H y F) y la ARNPII a los promotores de los genes blancos (Figura 1) (Thomas & Chiang, 2006).

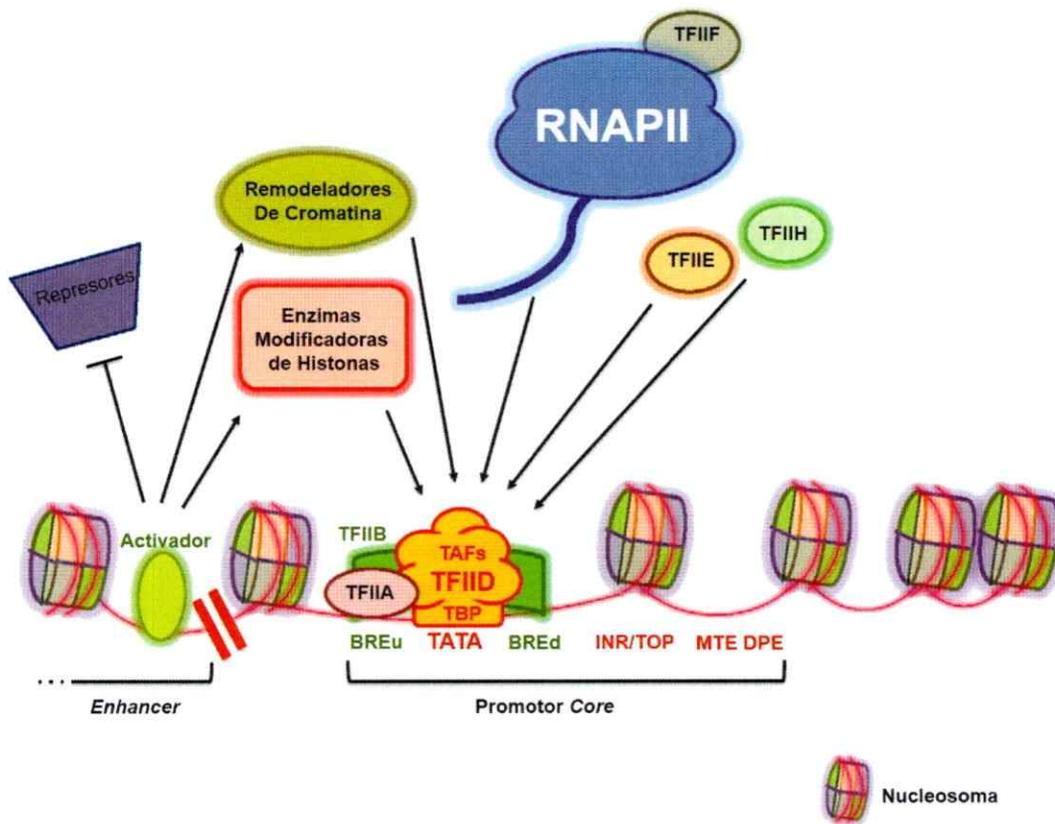


Figura 1. Reconocimiento del Promotor Core (PC). En alguna región potenciadora o “*enhancer*”, podría unirse un activador que eventualmente inhiba represores, o active complejos remodeladores de cromatina y/o modificadores de histonas, lo cual favorecería la unión del factor TFIID a los elementos presentes en los promotores, permitiendo el reclutamiento secuencial del resto de los factores generales junto a la ARNPII (Modificado de Shandilya & Roberts, 2012).

Una transcripción productiva está determinada por la formación del complejo de pre-iniciación (CPI) funcional en el promotor (Figura 2), seguido por la apertura de esta región, el escape desde allí de la RNAPII elongando los transcritos, y finalmente deteniéndose en determinados sitios de término, donde mediante la formación de un bucle en el DNA, los genes podrían estar sujeto a sucesivas rondas de transcripción basado en la eficiencia del ensamble de CPI y de la reiniciación (Yudkovsky y col., 2000; El Kaderi y col., 2009). Importante es mencionar que el procesamiento (que incluye el “*capping*”, seguido del corte y empalme, para escindir intrones) de los transcritos nacientes ocurre co-transcripcionalmente.

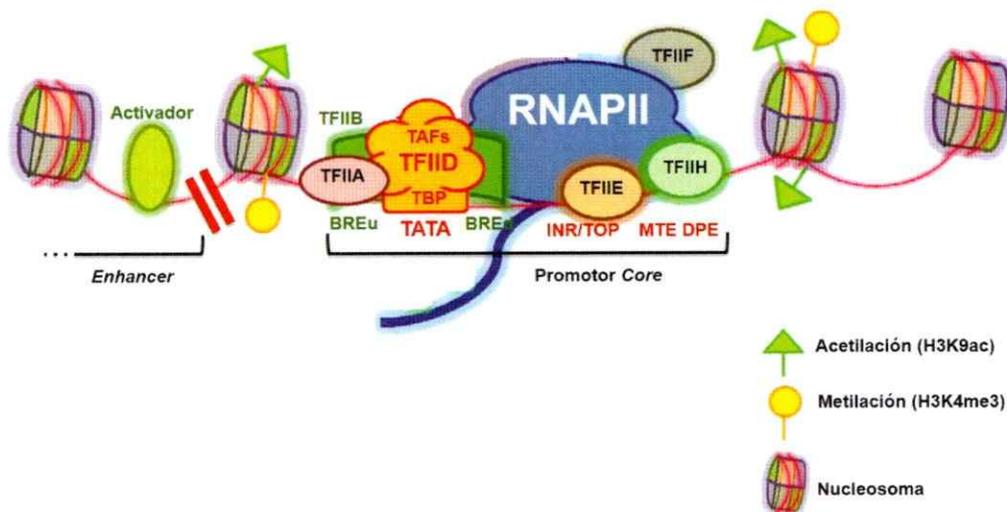


Figura 2. Ensamble del complejo de Pre-iniciación (CPI). Luego de la unión del factor TFIID a los elementos del promotor *core*, se favorece el reclutamiento secuencial de TFIIB, A, E, H y F junto a la RNAPII, formándose el CPI. Esto no necesariamente asegura el inicio de la transcripción, pero estabiliza a la polimerasa para que pueda ser fosforilada en la subunidad CDT, lo cual activa y permite llevar a cabo la elongación de la transcripción (Modificado de Shandilya & Roberts, 2012).

Existen varios puntos críticos en los procesos descritos anteriormente, por lo que en su conjunto el proceso está regulado por una amplia variedad de factores, incluyendo los factores de transcripción generales, proteínas de unión específica a promotores y/o *enhancers*, factores co-reguladores, complejos modificadores y remodeladores de cromatina, entre otros (Venters & Pugh, 2009). En última instancia, todas las señales de estos factores convergen al promotor durante la iniciación de la transcripción, proceso que podría ser de dos tipos principalmente; una iniciación focalizada o por otra parte una iniciación dispersa. En la primera, la transcripción comienza desde un único nucleótido o dentro de un rango restringido de nucleótidos, mientras que en la iniciación dispersa hay varios sitios débiles de inicio de la transcripción a lo largo de una región más amplia que la anterior (Lemon B. & Tjian, 2000; Juven-Gershon & Kadonaga, 2010). Los estímulos de los factores mencionados anteriormente, convergen finalmente en el promotor *core* (en adelante, "PC"); región que generalmente abarca alrededor de 100 pb de longitud conteniendo el sitio de inicio de la transcripción, y es por ello que dicha región resulta crucial en la expresión génica. Inicialmente se pensó que todos los PCs funcionaban vía un mecanismo común, pero se ha demostrado que éste es un componente estructural y funcionalmente muy diverso, dada la presencia de subregiones, elementos o motivos que le confieren propiedades específicas.

1.3 Motivos canónicos de PCs.

Estudios en PCs focalizados han llevado al descubrimiento de diversas secuencias, (denominadas desde ahora elementos o motivos) con características conservadas entre los diferentes reinos y/o especies. Uno de estos elementos es conocido como caja TATA, el cual posee con una secuencia consenso igual a TATA(A/T)AA, y cuya posición está ubicada alrededor del nucleótido -30 con respecto al sitio de inicio de la transcripción (SIT). Está presente desde arqueas hasta mamíferos, encontrándose en éstos últimos en aproximadamente sólo el 10-15% de todos los promotores (Kim y col., 2005; Carninci y col., 2006; Cooper y col., 2006). La secuencia de este motivo es reconocida por la subunidad TBP (del inglés, *TATA-binding protein*) del complejo TFIID (del inglés, *Transcription Factor II D*) de la ARN polimerasa II (ARNPII). Junto a la caja TATA, tanto río arriba como río abajo de ésta podría encontrarse otro motivo conocido como BRE (del inglés, *TFII B Recognition Element*), el cual se ha reportado que puede aumentar o disminuir los niveles basales de transcripción logrados por la caja TATA, dependiendo de las secuencias adyacentes (Deng & Roberts, 2005).

Otro motivo identificado en PCs de humanos, moscas y virus es INR (del inglés, *Initiator*), el cual abarca el sitio de inicio de transcripción, y es probablemente el motivo más común presente en promotores de genes de humanos y levaduras (Yang y col., 2007). Es reconocido por las subunidades TAF1 y TAF2 (del inglés, *TBP-Associated Factor*) del complejo TFIID de la ARNPII. En humanos su secuencia consenso es (C/T)(C/T)A(A/C/T/G)(A/T)(C/T)(C/T) y en moscas es TCA(T/G)T(C/T). A pesar de

que las diferencias entre las dos secuencias, se ha demostrado la funcionalidad del INR de moscas en cultivo de origen mamífero (Juven-Gershon y col., 2006).

Otro elemento descrito en la literatura es MTE (del inglés, *Motif Ten Element*), el cual se ha identificado en humanos y moscas río abajo de INR, entre las posiciones +18 y +27 con respecto al sitio de inicio de la transcripción, y cuya secuencia consenso es C(C/G)A(A/G)C(C/G)(C/G)AAC. Inmediatamente río abajo de MTE, entre las posiciones +28 y +33, podría encontrarse el motivo DPE (del inglés, *Downstream Promoter Element*), cuya secuencia consenso es (A/G)G(T/A)(C/T)(A/C/G)T. La unión de ambos elementos se han asociado a las subunidades TAF6 y TAF9 del complejo TFIID (Lim y col., 2004; Theisen y col., 2010). Tanto en estudios *in vitro* utilizando la maquinaria de transcripción y traducción purificadas, como en estudios *in vivo* en líneas celulares de origen mamífero, se ha visto que la expresión de GFP (del inglés, *Green Fluorescent Protein*) como reportero, comandado por un PC sintético portando una caja TATA, DPE y MTE, presenta mayor actividad en comparación a otros PCs de origen viral (Figura 3), y también mayor actividad en comparación a PCs en donde está presente sólo uno de los tres motivos a la vez, indicando que existe una complementariedad funcional entre dichos motivos (Juven-Gershon y col., 2006; Malecová y col., 2007).

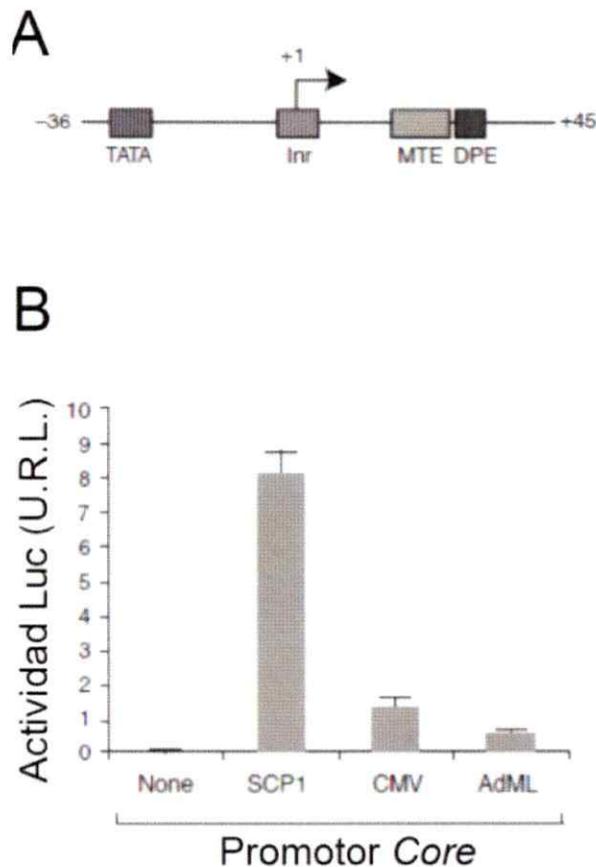


Figura 3. Actividades relativas de luminiscencia de cultivos celulares de origen animal (HeLa) transfectados con vectores de expresión con PCs sintéticos dirigiendo la expresión del gene reportero.

A) Diagrama del promotor *core* sintético construido a partir de elementos provenientes de diferentes especies; caja TATA de Citomegalovirus, INR de Adenovirus, MTE y DPE de mosca. **B)** Ensayo actividad Luciferasa transfectando células HeLa con vector pGL3 sin promotor (“none”), con el promotor *core* descrito en A) (“SCP1”), con el PC de CMV, y finalmente con el PC de adenovirus (“AdML”) (Modificado de Juven-Gershon y col., 2006)

1.4 Promotores de genes codificantes de proteínas ribosomales y el motivo 5'TOP.

Por otra parte, los ribosomas son organelos que catalizan la síntesis de proteínas en todos los seres vivos. En eucariontes, éstos están compuestos de cuatro moléculas de ARN (ARNr) y 79 proteínas distintas. En mamíferos, los genes que codifican los ARNr están en múltiples copias y agrupados en pocos *loci*, mientras que los genes que codifican las proteínas ribosomales están en una sola copia y dispersos en el genoma (Uechi y col., 2001).

En mamíferos, un 65% de los genes que codifican proteínas ribosomales carecen de caja TATA en su promotor, y ninguno cuenta con elementos como MTE o DPE. Además, todos los ARNm de estos genes tienen en su extremo 5' una secuencia denominada 5'TOP (del inglés, *5' Terminal OligoPyrimidine tract*) (Roepcke y col., 2006), conformada sólo de pirimidinas comenzando con una citocina (Levy y col., 1991), y cuyo reconocimiento no es por parte del complejo TFIID (Parry y col., 2010) como los otros motivos de promotores *core* descritos anteriormente, sino que es reconocido por el factor TRF2 (del inglés, *TBP-Related Factor 2*) (Wang y col., 2014). Además, se ha demostrado que esta secuencia es esencial en el control traduccional de dichos transcritos, para así mantener una cantidad adecuada de la maquinaria encargada de la síntesis proteica (Levy y col., 1991).

Dada la importancia que posee el ribosoma, y por ende de las proteínas que lo conforman, la expresión de éstas es mayor cuando hay abundancia de nutrientes (Hamilton y col, 2006), como lo que ocurre generalmente en los medios de cultivo de líneas celulares para la elaboración de productos biofarmacéuticos. Teniendo esto en consideración, no sorprende que en la actualidad exista patentado un vector de expresión generado en base al gen que codifica la proteína ribosomal S30. Este vector cuenta con una región promotora formada por el PC de dicho gen, al cual se le agregó una caja TATA, seguido del primer exón, primer intrón y parte del segundo exón. Se demostró su actividad transcripcional en células HEK293 (de origen humano), usando como gen

reportero al gen que codifica GFP, y se obtuvo magnitudes de expresión comparables con la del promotor del virus CMV (Figura 4) (Hitti y col., 2010).

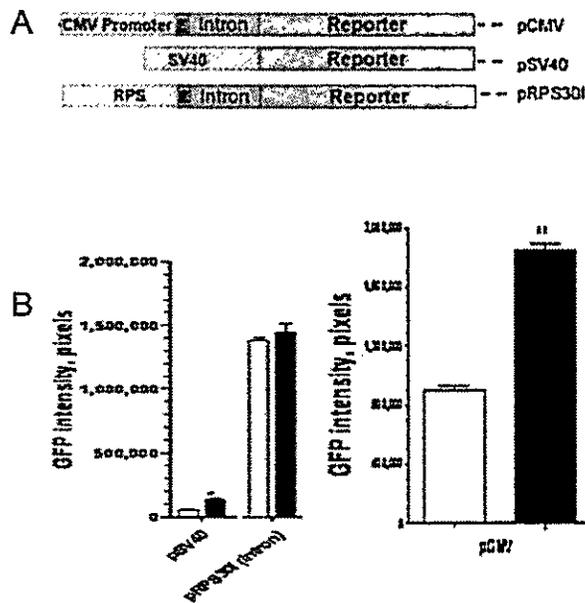


Figura 4. Vector de expresión construido en base a promotor de gen Rps30.

A) Diagrama de vectores de expresión construidos con promotores de CMV, de SV40 y del gen Rps30, de arriba hacia abajo, respectivamente.

B) Intensidad de GFP (en pixeles). La actividad alcanzada por el vector con el promotor SV40 es menor a la obtenida por el promotor RPS30, el cual alcanza casi 1,5 millones de pixeles, tanto en ausencia como en presencia de TNF. El vector construido en base a CMV alcanza alrededor de 800 mil y casi 2 millones de pixeles, en las mismas condiciones, respectivamente (Modificado de Hitti y col., 2010).

1.5 Importancia del primer intrón y primer exón en los niveles de transcripción.

La presencia del primer intrón en el vector patentado descrito anteriormente, se justifica pues se ha reportado que algunos transgenes carentes de intrones son expresados en menores cantidades respecto a los que sí poseen (Brinster y col., 1988), además también se ha reportado que la remoción de intrones de genes endógenos disminuye su expresión considerablemente (Furger y col., 2002). Los mecanismos involucrados en cómo los intrones pueden contribuir a aumentar la transcripción es un área de estudio aún en

desarrollo. Según uno de los modelos propuestos, el aumento en la expresión se debe a la presencia de secuencias específicas dentro del intrón que permitirían la unión de factores de transcripción o complejos activadores de la transcripción (Parra y col., 2011). Esto se ve reflejado principalmente en el primer intrón, cuya composición está enriquecida con respecto al resto de los intrones en nucleótidos C, G, sus dinucleótidos (Figura 5), además de islas CpG, cajas TATA, CAAT y GC (Li y col., 2012).

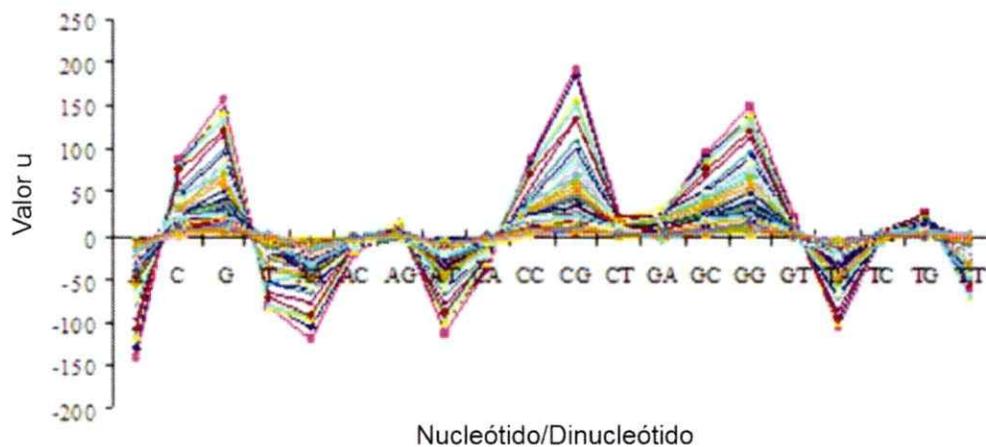


Figura 5. Enriquecimiento de nucleótidos y dinucleótidos en el primer intrón. Eligiendo un valor de significancia alfa igual a 0.05, valores de *u* superiores a 1.96 indican un enriquecimiento o sobrerrepresentación del nucleótido o dinucleótido en el primer intrón con respecto al resto de los intrones considerando el genoma humano (modificado de Li y col., 2012).

Un mecanismo complementario al anterior para explicar la potenciación de la transcripción ante la presencia de un intrón, se ha propuesto basándose en que el proceso de *splicing* en sí mismo favorecería la transcripción. Esto se logra, por ejemplo vía el complejo riboproteico U1 snRNP que participa en el *splicing*, pues éste se asocia con la

ARNPII, potenciando así la formación del complejo de pre-iniciación de la transcripción (Damgaard y col., 2008; Das y col., 2007). Además, otros componentes de la maquinaria de *splicing* han sido implicados en la estimulación de la elongación de la transcripción, favoreciendo finalmente la generación de una mayor cantidad de transcritos (Fong & Zhou, 2001; Lin y col., 2008).

Así como el primer intrón de cada gen puede tener una influencia sobre los niveles transcripción, también el primer exón parece tener una importancia en el proceso. Se ha demostrado que ciertas modificaciones en algunas histonas, asociadas a genes transcripcionalmente activos, como la tri-metilación de la lisina 4 de la histona 3 (H3K4me3) o la acetilación de la lisina 9 de la misma histona (H3K9ac) se ubican específicamente en los límites del primer exón-intrón. Estas modificaciones favorecen el reclutamiento de los factores generales de transcripción como TFIID a los promotores. Junto a lo anterior, se ha estudiado la relación entre el largo del primer exón con la abundancia de TFIID y ARNPII, encontrando que exones de hasta 250 pb de extensión tienen más del doble abundancia relativa de estos factores, respecto a exones con un largo entre 750 y 1000 pb (Figura 6) (Bieberstein y col., 2012).

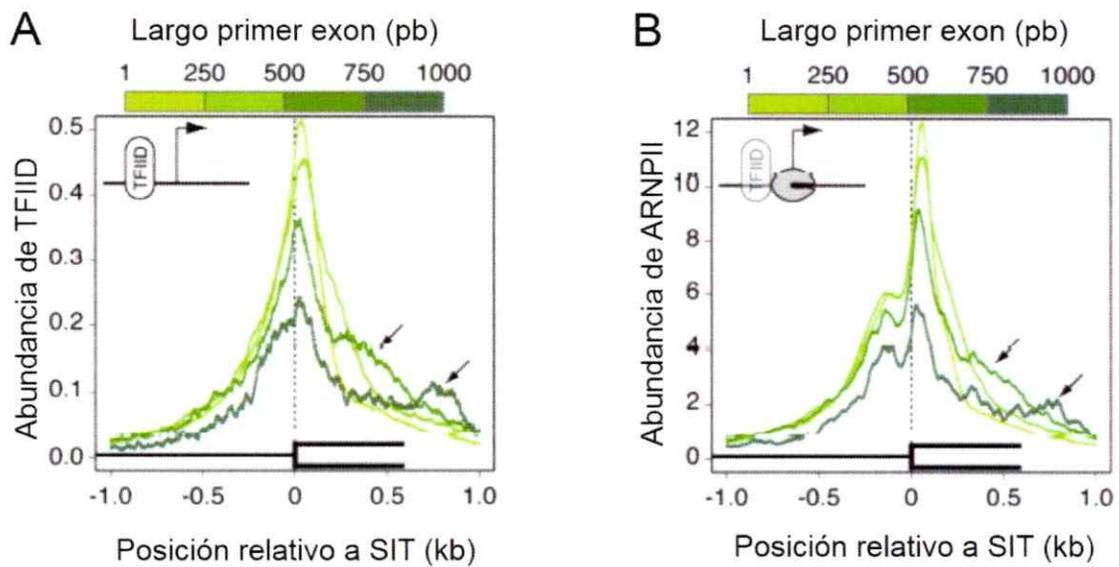


Figura 6. Largo del primer exón y presencia de TFIIID o ARNPII. Distribuciones promedio de TFIIID A) y ARNPII B). Utilizando los datos de CHIP-seq del proyecto ENCODE se calculó la abundancia de la maquinaria transcripcional relativo al sitio de inicio de tanscripción (SIT) en función del largo del primer exón (Modificado de Bieberstein y col., 2012)



Planteamiento del Problema

La producción de proteínas recombinantes utilizando vectores de expresión con promotores virales como CMV, se ve fuertemente disminuida luego de una serie de subcultivos debido a metilaciones de la región promotora y/o modificaciones de las histonas, surgiendo la necesidad de generar vectores de expresión con promotores modificados de tal modo que la expresión se mantenga estable.

Considerando los antecedentes expuestos, se postuló que los promotores de genes que codifican proteínas ribosomales podrían ser buenos candidatos para tomar como base e incorporar elementos de otros PCs. Por lo tanto, se buscó un criterio que permitiese decidir con cuál de los 80 genes existentes de este tipo trabajar. Frente a esto, un estudio comparativo de expresión a nivel transcripcional de estos genes utilizando tejidos humanos como retina, músculo, hígado, cerebro, útero y ovario, indicó que en ovario existe una elevada expresión del gen Rpl14 en comparación al resto (Bortoluzzi y col., 2001). Junto a lo anterior, teniendo presente que la línea celular utilizada como modelo de estudio (CHO-K1) deriva de tejido ovárico, y que además existe una alta homología de secuencia nucleotídica de dichos genes en mamíferos (Perry, 2005), se optó por utilizar como base el PC y una región que contiene los dos primeros exones, y el primer intrón del gen que codifica la proteína ribosomal L14.

HIPÓTESIS

Regiones del gen que codifica para la proteína ribosomal L14, con un promotor con iniciador de polipiridinas; TOP, junto a elementos DPE (*Downstream Promoter Element*) y MTE (*Motif Ten Element*), son funcionales en vectores de expresión de proteínas recombinantes en células CHO.

OBJETIVO GENERAL

Generar un vector de expresión con un promotor optimizado para la producción de anticuerpos recombinantes en células CHO-K1.

OBJETIVOS ESPECÍFICOS

1. Generar vectores de expresión que porten promotores *core* con los elementos MTE, DPE, caja TATA y TOP en distintas combinaciones, y luego incorporar la región EIE del gen Rpl14.
2. Evaluar la actividad transcripcional de los vectores generados mediante ensayos reporteros utilizando enzimas luciferasas, luego transfectar transientemente células CHO-K1.

2. MATERIALES Y MÉTODOS.

2.1. Materiales.

2.1.1 Cepa Bacteriana.

En este trabajo se utilizó la cepa comercial de *E. coli* DH5- α , se utilizó para la transformación y replicación de plasmidios. Se cultivó a 37°C con agitación constante durante toda la noche en medio completo LB líquido (Luria Bertani: 5 g/L extracto de levadura, 10 g/L triptona, 10 g/L NaCl, a pH 7.0). Cuando los cultivos se realizaron en placas, se adicionó agar al 1,5% al LB, y para aquellas cepas portadoras de plasmidios que confieren resistencia a ampicilina (Sigma-Aldrich, EE. UU.), se suplementó con dicho antibiótico a una concentración final de 100 μ g/ml. En el caso de la selección de transformantes con el vector pGEM-T Easy (Promega, EE. UU.), las placas se suplementaron además con el sustrato X-Gal a una concentración final de 30 μ g/ml, junto al inductor Isopropil β -D-1-tiogalactopiranosido (IPTG) (ambos de Sigma-Aldrich, EE. UU.) a una concentración final de 100 mM, de tal manera de seleccionar las colonias blancas. Esto debido a que la región de múltiple clonamiento se encuentra dentro del gen LacZ, el cual codifica la enzima β -galactosidasa por lo que se ve interrumpida la hidrolisis del X-Gal, evitando la coloración azul.

2.1.2 Células eucariontes.

El presente trabajó utilizó la línea celular derivada de ovario de hámster chino; CHO-K1, la cual siempre se cultivó, a menos que se indique lo contrario, en medio DMEM/F-12 (Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12) (Invitrogen,

EE. UU.), suplementado al 10% v/v de suero fetal bovino (SFB) (HyClone, EE. UU.), 1% v/v penicilina/estreptomicina (HyClone®, EE. UU.) y 1% v/v L-glutamina (HyClone®, EE. UU.), incubadas al 5% de CO₂ y a 37°C.

2.1.3 Plasmidios.

Las características generales de los vectores comerciales o generados previamente, tanto para subclonamiento como para expresión utilizados en el presente trabajo se encuentran resumidas en las siguientes tablas:

Tabla 1. Plasmidios comerciales^a utilizados en el presente seminario.

Plasmidios	Características principales	Origen
pGEM-T Easy	Vector de clonamiento. Confiere resistencia a ampicilina y porta LacZ como gen reportero.	Promega, EE. UU.
pGL 4.17	Vector de expresión. Confiere resistencia a ampicilina y Neomicina. Porta LUC2 como gen reportero, sin promotor.	Promega, EE. UU.
pGL CoreCMV	Generado a partir de la incorporación del promotor <i>core</i> de CMV en pGL 4.17	<i>Centro IBT^a.</i>
pGL CMV	Generado a partir de la incorporación del <i>enhancer</i> y promotor <i>core</i> de CMV en pGL 4.17	<i>Centro IBT^a.</i>
pGL 4.74	Vector de expresión. Porta hRluc como gen reportero bajo el promotor HSV-TK.	Promega, EE. UU.

^a Vectores generados por R. Zuñiga en el centro de InmunoBioTecnología (IBT) de la Facultad de Medicina, Universidad de Chile, previos al presente trabajo.

Tabla 2. Plasmidios generados en el presente seminario de título.

Plasmidios	Características principales (vector base: pGL 4.17)
pGL Core1	Elementos del promotor <i>core</i> : TATA y TOP
pGL Core2	Elemento de promotor <i>core</i> : TOP
pGL Core3	Elementos del promotor <i>core</i> : TATA y Inr
pGL Core4	Elemento de promotor <i>core</i> : Inr
pGL Core5	Elementos del promotor <i>core</i> : TATA, TOP, MTE y DPE
pGL Core6	Elementos del promotor <i>core</i> : TOP, MTE y DPE
pGL Core7	Elementos del promotor <i>core</i> : TATA, Inr, MTE y DPE
pGL Core8	Elementos del promotor <i>core</i> : Inr, MTE y DPE
pGL Core1-EIE	pGL Core1 + exon1-intron1-exón2
pGL Core2-EIE	pGL Core2 + exon1-intron1-exón2
pGL Core3-EIE	pGL Core3 + exon1-intron1-exón2
pGL Core4-EIE	pGL Core4 + exon1-intron1-exón2
pGL Core5-EIE	pGL Core5 + exon1-intron1-exón2
pGL Core6-EIE	pGL Core6 + exon1-intron1-exón2
pGL Core7-EIE	pGL Core7 + exon1-intron1-exón2
pGL Core8-EIE	pGL Core8 + exon1-intron1-exón2

2.1.4 Oligonucleótidos.

Todos los oligonucleótidos, usados tanto para la generación de los diferentes promotores *core* como partidores de PCR, se obtuvieron de IDT (Coralville, EE. UU.) y se encuentran detallados en la Tabla 3.

Tabla 3. Oligonucleótidos utilizados y/o diseñados en este trabajo.

Nombre	Secuencia (5' - 3')	Propósito
P1 TA-TC fw	ctagcgtataaaaggaccaggtcccatctgagattcttttctgtattcta	P. Core TATA TOP
P1 TA-TC rv	Cagaaaaa gaaatctcagatgggaaccctggctctttttatagcg	P. Core TATA TOP
P2 TA-TC fw	ggcttttagctgagtgaccctgcacccagtgcaaggctcgcactaaagccctagaata	P. Core TATA TOP
P2 TA-TC rv	gattcggcctgtctctcaccaggttcccatctgagattcttttctgtattcta	P. Core TATA TOP
P1 NTA-TC fw	ctagccggcccgccaggttcccatctgagattcttttctgtattcta	P. Core TOP
P1 NTA-TC rv	cagaaaaa gaaatctcagatgggaacctggctcaccggccggcg	P. Core TOP
P2 NTA-TC fw	ggcttttagctgagtgaccctgcacccaggttcccatctgagattcttttctgtattcta	P. Core TOP
P2 NTA-TC rv	gattcggcctgtctctcaccaggttcccatctgagattcttttctgtattcta	P. Core TATA INR
P1 TA-IN fw	ctagcgtataaaaggaccaggttcccatctgagattcttttctgtattcta	P. Core TATA INR
P1 TA-IN rv	Cagaagactgaatctcagatgggaacctggctcttttatagcg	P. Core TATA INR
P2 TA-IN fw	ggcttttagctgagtgaccctgcacccagtgcaaggctcgcactaaagccctagaata	P. Core TATA INR
P2 TA-IN rv	gattcggcctgtctctcaccaggttcccatctgagattcttttctgtattcta	P. Core TATA INR
P1 NTA-IN fw	ctagccggcccgccaggttcccatctgagattcttttctgtattcta	P. Core INR
P1 NTA-IN rv	cagaagactgaatctcagatgggaacctggctcaccggccggcg	P. Core INR
P2 NTA-IN fw	ggcttttagctgagtgaccctgcacccaggttcccatctgagattcttttctgtattcta	P. Core INR
P2 NTA-IN rv	gattcggcctgtctctcaccaggttcccatctgagattcttttctgtattcta	P. Core INR
P1 TA-TC-MD fw	ctagcgtataaaaggaccaggttcccatctgagattcttttctgtattcta	P. Core TATA TOP MTE-DPE
P1 TA-TC-MD rv	Cagaaaaa gaaatctcagatgggaacctggctcttttatagcg	P. Core TATA TOP MTE-DPE
P2 TA-TC-MD fw	tcgaaccgaaaggacgtcttgacatcgggtgagagacagggcca	P. Core TATA TOP MTE-DPE
P2 TA-TC-MD rv	gattcggcctgtctctcaccaggttcccatctgagattcttttctgtattcta	P. Core TATA TOP MTE-DPE
P1 NTA-TC-MD fw	ctagccggcccgccaggttcccatctgagattcttttctgtattcta	P. Core TOP MTE-DPE
P1 NTA-TC-MD rv	cagaaaaa gaaatctcagatgggaacctggctcaccggccggcg	P. Core TOP MTE-DPE
P2 NTA-TC-MD fw	tcgagccgaaaggacgtcttgacatcgggtgagagacagggcca	P. Core TOP MTE-DPE
P2 NTA-TC-MD rv	gattcggcctgtctctcaccaggttcccatctgagattcttttctgtattcta	P. Core TOP MTE-DPE
P1 TA-IN-MD fw	ctagcgtataaaaggaccaggttcccatctgagattcttttctgtattcta	P. Core TATA INR MTE-DPE
P1 TA-IN-MD rv	Cagaagactgaatctcagatgggaacctggctcttttatagcg	P. Core TATA INR MTE-DPE

(continuación tabla 3)

Nombre	Secuencia (5' - 3')	Propósito
P2 TA-IN-MD fw	tcgagccgaacggacgctctgacactcgggagggagagagagagggcca	P. Core TATA INR MTE-DPE
P2 TA-IN-MD rv	gactcggcctgtctcctcaccggagtgcaagactccgttcgggtcggatagaata	P. Core TATA INR MTE-DPE
P1 NTA-IN-MD fw	ctagccggcccggtgaccaggttcccatctgagattcagttcttctgtatteta	P. Core INR MTE-DPE
P1 NTA-IN-MD rv	cagaagactgaatctcagatgggaacctgggtcacgggcccggg	P. Core INR MTE-DPE
P2 NTA-IN-MD fw	tcgaaccgaacggagctctgcaactcgggtaggagagagagagggcca	P. Core INR MTE-DPE
P2 NTA-IN-MD rv	gactcggcctgtctcctcaccggagtgcaagactccgttcgggtcggatagaata	P. Core INR MTE-DPE
Int2 Fw	Ttttttagatctcagcttactgacttcatc	PCR Exon1-intron1-exon2
Int2 Rv	Tttttaagcttggtaaatgactgggcga	PCR Exon1-intron1-exon2
T7 Promoter fw	Attatgctgagtgatatccc	PCR en pGEM-T Easy
SP6 Promoter rv	Atttaggtgacactatagaa	PCR en pGEM-T Easy
LucRev01 fw	Cgtcggtaaaggcgaiggtg	PCR en pGL 4.17
RVprimer3 rv	Ctagcaaaataggctgtcc	PCR en pGL 4-17

2.1.5 Enzimas.

T4 ADN ligasa, tripsina (New England Biolabs, EE. UU.).

PNK, *NheI*, *BglII*, *HindIII*, Fosfasa alcalina, (Fermentas, EE. UU.).

Taq ADN polimerasa (Invitrogen, EE. UU.).

2.1.6 Marcadores de ADN.

Marcador de ADN cuantificado, escala de 100 pb (Winkler, Chile).

Marcador de ADN Lambda/*HindIII* (Winkler, Chile).

2.1.7 Otros reactivos y sistemas comerciales.

Agar (Becton Dickinson, EE. UU.).

Agarosa (Winkler, Chile).

Albúmina de Suero Bovina (BSA) (Sigma-Aldrich, EE. UU.).

Ampicilina, (Sigma-Aldrich, EE. UU.).

Bromuro de etidio (Sigma-Aldrich, EE. UU.).

dNTPs (Omega Bio-tek, EE. UU.).

Dual-Luciferase Reporter Assay System, (Promega, EE. UU.).

Etanol, Glicerol, IPTG, X-Gal (Merck, Alemania).

Membrana de nitrocelulosa (Millipore).

Penicilina/estreptomicina, L-glutamina, Suero Fetal Bovino (Invitrogen, EE. UU.).

PureYield™ Plasmid Miniprep System (Promega, EE. UU.).

Todas las sales, ácidos, bases: Winkler, Chile o Merck, Alemania.

Wizard® SV Gel and PCR Clean-Up System, (Promega, EE. UU.).

pGEM®-T and pGEM®-T Easy Vector Systems (Promega, EE. UU.)

2.1.8 Soluciones y Medios de cultivo.

PBS: NaCl a 250 mM, KCl a 20 mM, Na₂HPO₄ a 48 mM, KH₂PO₄ a 1.5 mM. Se autoclavó antes de uso.

DMEM/F12 (Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12): cada sobre de medio DMEM/F12 se disolvió en 1 L de agua bidestilada estéril, se le agregó 1.42 g/L de NaHCO₃. Se ajustó a pH 7.0-7.2, se filtró y autoclavó.

LB (Caldo Luria Bertani): 5 g/L de Extracto de levadura, 10 g/L de Triptona, 10 g/L de NaCl. Se ajustó a pH 7.0 y autoclavó.

Placas LB Agar: 5 g/L de Extracto de levadura, 10 g/L de Triptona, 10 g/L de NaCl, 15 g/L de agar y se autoclavó.

LB-Glicerol: LB líquido + Glicerol al 15% (v/v). Se autoclavó.

Amortiguador TBE 5x: Tris base a 89 mM, Ácido bórico a 89 mM, EDTA a 2 mM, se ajustó a pH 8.3 - 8.7.

Amortiguador TE: Tris-HCL a 10 mM, EDTA a 1 mM. Se ajustó a pH 8.0 y autoclavó.

Amortiguador de carga de DNA 6x: 10mM de Tris-HCl (pH 7.6), 0.03% de azul bromofenol, 0.03% de xilencianol, 60% de glicerol, 60mM de EDTA.

Amortiguador Carbonato: Na_2CO_3 a 0.015 M, NaHCO_3 a 0.0035 M, pH 9.5.

PEI (Polietilenimina): de peso molecular 25 kDa 1,0 $\mu\text{g}/\text{mL}$ en su forma lineal (Sigma-Aldrich®, EE. UU.).

2.2 Métodos.

2.2.1 Generación de vectores “pGL *core* 1-8”.

Tomando como base la secuencia de ADN del PC del gen Rpl14 (Gene ID: 100774219. GenBank, NCBI), se agregaron diferentes combinaciones de los motivos caja TATA, TOP, INR, MTE y DPE, y se construyeron 8 PCs. Para ello, se diseñaron secuencias de oligonucleótidos de tal modo que cada uno de los ocho promotores se formase mediante dos secuencias de ADN de doble hebra (ADNdh) de alrededor de 50 pb cada una, donde el extremo 3' de un ADNdh era cohesivo y complementario al extremo 5' del otro ADNdh. A su vez, cada ADNdh se formó por dos hebras simples de oligonucleótidos complementarias entre sí. Finalmente, cada promotor formado por las dos secuencias de

ADNdh, tenía extremos 5' y 3' compatibles para ser ligado directamente en un vector digerido por las enzimas *NheI* y *BglII*, respectivamente. Todo el proceso se esquematiza en la Figura 7 y la secuencia de los 8 PCs con sus respectivos motivos se observa en la figura 8. Brevemente, cada oligonucleótido fue independientemente fosforilado por la enzima PNK a 37°C durante una hora. Luego, para formar las secuencias de ADNdh se mezcló cada pareja de oligonucleótidos complementarios entre sí en un tubo *eppendorf* diferente, se calentó en un "termoblock" hasta 100°C por cinco minutos, y posteriormente se dejó enfriar a temperatura ambiente. Las dos secuencias de ADNdh junto al vector pGL 4.17 (previamente digerido con las enzimas *NheI* y *BglII* durante 3 horas a 37°C, separado mediante electroforesis en gel de agarosa al 1% y purificado utilizando el kit *Wizard® SV Gel and PCR Clean-Up System*) se utilizaron para las reacciones de ligación, es decir se ligaron 3 moléculas de ADN a la vez, utilizando la enzima T4 ADN ligasa a 16°C durante toda la noche. La reacción de ligación se componía de 5 µL de oligonucleótidos fosforilados a 100 µM, 2 µL de amortiguador 10X de T4 DNA ligasa, 1 µL de ligasa, 100 ng de plasmidio digerido y se completó con agua hasta los 20 µL.

Posteriormente las ligaciones se dializaron contra agua bidestilada durante 15 min, utilizando membranas de nitrocelulosa "Millipore Type-VS" para luego llevar a cabo la electrotransformación de *E. coli* DH5α y posterior chequeo mediante PCR colonia.

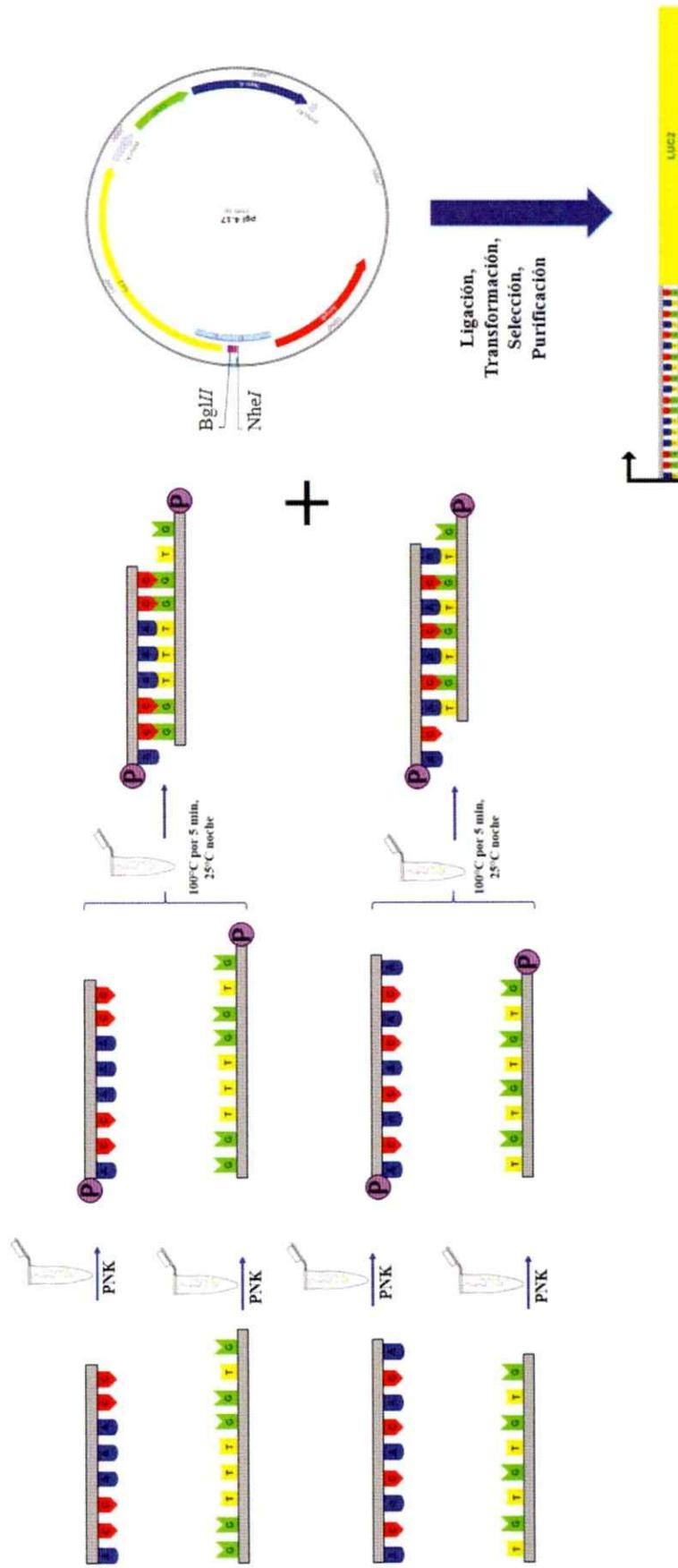


Figura 7. Esquema generación vectores "pgL Core".

Para cada uno de los ocho promotores *core*, se realizó el mismo procedimiento esquematizado arriba, el cual cuenta con una fase de fosforilación del extremo 5' del oligonucleótido, luego una etapa de desnaturalación seguido de una hibridación. Posteriormente los fragmentos de ADN de doble se ligaron en el vector pGL 4.17, generándose una construcción con el promotor (hecho en base a dos fragmentos) río arriba del gen LUC2.

Promotor Core1, Elementos: TATA y TOP
5' -GCTAGCCGTATAAAAAGGACCAGGTTCCCATCTGAGATTTCTTTTCTGTATTCTAGGCTTTAGTCGATGGACCTTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGCGATATTTTCTGGTCCAAGGGTAGACTCTAAAGAAAAGACATAAAGATCCGAAATCAGTACCTGGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core2, Elemento: TOP
5' -GCTAGCCCGGGCCGTGACCAGGTTCCCATCTGAGATTTCTTTTCTGTATTCTAGGCTTTAGTCGATGGACCTTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGGGCCGGCACTGGTCCAAGGGTAGACTCTAAAGAAAAGACATAAAGATCCGAAATCAGTACCTGGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core3, Elementos: TATA y INR
5' -GCTAGCCGTATAAAAAGGACCAGGTTCCCATCTGAGATTCAGTCTTCTGTATTCTAGGCTTTAGTCGATGGACCTTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGCGATATTTTCTGGTCCAAGGGTAGACTCTAAGTCAGAAAAGACATAAAGATCCGAAATCAGTACCTGGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core4, Elemento: INR
5' -GCTAGCCCGGGCCGTGACCAGGTTCCCATCTGAGATTCAGTCTTCTGTATTCTAGGCTTTAGTCGATGGACCTTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGGGCCGGCACTGGTCCAAGGGTAGACTCTAAGTCAGAAAAGACATAAAGATCCGAAATCAGTACCTGGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core5, Elementos: TATA, TOP, MTE y DPE
5' -GCTAGCCGTATAAAAAGGACCAGGTTCCCATCTGAGATTTCTTTTCTGTATTCTCGAACCCGAACGGACGTCCTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGCGATATTTTCTGGTCCAAGGGTAGACTCTAAAGAAAAGACATAAAGATAGCTTGGCTTGCCTGCAGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core6, Elementos: TOP, MTE y DPE
5' -GCTAGCCCGGGCCGTGACCAGGTTCCCATCTGAGATTTCTTTTCTGTATTCTCGAACCCGAACGGACGTCCTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGGGCCGGCACTGGTCCAAGGGTAGACTCTAAAGAAAAGACATAAAGATAGCTTGGCTTGCCTGCAGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core7, Elementos: TATA, INR, MTE y DPE
5' -GCTAGCCGTATAAAAAGGACCAGGTTCCCATCTGAGATTCAGTCTTCTGTATTCTCGAACCCGAACGGACGTCCTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGCGATATTTTCTGGTCCAAGGGTAGACTCTAAGTCAGAAAAGACATAAAGATAGCTTGGCTTGCCTGCAGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Promotor Core8, Elementos: INR, MTE y DPE
5' -GCTAGCCCGGGCCGTGACCAGGTTCCCATCTGAGATTCAGTCTTCTGTATTCTCGAACCCGAACGGACGTCCTGCACCTCGGGTGAGGAGACAGGCCAGATCT-3'
3' -CGATCGGGCCGGCACTGGTCCAAGGGTAGACTCTAAGTCAGAAAAGACATAAAGATAGCTTGGCTTGCCTGCAGAACGTGAGCCCACTCCTCTGTCCGGTCTAGA-5'

Figura 8. Secuencia de ADN de los 8 promotores core. Los motivos TATA, TOP, INR y MTE-DPE están destacados con fondo amarillo, gris, verde y calipso, respectivamente. En los extremos 5' y 3' de cada secuencia se encuentran en fuente de color rojo los sitios de restricción reconocidos por las enzimas NheI y BglII, respectivamente.

2.2.1.2 Electrotransformación de *E. coli* DH5 α .

Con el propósito de transformar *E. coli* DH5 α con cada una de las construcciones, se fijó el electroporador en 2,5 KVolts y se generó una descarga durante 4.5 mseg, agregando inmediatamente 1 mL de LB en condiciones de esterilidad al lado de mechero. El contenido de la cubeta se transfirió a un nuevo tubo *eppendorf* y se incubó a 37°C durante una hora. Luego diferentes diluciones se sembraron sobre placas LB-agar más ampicilina. Una vez secas las placas, éstas se incubaron de forma invertida durante toda la noche a 37°C.

2.2.1.3 PCR Colonia.

La búsqueda de transformantes que incorporaron los promotores-en el vector pGL 4.17 se realizó mediante PCR colonia, utilizando los partidores LucRev01fw sentido y RVprimer3rv antisentido, los cuales amplifican alrededor de 300 pb a partir del vector pGL 4.17 sin inserto, y alrededor de 400 pb si el clonamiento de los promotores fuese exitoso. Para cada reacción de amplificación de ADN con la enzima *Taq* ADN polimerasa se mezclaron los siguientes reactivos: tampón PCR de enzima 10X (2.5 μ L), MgCl₂ 50 Mm (0.75 μ L), mezcla dNTPs 10 mM (0.5 μ L), partidor sentido 10 μ M (0.5 μ L), partidor antisentido 10 μ M (0.5 μ L), enzima ADN polimerasa 5U/L (0.5 μ L), agua c.s.p. 25 μ L, y con una punta estéril se picó una colonia y se introdujo en la mezcla de PCR descrita. Se utilizó el programa de PCR consistente en 1 ciclo de 94°C por 5 min,

35 ciclos de 94°C por 30 s, 55°C por 30s, 72°C por 60 s; seguido por 1 ciclo de 72°C por 10 min.

2.2.1.4 Extracción ADN plasmidial.

Una vez confirmado por PCR colonia la incorporación del clonamiento del inserto de interés, se procedió a cultivar las colonias positivas en LB líquido más ampicilina durante toda la noche, con el propósito de extraer ADN plasmidial mediante el *kit* comercial *PureYield™ Plasmid Miniprep System*, de acuerdo a las instrucciones del fabricante. Usando el ADN plasmidial, nuevamente se realizó PCR para detectar la presencia del promotor en el vector pGL4.17, y luego se confirmó mediante secuenciación automática (OMICS Solutions, Pontificia Universidad Católica de Chile).

2.2.2.1 Generación de vector “pGEMT Easy-EIE”.

A partir de la secuencia del gen Rpl14 disponible en la base de datos del NCBI (Gene ID: 100774219), se diseñaron los partidores Int2Fw sentido y Int2Rv antisentido, que amplifican la región comprendida desde la posición +69 hasta +1175, exón 1, intrón 1 y exón 2 del gen Rpl14 (de ahora en adelante; “región EIE”), las con respecto al sitio de inicio de la transcripción, e incorporan sitios de restricción para enzimas *BglI* y *HindIII*, respectivamente. A partir de ADN genómico extraído de células CHO-K1, se amplificó la región EIE y el tamaño esperado del producto de PCR se chequeó mediante electroforesis en gel de agarosa al 1%, y posteriormente se purificó utilizando el *kit Wizard® SV Gel and PCR Clean-Up System*. Una vez confirmado el tamaño del

producto de PCR, se procedió a subclonarlo en el vector pGEMT Easy, de acuerdo a las instrucciones del fabricante. Con el producto de esta ligación, se electrotransformó *E. coli* DH5 α (utilizando el protocolo anteriormente descrito) y se sembraron en placas LB/Agar/Ampicilina/X-Gal/IPTG. Esta vez las colonias positivas se seleccionaron en función del color, pues la incorporación del inserto interrumpe la región codificante del péptido- α de la enzima β -galactosidasa, debido a lo cual no es posible hidrolizar el X-Gal y no se produce la coloración azul. Luego, para confirmar la incorporación del producto de PCR en el vector, se procedió a realizar PCR colonia utilizando los mismos partidores empleados en la amplificación desde ADN_g; Int2Fw sentido y Int2Rv antisentido. Las colonias que resultaron ser positivas fueron posteriormente confirmadas mediante secuenciación automática utilizando los partidores T7promoter sentido y SP6 antisentido (OMICS Solutions, Pontificia Universidad Católica de Chile).

2.2.2.2 Clonamiento de la región EIE en los vectores “pGL core 1-8”.

Con el propósito de incorporar la región EIE río abajo de los 8 promotores antes construidos, se procedió a liberar el inserto a partir del vector pGEM-T Easy que lo contenía. Para esto se llevó a cabo una digestión doble utilizando las enzimas BglIII y HindIII durante 3 horas a 37°C. Posteriormente, se separó mediante electroforesis en gel de agarosa al 1% y se purificó utilizando el *kit Wizard® SV Gel and PCR Clean-Up System*. Paralelamente, los 8 vectores con los distintos promotores se digirieron con las mismas enzimas de restricción y purificados de igual modo. Luego se procedió a ligar cada vector digerido con el inserto durante toda la noche, como lo descrito

anteriormente. Posteriormente la ligación se dializó, y se electrotransformó *E. coli* DH5 α . La presencia de la región EIE se chequeó mediante PCR colonia utilizando los partidores utilizados anteriormente para amplificar dicha región a partir de ADNg; Int2Fw sentido y Int2Rv antisentido. La presencia del inserto se verificó mediante secuenciación automática utilizando los partidores LucRev01fw sentido y RVprimer3rv antisentido. Todas las colonias cuya secuencia fue confirmada por secuenciación fueron seleccionadas y se almacenaron a -80°C en medio LB líquido-Glicerol 15% v/v.

2.2.3 Optimización cantidad PEI/ADN en transfección de células CHO-K1.

Con el propósito de determinar cuáles era las cantidades de ADN y del polímero catiónico usado como transportador (en este caso polietilenimina (PEI)), con las cuales se debía transfectar las células CHO-K1 para obtener mayores unidades relativas de luminiscencia a las 24 horas, se procedió a efectuar una optimización de estas variables. En este caso el ADN utilizado fue el vector pGL CMV, el cual contiene el *enhancer* y promotor CMV comandando la expresión del gen reportero que codifica para la enzima luciferasa de luciérnaga, y para normalizar se usó el vector pGL 4.74 cuyo gen reportero codifica una luciferasa diferente dirigido por el promotor HSV-TK.

Previo a la optimización, un ensayo cualitativo se efectuó para evaluar de modo general la supervivencia del cultivo celular frente a cantidades crecientes de ADN y PEI. En este ensayo se encontró que la cantidad de ADN no resultaba evidentemente relevante en la sobrevivencia del cultivo, a diferencia de lo observado para el PEI, puesto que cantidades

superiores a 12 μg del compuesto resultaba en una sobrevida fuertemente reducida (datos no mostrados) acorde a lo reportado en la literatura (Floreza y col., 2002; Choi y col., 2010).

Para esta optimización se utilizó un diseño central compuesto. Este diseño es uno de los más utilizados para la optimización de variables cuantitativas mediante superficies de respuesta. Dicho método utiliza dos variables cuantitativas, cada una con 3 niveles (-1, 0, 1), y se evalúan todas las combinaciones resultantes formadas por un nivel de una variable con un nivel de la otra. Los niveles y cantidades utilizadas en la optimización se muestran en la siguiente tabla:

Tabla 4. Cantidades de ADN y PEI (μg), clasificados en 3 niveles.

	Nivel		
	-1	0	1
ADN (μg)	0.5	1.75	3
PEI (μg)	4	8	12

Las transfecciones utilizando las diferentes cantidades de ADN y PEI, se efectuaron siguiendo las mismas condiciones de incubación descritas en la siguiente sección.

2.2.4 Transfección células CHO-K1.

Se cultivaron células CHO-K1 hasta una confluencia del 90%, en una botella *T-flask* T-25 con 10 mL de medio de cultivo (sección Materiales), en una incubadora con CO_2 al 5% y a 37°C. Luego, se extrajo el medio de cultivo y se lavó suavemente 2 veces con tampón PBS, seguido a lo cual se agregó 1 mL de tripsina 1x y se incubó durante 3

minutos a 37°C, tiempo después del cual se agregó medio de cultivo como bloqueador. Se transfirió el contenido de la botella a un tubo *falcon* de 15 mL, y se colectaron las células por centrifugación durante 5 min a 1200 rpm, se contaron las células en una cámara de Neubauer y se sembraron 200.000 células en cada uno de los 24 pocillos de la placa utilizada, agregando medio de cultivo hasta un volumen final de 500 uL. Se incubó durante 24 horas. Posteriormente se extrajo el medio de cultivo, y se lavó 2 veces con tampón PBS. El PEI se incubó junto al vector a transfectar durante 15 min, y luego fue transferido sobre el cultivo celular recién lavado con PBS, y se agregó medio de cultivo sin SFB hasta un volumen final de 500 mL. Se cultivó durante 3 horas, y luego se procedió a cambiar el medio de cultivo por uno nuevo suplementado con SFB al 10% v/v. A las 24 horas se midió la actividad luciferasa con el *kit Dual-Luciferase Reporter Assay System*, siguiendo las instrucciones del fabricante.

2.2.5 Co-Transfección vectores portando los genes reporteros Luc.

Todos los vectores construidos en el presente seminario de título utilizaron como base el vector de expresión pGL 4.17, el cual contiene el gen que codifica para una luciferasa de luciérnaga (*Photinus pyralis*). Con el propósito de normalizar la luminiscencia, cada una de estas construcciones se cotransfectó con el vector pGL 4.74, el cual contiene un gen codificante de una luciferasa distinta; uno de *Renilla reniformis* o “pensamiento de mar”, dirigido por el promotor viral HSV-TK. Cabe mencionar que cada una de estas dos luciferasas tiene un sustrato distinto por lo que la reacción enzimática de una para producir luminiscencia no interfiere con la de la otra. Ante lo recién dicho, es necesario

aclarar que cada una de las luminiscencias de los diferentes vectores construidos se reporta relativo a la luminiscencia alcanzada por la luciferasa de *Renilla reniformis*. Las condiciones a las cuales es sometido el cultivo de células CHO-K1 para realizar las co-transfecciones, son las mismas a las descritas en la sección 2.2.4, exceptuando que la cantidad total de ADN entre los dos vectores (vector con promotor construido en este seminario y vector con promotor HSV-TK, dirigiendo la transcripción de luciferasa de luciérnaga y *Renilla reniformis*, respectivamente) alcanza los 500 ng, los cuales se reparten en dos cantidades equimolares de cada vector, y la cantidad de PEI fue de 12 µg, masas determinadas de acuerdo al ensayo de optimización.

2.2.6 Medición de luminiscencia.

La luminiscencia de las células co-transfectadas fue medida utilizando el *kit* comercial *Dual-Luciferase Reporter Assay System* (que contiene los compuestos *LARII* y *Stop&Glo* y tampón de lisis mencionados en esta sección) (Promega, EE. UU.), de acuerdo a las instrucciones del fabricante. Este kit contiene Brevemente, 24 horas después de la co-transfección, se removió el medio de cultivo de los pocillos de la placa, y las células fueron lavadas 2 veces con tampón PBS, procurando extraer todo el líquido. Posteriormente se agregó 100 uL de tampón de lisis, y se incubó durante 15 min en agitación en balanza a temperatura ambiente. De este lisado, se transfirió 20 uL a una placa blanca de medición de luminiscencia, se agregó *LARII* (sustrato de luciferasa de luciérnaga), se midió la luminiscencia en luminómetro e inmediatamente después se agregó *Stop&Glo* (inhibidor de luciferasa de luciérnaga y sustrato de luciferasa de

Renilla reniformis) y se volvió a medir en el luminómetro. La medida final de luminiscencia reportada y graficada es la luminiscencia de luciferasa de luciernaga dividida por la luminiscencia de luciferasa de *Renilla reniformis*, (Luc/Ren).

2.2.7 Cálculo de la Concentración de ADN por espectrofotometría.

La concentración de ADN se obtuvo utilizando un espectrofotómetro midiendo a una longitud de onda de 260 nm. Se calibró el espectrofotómetro para 0% de Absorbancia con agua ultrapura estéril y se midió la concentración de la muestra. La pureza de la solución de ADN se determinó obteniendo la razón de Absorbancia 260nm/280nm. Para una solución pura de DNA de doble hebra esta razón debe estar entre 1,7 y 1,9.

2.2.8. Electroforesis en Gel de Agarosa.

Para un gel de agarosa al 1% p/v se disolvió 1 g de agarosa en 100 mL de tampón TBE 1x (preparado según lo descrito en sección 2.1.8), se agregó el bromuro de etidio al 0,003% v/v e inmediatamente se depositó sobre la cámara de electroforesis. Posteriormente se cargó el volumen de las muestras de DNA escogido, disueltas en buffer de carga de DNA 1x. La electroforesis se realizó a un voltaje constante de 90 V por aproximadamente 1,0-1,5 h.

2.2.9. Reacción de ligación.

La reacción de ligación se realizó en un volumen final de 10 uL utilizando 1,5 U de enzima T4 ADN ligasa, 100 ng de vector digerido, DNA inserto en una relación 3 a 1 e

incubando por 14-16 h a 4°C en el caso del vector pGEM-T Easy, según instrucciones del fabricante, y a 16 °C en el caso de los vectores pGL.

2.2.10. Extracción de DNA plasmidial de *E. coli*.

Para la extracción de DNA plasmidial desde cultivos de *E. coli*, las bacterias se cultivaron en medio LB líquido suplementado con ampicilina y la extracción de ADN plasmidial se realizó mediante la utilización del *kit* comercial *PureYield™ Plasmid Miniprep System* (Promega, EE. UU.), según las instrucciones del fabricante.

2.2.11. Análisis estadístico

Los resultados mostrados corresponden al promedio \pm desviación estándar de al menos tres experimentos independientes para determinar la actividad luciferasa. El estadígrafo utilizado para comparar múltiples grupos fue ANOVA de una vía, seguido de la prueba de Newman-Keuls para las comparaciones a posteriori. Todos los datos se analizaron con el *software GraphPad Prism*, versión 5.0 (EE. UU.), excepto el análisis de optimización que se realizó en el *software “JMP, Academic version”* (EE.UU). Los valores de $p < 0.05$ se consideraron estadísticamente significativos.

2.3 Recursos Materiales y financiamiento.

El presente seminario de título fue financiado dentro del marco del Proyecto FONDEF D09I1190: “Desarrollo de un sistema de expresión optimizado para la producción de anticuerpos terapéuticos”.

3 RESULTADOS

3.1 Objetivo Especifico 1: Generar vectores de expresión que porten promotores core con los elementos MTE, DPE, caja TATA y TOP en distintas combinaciones, y luego incorporar la región EIE del gen Rpl14.

Los oligonucleótidos que mediante complementariedad y sobreposición conformaban los 8 promotores con diferentes elementos, se clonaron en el vector pGL 4.17 como se detalla en Materiales y Métodos (sección 2.2.1). Dado que cada uno de los promotores se formaba por dos fragmentos de ADN de doble hebra de alrededor de 50 pb, el clonamiento de cada promotor en el vector pGL 4.17 implicaba una reacción en donde se ligasen tres moléculas de ADN, lo cual no fue fácil de lograr, y se probaron diferentes proporciones fragmentos:vector, puesto que la razón 3:1 recomendada por el fabricante no dio resultados positivos en nuestro caso. Una vez logrado el clonamiento del primer promotor, el protocolo se repitió para los 7 restantes y se obtuvieron los transformantes en un menor tiempo que el empleado para el primero. Luego, se analizó las colonias mediante PCR, y luego a partir de las colonias positivas se purificó el ADN plasmidial con el fin de repetir el PCR usándolo como ADN molde. En la Figura 9, se muestra el resultado de la amplificación a partir del PCR usando ADN plasmidial, donde se observa que en ausencia de inserto se debiese obtener un producto de amplificación de alrededor de 300 pb (283 pb exactamente), y en presencia de cada uno de los promotores clonados el tamaño debiese aumentar a alrededor de 400pb (387pb, exactamente). Como control positivo de PCR, se utilizó el vector pGL 4.17 sin promotor, y como control negativo no

se agregó ADN molde. Posterior a la confirmación por PCR, se procedió a la verificación de la secuencia clonada mediante secuenciación automática, para lo cual se enviaron dos clones por cada promotor distinto. En la figura 10 se muestran los mejores alineamientos de las secuencias teóricas y experimentales. Los promotores core 2, 3, 4, 5, 6 evidenciaron discordancias en 2, 2, 1, 1, 1 nucleótidos con respecto a la secuencia teórica (de 104 pb de largo), respectivamente, pero ninguno de los motivos relevantes (TATA, TOP, INR, MTE-DPE y sitios de restricción) tuvo mutaciones.

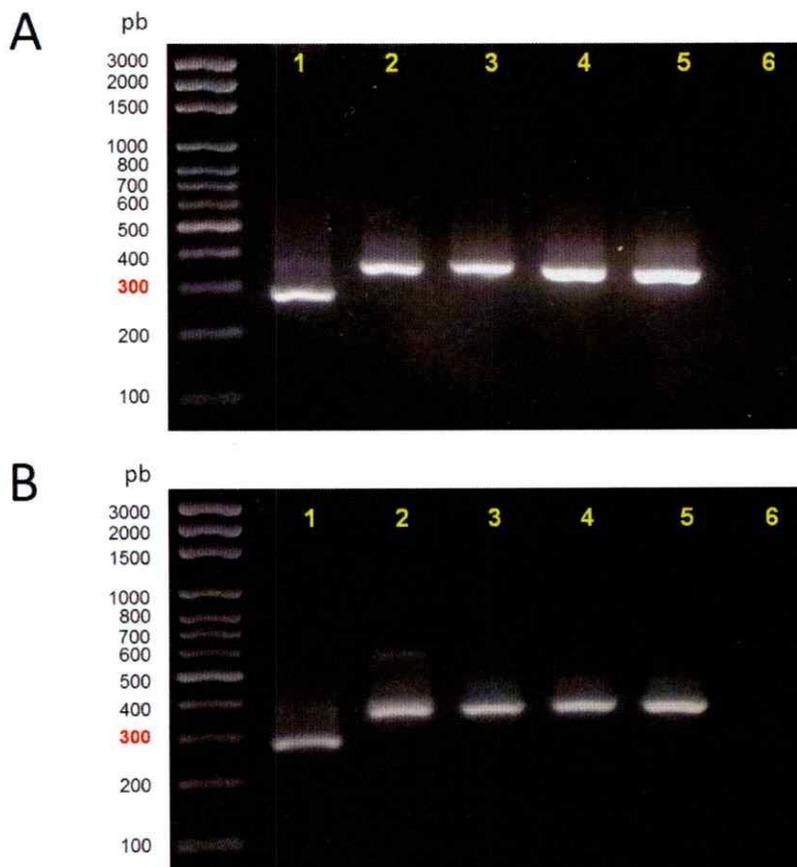


Figura 9. PCR usando como ADN molde los vectores "pGL Core1-8" purificados.

Panel A: Carril 1; pGL4.17 sin promotor clonado (control positivo de PCR), carril 2 a 5; Promotores con los elementos TATA/TOP, TOP, TATA/INR, INR, respectivamente, clonados en el vector pGL 4.17 y carril 6; control negativo.

Panel B: Carril 1; control positivo de PCR, pGL 4.17 sin promotor, carril 2 a 5; Promotores con los elementos TATA/TOP/MTE-DPE, TOP/MTE-DPE, TATA/INR/MTE-DPE, INR/MTE-DPE respectivamente, clonados en el vector pGL 4.17 y carril 6; control negativo.

Promotor Core1

		Nhe I	TATA			TOP						
Experimental	GCTAGCGCTATAAAAGGACCAGGTTAACATCTGAGATTCTTTTTCTGATTCTTAGGCTTTAGTCGATGGACCTTGCACT	60	70	80	90	100	110	120	130			
											
Teórica	GCTAGCGCTATAAAAGGACCAGGTTCCCATCTGAGATTCTTTTTCTGATTCTTAGGCTTTAGTCGATGGACCTTGCACT	60	70	80	90	100	110	120	130			
											
		10	20	30	40	50	60	70	80			
Experimental	CGGGTGAGGAGACAGGCCAGATCT	140	150									
											
Teórica	CGGGTGAGGAGACAGGCCAGATCT	140	150									
											
		90	100									

Bgl II

Promotor Core2

		Nhe I	TATA^{mut}			TOP						
Experimental	GCTAGCGCGGGCCCGTACCAGGATCCCATCTGAGATTCTTTTTCTGATTCTTAGGCTTTAGTCGATGGACCTTGCACT	60	70	80	90	100	110	120	130			
											
Teórica	GCTAGCGCGGGCCCGTACCAGGTTCCCATCTGAGATTCTTTTTCTGATTCTTAGGCTTTAGTCGATGGACCTTGCACT	60	70	80	90	100	110	120	130			
											
		10	20	30	40	50	60	70	80			
Experimental	CGGGTGAGGAGACAGGCCAGATCT	140	150									
											
Teórica	CGGGTGAGGAGACAGGCCAGATCT	140	150									
											
		90	100									

Bgl II

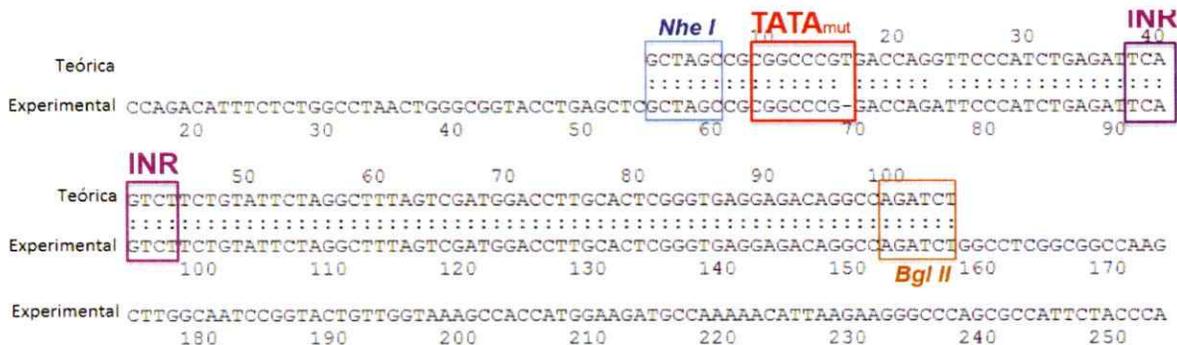
Promotor Core3

						Nhe I	TATA					INR		
Teórica						GCTAGCGCTATAAAAGGACCAGGTTCCCATCTGAGATTCA	20	30	40					
													
Experimental	AAAACATTTTCTCTGGCCTAACTGGCCGTACCTGAGCTCGCTAGCGCTATAAAAGGACCAGGTTCCCATCTGAGATTCA	20	30	40	50	60	70	80	90					
													
						60	70	80	90					
Teórica	INR GTCTTCTGTATTCTAGGCTTTAGTCGATGGACCTTGCACTCGGGTGAGGAGACAGGCCAGATCT	50	60	70	80	90	100							
													
Experimental	GTCTTCTGTATTCTAGGCTTTACTCGATGGACCTTGCACTCGGGTGAGGAGACAGGCCAGATCTGGCCTCGGCGGCCAAG	50	60	70	80	90	100	110	120	130	140	150	160	170
													
		100	110	120	130	140	150	160	170					

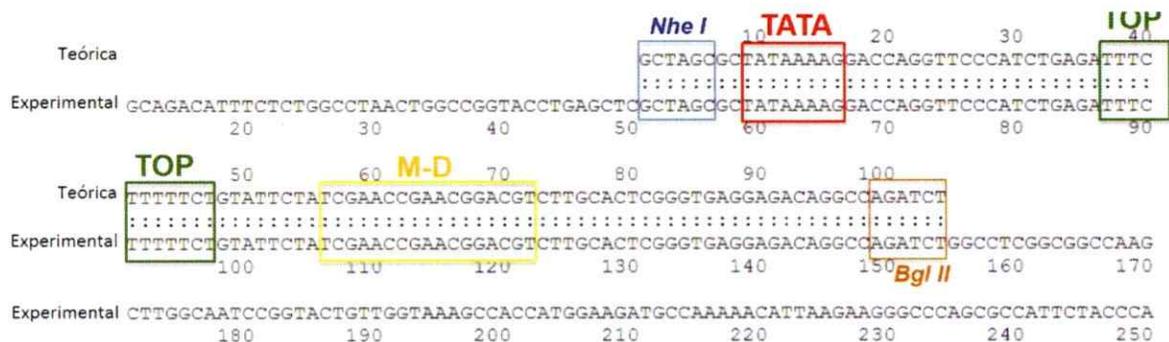
Bgl II

Experimental	CTTGCAATCCGGTACTGTGGTAAAGCCACCATGGAAGATGCCAAAAACATTAAGAAGGCCAGGCCATTCTACCCA	180	190	200	210	220	230	240	250
--------------	---	-----	-----	-----	-----	-----	-----	-----	-----

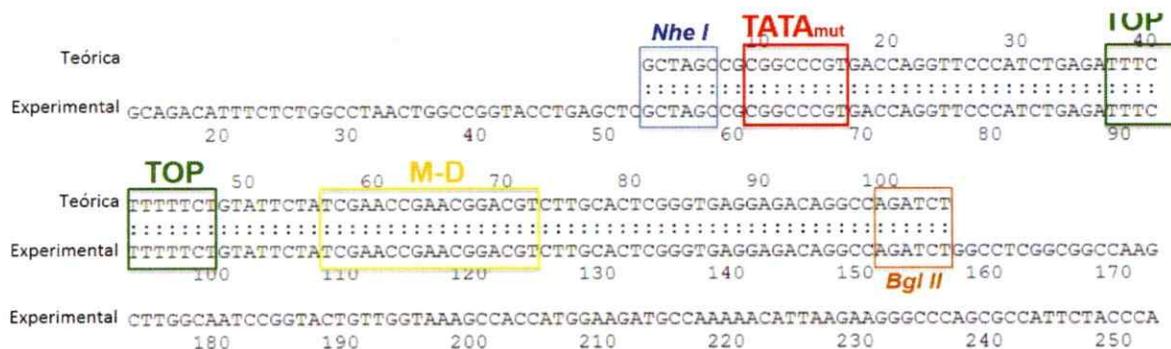
Promotor Core4



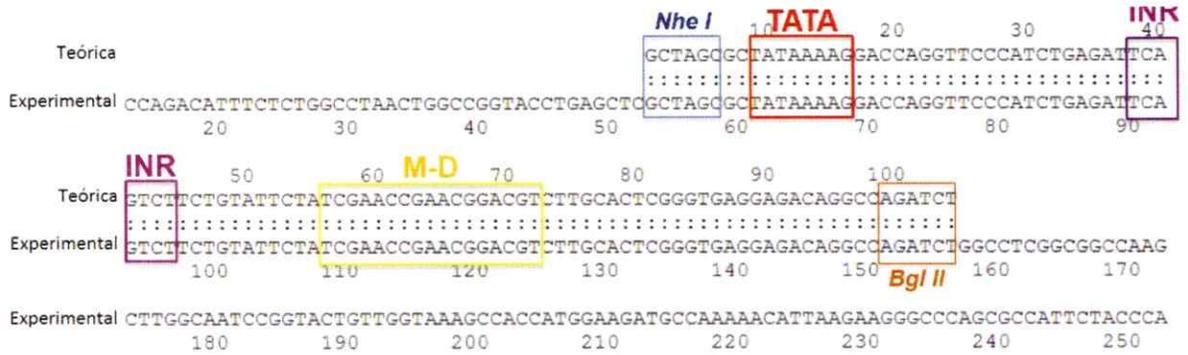
Promotor Core5



Promotor Core6



Promotor Core7



Promotor Core8

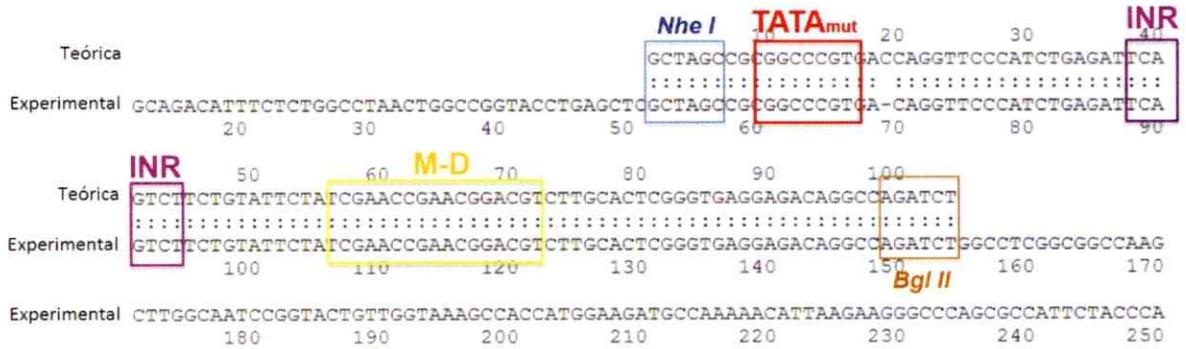


Figura 10. Alineamiento de secuencias teóricas y experimentales de los promotores core. Encerrado en rectángulos se encuentran los motivos TATA, TOP, INR y M-D (MTE y DPE). “TATAmut” corresponde a una secuencia que reemplazó a la caja TATA.

Con el propósito de clonar la región EIE (exón 1, intrón 1 y exón 2 del gen Rpl14) río abajo de cada uno de los 8 promotores de los vectores construidos en el punto anterior, se procedió a amplificar dicha región a partir de ADNg de células CHO-K1, para generar el vector de subclonamiento “pGEMT-EIE”, y posteriormente clonar dicha región río abajo de los promotores de los vectores “pGL Core 1-8” según lo descrito en Materiales y Métodos (sección 2.2.2.1 y 2.2.2.2, respectivamente).

La amplificación de la región EIE a partir de ADNg y su posterior subclonamiento en el vector pGEM-T Easy resultó mucho más rápida que el clonamiento de un promotor *core*, probablemente dado que se clonaba un producto de PCR, y no un producto de hibridación, además de la diferencia de tamaños entre los insertos; alrededor de 50 pb y 1000 pb para los insertos de los promotores y la región EIE, respectivamente.

Otro factor que hizo más rápida la obtención del clonamiento de la región EIE a diferencia del clonamiento de los promotores fue que el vector en donde se clonó, y más específicamente, la región en donde se insertó, interrumpía la región codificante de la enzima β -galactosidasa, por lo cual se podía realizar una selección inicial de las colonias positivas en base al color blanco, las cuales luego eran chequeadas vía PCR colonia. Es importante mencionar que el partidor sentido con el cual se amplificó la región EIE portaba el sitio de restricción para la enzima *BglIII* y dicho sitio hibridaba en la región del primer exón en donde se encontraba una secuencia ATG. De modo similar, el partidor antisentido, que incorporaba el sitio de restricción para la enzima *HindIII*, también modificaba una secuencia ATG presente en el nucleótido 15 del segundo exón.

Ambas modificaciones se realizaron para prevenir que la traducción del ARNm comenzase de modo prematuro y el marco de lectura del gen LUC2 se viese corrido, y que por ende no se produjese la proteína funcional.

A modo de ejemplo, en la Figura 11A se muestra la verificación de la presencia de la región EIE en el vector pGL TATA, formándose el vector pGL TATA-EIE. A partir de ADN plasmidial purificado, se realizó PCR usando partidores externos al inserto; RVprimer3 sentido y LucRv antisentido. En el vector pGL TATA, el producto de amplificación es alrededor de 400 pb (387 pb exactamente), y con la incorporación de la región EIE dicho tamaño aumenta a alrededor de 1400 pb (1493 pb exactamente). En la Figura 11B, se muestra la liberación de la región EIE a partir del vector pGL TATA-EIE, luego de una digestión doble, como segundo modo de confirmación de correcta incorporación, tanto en tamaño como en sentido de inserción.

En la Figura 12, se muestra un esquema de las construcciones generadas; cada uno de los promotores *core* con la región EIE río abajo, comandando la expresión del gen LUC2.

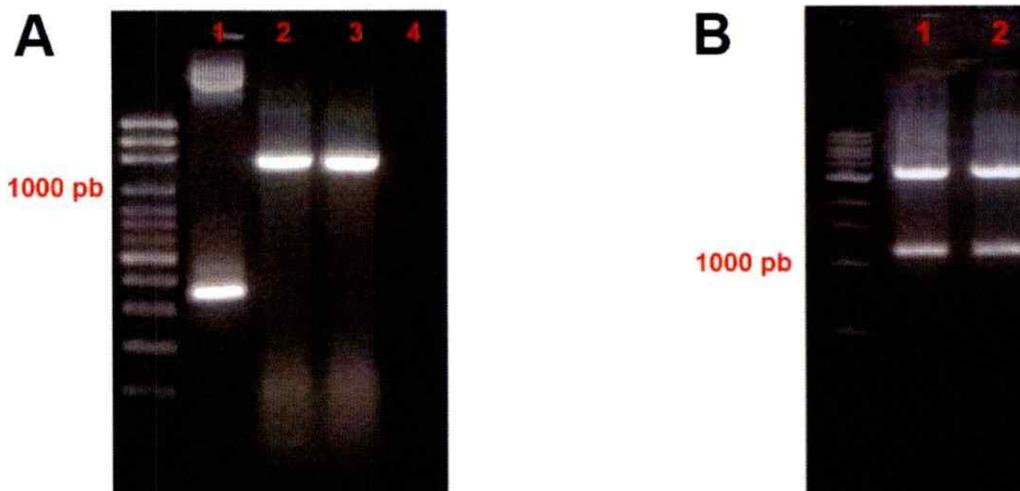


Figura 11. Chequeo de la incorporación de la región EIE en el vector pGL TATA.

Panel A: PCR usando ADN plasmidial como molde. Carril 1; pGL TATA, carril 2-3; pGL TATA-EIE, carril 4: control negativo.

Panel B: Doble digestión de pGL TATA-EIE. Durante 3 horas se digirió el vector con las enzimas BglIII y HindIII a 37°C.

Como marcador de tamaño de banda en pares de bases (pb), se utilizó 100pb y Lambda/HindIII (Winkler, Chile) en A y B, respectivamente, en geles de agarosa al 1%.

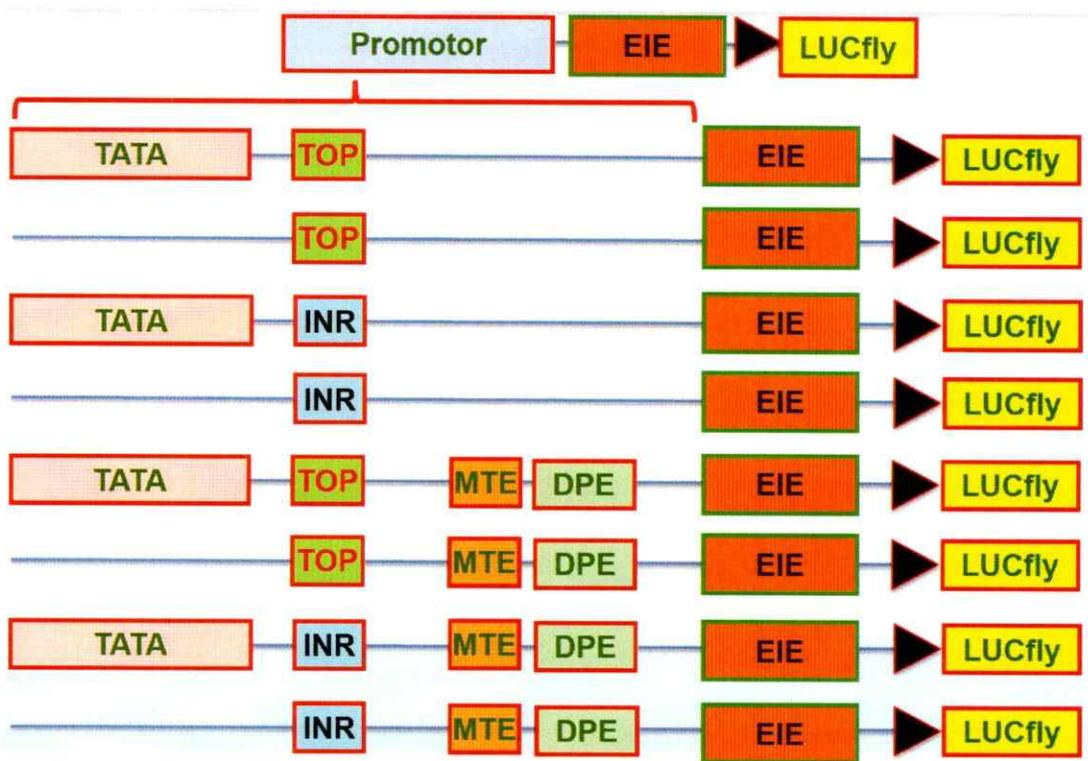


Figura 12. Esquemas vectores “pGL Core-EIE”.

Representación esquemática de los 8 promotores junto a la región EIE. Fueron generados en base a los vectores “pGL Core 1-8”, los cuales portan el gen reportero LUC2 que codifica la luciferasa de luciernaga (“LUCfly” en el esquema).

En la zona superior de la imagen un esquema general de las construcciones, las cuales cuentan de un promotor core, seguido de la región EIE, y luego el gen reportero.

Luego, de arriba hacia abajo se encuentran descritos los elementos de los promotores cuyos motivos son TATA/TOP; TOP; TATA/INR; INR; TATA/TOP/MTE-DPE; TOP/MTE-DPE; TATA/INR/MTE-DPE; y INR/MTE-DPE. Todos estos promotores tienen una extensión de 104 pb, y la región EIE de 1106 pb.

3.2 Objetivo Específico 2: Evaluar la actividad transcripcional de los vectores generados mediante ensayos reporteros utilizando enzimas luciferasas, luego transfectar transientemente células CHO-K1.

3.2.1 Optimización cantidad PEI\ADN en transfección de células CHO-K1.

EL proceso de optimización de las cantidades de PEI y ADN, se llevó a cabo según lo descrito en Materiales y Métodos (sección 2.2.3). Los datos de las diferentes combinaciones de cantidades de PEI y ADN, y las respectivas luminiscencias relativas obtenidas a las 24 horas, fueron analizadas usando “*JMP software, Academic version*” (EE.UU).

Se estimó un modelo que explicaba el 84% de la variabilidad ($R^2=0.844513$), y que tenía los siguientes coeficientes:

$$\text{Luc/Ren} = -0.85221\text{ADN*PEI} + \text{ADN*6.9246} + \text{PEI*2.2494} - \text{ADN}^2*0.5705 + \text{PEI}^2*0.0300$$

La superficie de respuesta generada por el modelo se ve en la Figura 13.

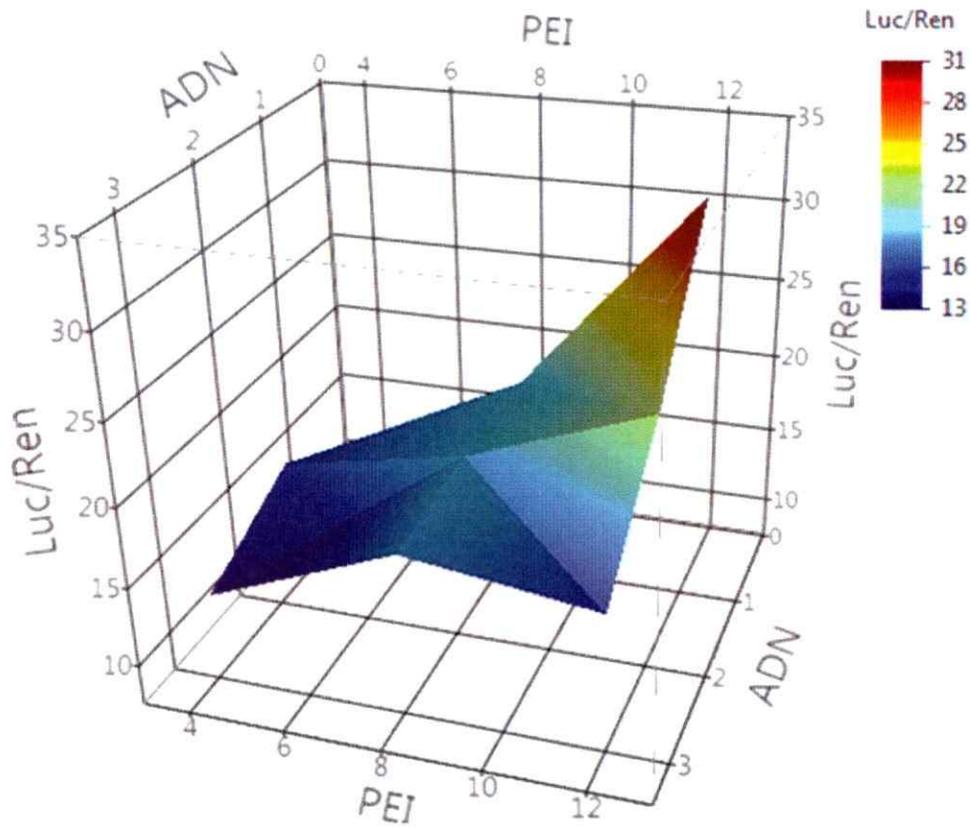


Figura 13. Optimización cantidades de PEI y ADN.

Superficie de respuesta del diseño central compuesto utilizado. A partir de las dos variables independientes; cantidad de PEI y ADN, se modeló la variable luminiscencia (luminiscencia de luciferasa de luciérnaga dividida luminiscencia de luciferasa de “pensamiento de mar”) utilizando un diseño central compuesto, en donde cada variable presenta tres niveles, y se ensayan todas las combinaciones posibles para luego generar una superficie modelo que trata de explicar la variable respuesta.

3.2.2 Co-Transfeccion vectores Luc y medición actividad luciferasa.

El propósito principal del presente seminario de título fue determinar si promotores *core* con diferentes elementos, principalmente TOP, DPE y MTE, eran funcionales en vectores de expresión utilizados en células CHO-K1. Para ello se empleó la actividad luciferasa como reportero. La Figura 14 muestra la luminiscencia relativa de los ocho PCs generados, además de la actividad de un vector sin promotor (pGL 4.17) y la del vector pGL CoreCMV, el cual contiene como promotor del gen reportero al promotor *core* de CMV. En la Figura 15 se muestra los 8 promotores *core* junto a sus respectivas versiones que incorporaron la región EIE. Dado que se esperaba que la actividad luciferasa fuese mayor, se midió junto al vector que contenía *enhancer* y promotor completo de CMV.

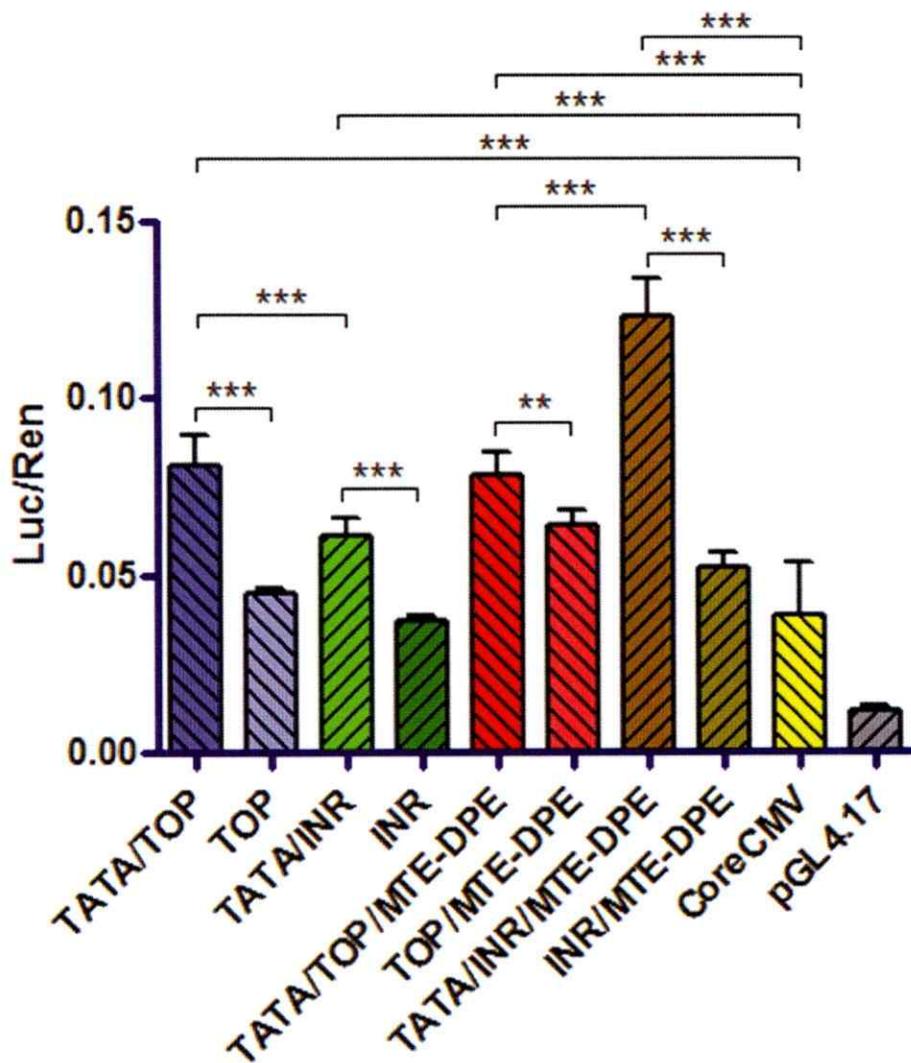


Figura 14. Actividad luciferasa en vectores “pGL Core 1-8”.

Cada uno de los vectores con diferentes promotores comandando el gen codificante de luciferasa de luciérnaga (LUC2) indicados en el eje de la abscisa fue co-transfectado con el vector pGL 4.74, el cual porta el promotor HSV-TK comandando la expresión del gen que codifica la luciferasa de *Renilla reniformis* (Ren). El eje de la ordenada representa el cociente entre estas dos luminiscencias (Luc/Ren). Cada experimento se realizó al menos tres veces. Se realizó ANOVA de una vía para las comparaciones múltiples, seguido de la prueba de Newman-Keuls para las comparaciones a posteriori. Los niveles de significancia están representados por: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$.

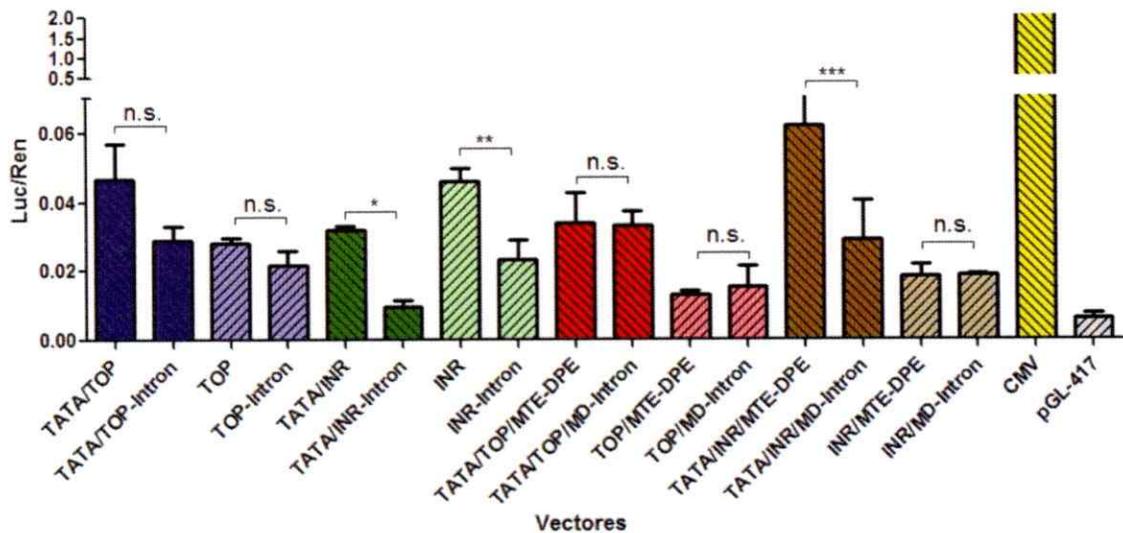


Figura 15. Actividad luciferasa en vectores “pGL Core 1-8- EIE”.

Los mismos vectores “pGL Core” fueron ensayados junto a su versión que incluye la región EIE. Al igual que en la figura 13, cada uno de los vectores indicados en el eje de la abscisa (portando el gen LUC2) fue co-transfectado con el vector pGL 4.74, el cual porta el promotor HSV-TK comandando la expresión del gen que codifica la luciferasa de *Renilla reniformis* (Ren). El eje de la ordenada representa el cociente entre estas dos luminiscencias (Luc/Ren). Cada experimento se realizó al menos tres veces. Se realizó ANOVA de una vía para las comparaciones múltiples, seguido de la prueba de Newman-Keuls para las comparaciones a posteriori. Los niveles de significancia están representados por: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$; ns: no significativa estadísticamente la diferencia.

4. DISCUSIÓN

Los vectores de expresión son herramientas esenciales en la producción de productos biofarmacéuticos como los anticuerpos recombinantes. Es por ello que el desarrollo de sistemas de expresión que mantengan niveles de producción de proteínas recombinantes acordes con las necesidades de manufacturación es una valiosa fuente de innovación. El presente seminario de título tuvo como propósito explorar la utilidad de diferentes promotores *core* portando combinaciones de los elementos TOP, caja TATA, TOP, INR, MTE y DPE, en uno de los sistemas de cultivo más empleados en la industria biotecnológica como lo son las líneas celulares derivadas de las células CHO. Dichos promotores fueron clonados en vectores de expresión que carecían de una región promotora que dirigiese la transcripción del gen reportero LUC2 que codifica la luciferasa de luciérnaga. Mediante transfecciones transientes y la medición de luminiscencia se evaluó la actividad transcripcional lograda por cada uno de los promotores sintéticos construidos.

Como se mencionó anteriormente, en base a la literatura estudiada se eligió como secuencia base a la cual agregar o quitar elementos a la región del promotor *core* del gen Rpl14 (codificante de la proteína ribosomal L14), el cual sólo contiene el elemento TOP ubicado en el sitio de inicio de la transcripción. Esta elección se efectuó basándose en los siguientes motivos principalmente; la elevada homología de secuencia nucleotídica entre los genes Rpl14 de humano y hámster chino (Perry, 2005), la evidencia que en

tejido ovárico dicho gen es uno de los genes ribosomales con mayor expresión (Bortoluzzi y col., 2001), y el aumento de expresión a la cual se ve sujeto este tipo de genes en presencia de un medio de cultivo rico en nutrientes, dada la naturaleza del organelo del cual forman parte (Powers & Walter, 1999).

Sin duda alguna, en un contexto silvestre o sin intervención, la alta expresión del gen Rpl14 o de cualquier gen no se debe solamente al promotor, ya que también son cruciales otras regiones regulatorias como *enhancers*, aisladores, y regiones transcritas no traducidas, entre otros motivos. Cabe señalar que este trabajo se enmarcó en la línea de investigación del proyecto “Desarrollo de un sistema de expresión optimizado para la producción de anticuerpos terapéuticos” del Centro de ImunoBioTecnología, de la Facultad de Medicina, de la Universidad de Chile, en la cual también se trabaja en la modificación de regiones diferentes al promotor, y en ese sentido este seminario fue complementario al resto de la investigación, y se limitó a estudiar sólo el efecto del promotor *core* como uno de todos los posibles blancos de mejoramiento de los vectores de expresión.

Las secuencias utilizadas de los motivos de promotores fueron las secuencias canónicas descritas en la literatura, excepto para los motivos MTE y DPE para los cuales en vez de usar la secuencia canónica se usaron secuencias alternativas, obtenidas de un trabajo de optimización de dichos motivos, en donde se reemplazó cada nucleótido de estos dos elementos por una de las cuatro bases nitrogenadas presentes en el ADN y se evaluó su actividad transcripcional (Theisen y col., 2010). Finalmente, las secuencias de los

elementos utilizadas fueron; caja TATA: TATAAAAG, INR: TCAGTC, TOP: TTCTTTTCT, MTE-DPE: TCGAACCGAACGGACGT.

4.1 Optimización cantidades de PEI y ADN.

Una vez generados los vectores con los distintos promotores *core*, y sus versiones respectivas con la región EIE incorporada, fue necesario establecer las condiciones de co-transfección para evaluar la actividad transcripcional de las construcciones hechas.

En la actualidad existe una variedad de agentes químicos para la transfección, los cuales podrían ser clasificados en lipídicos y no lipídicos. Dentro del primer grupo, los lípidos catiónicos son capaces de unirse a las células mediante interacciones electrostáticas con componentes de la membrana cargados negativamente (Van der Aa y col., 2007), desestabilizando la bicapa lipídica para permitir el ingreso del ADN (Wattiaux y col., 1997) mediante la fusión de los liposomas catiónicos con las vesículas endocíticas (Zuhorn y col., 2005). Dentro del segundo grupo de agentes de transfección se encuentra el polímero catiónico PEI (polietilenimina), el cual a diferencia de la transferencia vía vesículas endocíticas como el caso anterior, lo hace a través de la vía de las caveolas, evitando su degradación vía endosomas tempranos y tardíos (Takei & Haucke, 2001), permitiendo su posterior difusión al núcleo, especialmente durante la mitosis (Pollard y col., 1998).

El PEI forma complejos con el ADN mediante una contra-condensación, generando estructuras con un centro eléctricamente neutro y una superficie con carga positiva

necesaria para la unión a la célula (Behr y col., 1989). La magnitud de esta carga positiva superficial viene dada por la razón de las cargas de cationes presentes en el PEI y de aniones presentes en el ADN, es por ello que es de crucial importancia establecer una adecuada razón de las cantidades de PEI y ADN.

Para transfecciones en células de origen mamífero, altos valores de la razón PEI/ADN son generalmente usados (Boussif y col., 1995), a pesar de que más del 80% del PEI permanece sin unir ADN (Clamme y col., 2003). Este exceso de PEI es esencial para la entrega efectiva de ADN, pero también está asociada a una elevada toxicidad celular (Boeckle y col., 2004). En nuestro caso, cantidades superiores a 12 μg de PEI, comenzaron a producir mortalidades superiores al 50%, por lo cual se tomó ese valor como límite superior en el intervalo de cantidades de PEI en el proceso de optimización. Como resultado del ensayo de optimización, se determinó que 0.5 μg de ADN y 12 μg de PEI producían la mayor cantidad de luminiscencia relativa luego de 24 h posterior a la transfección. Dichas cantidades representan una razón PEI/ADN igual a 24/1 (p/p).

Es importante mencionar que las cantidades óptimas dependen de la naturaleza celular del sistema de cultivo a transfectar, así por ejemplo, para las células de origen humano HEK293T, la razón óptima de PEI:ADN ha sido descrita alrededor de 2:1 (Aricescu y col., 2006).

En cuanto a optimizaciones de este tipo usando células CHO, existe un trabajo en donde se determinó la razón PEI/ADN óptima para la línea celular CHO-DXB11 (la cual deriva de la línea CHO-K1) y resultó ser 36/1 (Reisinger y col., 2009) valor no tan

alejado de 24/1. Interesante mencionar que la razón 24/1 fue ensaya por Reisinger y col., y representó el 90% de viabilidad, y a pesar de no ser la razón que presentó mayor actividad reportera, estuvo entre las que más lo hacían. La diferencia en la efectividad de la transfección entre razones 36/1 y 24/1 podría deberse principalmente al tiempo post-transfección transcurrido para medir la actividad reportera; en nuestro caso ésta fue medida a las 24 h, y en Reisinger y col. fue a las 72 h. Así, a pesar del número de células muertas a causa de la mayor cantidad de PEI (36/1), el número de células transfectadas exitosamente fue mayor, y ellas tuvieron tiempo para reproducirse 3 veces (en vez de 1 como en nuestro caso, puesto que luego de 24 horas el cultivo se reproduce) y elevar el número de células que produjesen la actividad reportera.

4.2 Actividad luciferasa vectores “pGL Core” y “pGL Core-EIE”.

El propósito final del presente seminario fue poder determinar la veracidad de la hipótesis planteada inicialmente. Para ello se construyeron promotores que contaban con la presencia de elementos TOP, MTE y DPE. Además se incorporaron los elementos caja TATA e INR. Todos estos elementos se combinaron de diferentes modos, y se construyeron 8 promotores, los cuales fueron clonados en el vector pGL 4.17 para que dirigiesen la expresión del gen codificante de la enzima luciferasa.

Como se puede apreciar en la Figura 14, los ocho vectores con promotores *core* construidos tuvieron una luminiscencia relativa (Luc/Ren) mayor que la del vector sin promotor (pGL 4.17), con lo cual podríamos decir que todas las combinaciones fueron

funcionales en los ensayos de transfección transientes. El promotor que logró la menor luminiscencia fue aquel que sólo portaba el elemento INR. Por su contraparte, el promotor que logró la mayor luminiscencia fue aquel que contenía la caja TATA, INR, MTE y DPE. La funcionalidad de esta última conformación ya había sido reportada en las células de origen humano HeLa S3 (Juven-Gershon y col., 2006). En aquel estudio se utilizó la secuencia canónica de los motivos que conformaban al promotor *core*, y en el presente trabajo se utilizó la secuencia optimizada de los motivos MTE, DPE y la región entre ellos, puesto que mediante ensayos de foto-entrecruzamiento y digestión con DNaseI se demostró un aumento de la afinidad de TFIID por dicha secuencia (Juven-Gershon & Kadonaga, 2010).

Comparando los promotores que sólo tienen los motivos TOP e INR, los cuales ocupan la posición alrededor del sitio de inicio de la transcripción, podemos ver que entre ellos no hubo una diferencia significativa, lo cual podría deberse a que las afinidades de TRF2 y TAF1/2 del complejo TFIID de la maquinaria basal, por los motivos TOP e INR respectivamente, son similares. Ahora bien, si a esos mismos promotores se les incorporan los elementos MTE y DPE posicionados a partir del +18 hasta el +33, se observa que los niveles de luminiscencia relativa aumentan significativamente ($p < 0.01$) en ambos casos, y estos aumentos no son significativamente diferentes entre ellos. Considerando el promotor con los motivos INR, DPE y MTE, el efecto positivo en el aumento de la actividad reportera es concordante con efectos sinérgicos entre INR y DPE reportados en la literatura (Burke y Kadonaga, 1997; Zhou & Chiang 2001), al

igual que el efecto sinérgico entre INR y MTE en células de origen humano (Lim y col., 2004; Theisen y col., 2010).

Por otra parte, si a los promotores que sólo tienen los motivos TOP o INR, se les incorporan la caja TATA, la actividad reportera aumenta en ambos significativamente ($p < 0.001$), siendo mayor el incremento en la combinación TATA-TOP con respecto a TATA-INR. El comportamiento cooperativo entre TATA e INR ha sido reportado y se debe principalmente a un aumento de la afinidad de complejo TFIID por el promotor evidenciado mediante ensayos de entrecruzamiento ADN-Proteína y protección del promotor de la digestión por DNasa I (O'Shea-Greenfield & Smale, 1992; Emami y col., 1997). Es importante recordar que las proteínas que unen la caja TATA e INR, son subunidades del mismo complejo, lo cual podría ser la razón principal por la cual se permite la estabilización de la interacción de todo el complejo en la región promotora, favoreciendo luego una formación del complejo de pre-iniciación (CPI) funcional en el promotor. En ese sentido, es interesante el aumento mostrado por el promotor TATA-TOP, ya que el motivo TOP es reconocido por TRF2, el cual no forma parte del complejo TFIID, y a pesar de ello, se logra una cooperación entre ambas proteínas para lograr un mayor nivel de transcripción. A pesar de desconocer el mecanismo por el cual TATA y TOP pueden actuar sinérgicamente, en la práctica este hecho se ha utilizado con propósitos biotecnológicos; Hitti y col., incorporó una caja TATA en el promotor (carente de ella) del gen S30 (codificante de la proteína ribosomal S30) logrando una mayor expresión, y un vector patentado.

Ahora bien, si a los promotores con los motivos TATA-TOP o TATA-INR, se les agrega los elementos MTE y DPE, vemos que el primero no aumenta, a diferencia del segundo que logra un incremento significativo ($p < 0.001$), convirtiéndose en el promotor *core* que consigue la expresión de luminiscencia relativa más alta de los vectores ensayados (Figura 14). El hecho de que los niveles de expresión logrado por los promotores con motivos TATA-TOP y TATA-TOP-MTE-DPE no sean diferentes entre ellos, pero mayores que el vector que sólo contiene el motivo TOP o sólo TATA (dado por el vector que porta el promotor *core* de CMV, "CoreCMV"), podría explicarse bajo un escenario putativo en el cual TRF2 se une a TOP, y TBP se une a la caja TATA, pero las subunidades TAF6 y TAF9, están viendo impedido su unión a MTE-DPE. Si el mismo análisis se hace para los promotores con motivos TATA-INR y TATA-INR-MTE-DPE, vemos que el primero logra una expresión mayor a los vectores que sólo portan la caja TATA o sólo el motivo INR, y que luego al incorporar los motivos MTE-DPE la expresión aumenta nuevamente, lo cual podría deberse a que las subunidades TAF6/9 pueden unirse a MTE-DPE, ya que TRF2 no fue reclutado puesto que no se encontraba ningún motivo TOP.

También fue posible determinar que al comparar las luminiscencias relativas de los promotores con motivos TATA-TOP, TATA-INR, TATA-INR-MTE-DPE, y TATA-TOP-MTE-DPE, con el promotor *core* de CMV (el cual sólo contiene caja TATA como elemento), se observó que cada uno de los cuatro promotores presentaba una actividad

reportera significativamente mayor ($p < 0.001$), reforzando la idea de que TATA actúa sinérgicamente con los elementos TOP o INR o cualquiera de ellos más MTE y DPE.

Luego de analizar la actividad reportera entre los promotores *core*, se quiso estudiar el efecto de la incorporación de la región EIE, la cual contenía parte del primer exón, primer intrón y parte del segundo exón del gen Rpl14. Para ello se volvió a transfectar utilizando células frescas recién descongeladas. Se transfectó utilizando las mismas condiciones, y ahora el análisis se volcó en la comparación entre cada promotor *core* y su versión respectiva con región EIE. A partir de la figura 15, podemos observar que la luminiscencia relativa en general fue menor a los ensayos realizados previamente, quizás por el hecho de exponer cultivos tan frescos a agentes tóxicos como el PEI. La presencia de la región EIE, en general presentó tendencias a la baja en la expresión de la actividad reportera, pero esta disminución sólo fue significativa en promotores cuya expresión era alta, como lo es el caso de TATA-INR-MTE-DPE, el cual era el que sin la región EIE lograba los niveles más altos de luminiscencia. Estos resultados son contrario a lo que se esperaba según los reportes de potenciamiento de la transcripción ante la presencia de una región intrónica mostrados inicialmente. Las razones por las cuales la incorporación de esa región entre el promotor y el gen reportero tuvo implicancias negativas podrían ser: i) que la región EIE portase secuencias de ADN que sean reconocidas por represores de la transcripción, ii) que ocurriese un corte y empalme alternativo al que se esperaba. La primera de estas posibles causas, fue en alguna medida prevista y se modificaron las

secuencias ATG, presentes en el primer y segundo exón de la región EIE mediante la incorporación de sitios de restricción de las enzimas BglIII y HindIII en los partidores utilizados para la amplificación de dicha región, y tal modificación fue corroborada vía secuenciación automática.

La posibilidad de que hubiese ocurrido un corte y empalme alternativo al esperado acorde lo publicado para el gen Rpl14, no fue evaluado, pero podría haberse determinado mediante la extracción del ARN y su posterior uso en un RT-PCR utilizando partidores que hibridasen en las regiones correspondientes a los exones, y luego visualizar los productos de PCR en un gel de agarosa de tal modo que se hubiese visto si el tamaño era el que correspondía o no luego de la escisión del intrón de alrededor de 1000 pb.

Con respecto a la presencia de secuencias que unan represores, inicialmente no se evaluó. La búsqueda de sitios de unión a factores de transcripción con actividad represora es posible de modo bioinformático. Dicha información, hubiese entregado posibles candidatos de represores, con respecto a lo encontrado en organismos modelos más estudiados como *Mus musculus*, en donde ya está ampliamente caracterizada este tipo de información en bases de datos (Dreos y col, 2015) puesto que dicho genoma se conoce desde mucho antes que el de hámster chino que se obtuvo en años más recientes además, dicha información no puede ser tomada como cierta en su totalidad para el hámster chino, ya que no todos sus genes son homólogos al del ratón, además no todos

los genes del genoma de hámster chino se expresa en las células CHO-K1 (Lewis y col., 2013).

La posibilidad de que motivos en la secuencia de un intrón reprima la transcripción no es algo excepcional, y existen estudios que han reportado dicho fenómeno (Carstens y col., 1998; Deyholos & Sieburth, 2000) por lo cual no podemos descartar que esto sea la causa de la disminución vista en nuestros ensayos.

Los sistemas biológicos son sistemas complejos, y la expresión génica puede ser regulada en diferentes niveles. Uno de ellos es el inicio de la transcripción. Sólo fijándonos en ese nivel ya existen múltiples factores que influyen en el proceso, entre los cuales, el presente seminario de título pretendió abordar la contribución del promotor *core*. A pesar lo que se pensaba en un comienzo, esta región de los genes puede presentar diferentes motivos o elementos, y en muchos casos muy diferentes a la bien estudiada caja TATA. Estudios a nivel genómico que busquen nuevos motivos en los promotores evidenciarán elementos desconocidos y representarán nuevos desafíos para entender el modo en el que operan. Recientemente, Marbach-Bar y col, ha descubierto un nuevo elemento al cual denominaron DTIE, presente en promotores sin caja TATA, contribuyendo al entendimiento de ese gran grupo de genes carentes de caja TATA que aun así reclutan la maquinaria transcripcional.

La diversidad en los motivos que pueden conformar los promotores, viene de la mano con la diversidad de proteínas que son capaces de unirse a ellos y reclutar la maquinaria

transcripcional para lograr un ensamble funcional del complejo de pre-iniciación. En ese aspecto, cobra sentido la idea de que los “factores de transcripcional generales” en realidad no son tan generales como se creyó, y en conjunto con los promotores *core* conforman una nueva fuente de riqueza en la regulación de familias de genes ya sea dentro de una misma células o entre diferentes tipos celulares de un mismo organismo (Goodrich & Tjian, 2010).

El entendimiento del funcionamiento de los diferentes motivos de promotores *core*, puede ser aprovechado para el diseño racional de promotores, que junto a regiones más amplias como los *enhancers*, puedan servir para construir vectores de expresión, en base a secuencias de ADN propias del organismo y así ser utilizado en sistemas de expresión de proteínas de interés y de este modo evitar las adversidades presentes cuando se usan regiones promotoras de origen viral.

En la actualidad existen métodos directos para encontrar genes con alta expresión cuyas regiones promotoras puedan ser utilizadas con propósitos biotecnológicos, estos métodos son más sofisticado y costosos que los empleados en biología molecular clásica como el “*promoter trapping*” (von Melchner & Ruley, 1989), y requieren de un grado de conocimiento previo del organismo de estudio. En ese escenario, dado que ya se cuenta con el genoma anotado del hámster chino (Xu y col, 2011; Rupp y col, 2014), se podrían efectuar ensayos de RNA-seq para identificar los genes más expresados y/o realizar ensayos de CHIP-seq contra alguna de las subunidades de la RNA polimerasa II, de tal modo de identificar los genes que mayormente unen a la polimerasa y con ello

posiblemente ser los más expresados, y a partir de ellos analizar sus secuencias regulatorias con el propósito de utilizar posteriormente en la construcción de vectores de expresión.

5. Conclusiones

En el presente trabajo se pudo determinar mediante transfección transiente y ensayos reporteros en base a actividades luciferasas, que en células CHO-K1 al igual que en otros sistemas reportados, existe una complementariedad entre TOP y MTE-DPE, la caja TATA y TOP, TATA e INR, y este último par con los motivos MTE y DPE. A partir de ello podemos concluir que a pesar de que naturalmente dichos elementos no se encuentren siempre presente en los mismos promotores, podemos mediante biología sintética rediseñar promotores en base a combinaciones de motivos que presenten elevadas actividades reporteras y utilizarlas en conjunto con otras regiones regulatorias, como por ejemplo *enhancers*, para construir vectores de expresión acorde a las necesidades del modelo empleado.

6. Bibliografia

Aricescu, A. R., Lu, W., Jones, E. Y. 2006. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**: 1243-1250.

Behr, J.P., Demeneix, B., Loeffler, J.P., Perez-Mutul, J. 1989. Efficient gene transfer into mammalian primary endocrine cells with lipopolyamine-coated DNA. *Proc. Natl. Acad. Sci. U.S.A.* **86**: 6982-6986.

Bieberstein, N., Carrillo F., Straube O., Neugebauer, K.M. 2012. First Exon Length Controls Active Chromatin Signatures and Transcription. *Cell Rep.* **2**: 62-68.

Boeckle, S., von Gersdorff, K., van der Piepen, S., Culmsee, C., Wagner, E., Ogris, M. 2004. Purification of polyethylenimine polyplexes highlights the role of free polycations in gene transfer. *J. Gene Med.* **6**: 1102-1111.

Bortoluzzi, S., d'Alessi, F., Romualdi, C., Danieli, G.A. 2001. Differential expression of genes coding for ribosomal proteins in different human tissues. *Bioinformatics.* **17**: 1152-1157.

Boussif, O., Lezoualc'h, F., Zanta, M.A., Mergny, M.D., Scherman, D., Demeneix, B., Behr, J.P. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. U.S.A.* **92**: 7297-7301.

Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E., Palmiter, R.D. 1988. Introns increase transcriptional efficiency in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **85**: 836-840.

Burke, T.W., Kadonaga, J.T. 1997. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes Dev.* **11**: 3020-3031.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A., Taylor, M.S., Engström, P.G., Frith, M.C., Forrest, A.R., Alkema, W.B., Tan, S.L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S.M., Wells, C.A., Orlando, V., Wahlestedt, C., Liu, E.T., Harbers, M., Kawai, J., Bajic, V.B., Hume, D.A., Hayashizaki, Y. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nat. Genet.* **38**: 626-635.

Carstens, R.P., McKeehan, W.L., Garcia-Blanco, M.A. 1998. An intronic sequence element mediates both activation and repression of rat fibroblast growth factor receptor 2 pre-mRNA splicing. *Mol. Cell. Biol.* **4**: 2205-17.

Choi, Y.J., Kang, S.J., Kim, Y.J., Lim, Y.B., Chung, H.W. 2010. Comparative studies on the genotoxicity and cytotoxicity of polymeric gene carriers polyethylenimine (PEI) and polyamidoamine (PAMAM) dendrimer in Jurkat T-cells. *Drug Chem. Toxicol.* **33**: 357-366.

- Clamme, J.P., Azoulay, J., Mély, Y.** 2003. Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophys. J.* **84**: 1960-1968.
- Cooper, S.J., Trinklein, N.D., Anton, E.D., Nguyen, L., Myers, R.M.** 2006. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res.* **16**: 1-10.
- Damgaard, C.K., Kahns, S., Lykke-Andersen, S., Nielsen, A.L., Jensen, T.H., and Kjems, J. A.** 2008. 5' splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol. Cell.* **29**: 271-278.
- Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P., Reed, R.** 2007. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell.* **26**: 867-881.
- Deng, W., Roberts, S.G.** 2005. A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev.* **19**: 2418-2423.
- Deyholos, M.K., Sieburth, L.E.** 2000. Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron. *Plant Cell.* **12**: 1799-1810.
- Dreos, R., Ambrosini, G., Périer, R.C., Bucher, P.** 2015. The Eukaryotic Promoter Database: expansion of EPDnew and new promoter analysis tools. *Nucleic Acids Res.* **43**: D92-D96.

El Kaderi, B., Medler, S., Raghunayakula, S., Ansari, A. 2009. Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. *J. Biol. Chem.* **284**: 25015-25025.

Emami, K.H., Jain, A., Smale, S.T. 1997. Mechanism of synergy between TATA and initiator: synergistic binding of TFIID following a putative TFIIA-induced isomerization. *Genes Dev.* **11**: 7-3019.

Florea, B.I., Meaney, C., Junginger, H.E., Borchard, G. 2002. Transfection efficiency and toxicity of polyethylenimine in differentiated Calu-3 and nondifferentiated COS-1 cell cultures. *AAPS PharmSci.* **4**: 1-11.

Fong, Y.W., Zhou, Q. 2001. Stimulatory effect of splicing factors on transcriptional elongation. *Nature.* **414**: 929-933.

Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A., Proudfoot, N.J. 2002. Promoter proximal splice sites enhance transcription. *Genes Dev.* **16**: 2792-2799.

Goodrich, J., Tjian, R. 2010. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nat. Rev. Genet.* **11**: 549-558.

Hamilton, T.L., Stoneley, M., Spriggs, K.A., Bushell, M. 2006. TOPs and their regulation. *Biochem. Soc. Transac.* **34**: 12-16.

Hauser, H., y Wagner, R. 1997. Heterologous expression of genes in mammalian cells, in *Mammalian cell biotechnology in protein production*. pp 3-32. Walter de Gruyter, Berlin.

- Hitti, E., Al-Yahya, S., Al-Saif, M., Mohideen, P., Mahmoud, L., Polyak, S.J., Khabar, K.S.** 2010. A versatile ribosomal protein promoter-based reporter system for selective assessment of RNA stability and post-transcriptional control. *RNA*. **16**: 1245-1255.
- Juven-Gershon, T., Cheng, S., Kadonaga, J.T.** 2006. Rational design of a super core promoter that enhances gene expression. *Nat. Methods*. **3**: 917-922.
- Juven-Gershon, T., Kadonaga, J.T.** 2010. Regulation of Gene Expression via the Core Promoter and the Basal Transcriptional Machinery. *Dev. Biol.* **339**: 225-229.
- Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., Ren, B.** 2005. A high-resolution map of active promoters in the human genome. *Nature*. **436**: 876-880.
- Lemon, B., Tjian, R.** 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* **14**: 2551-2569.
- Levy, S., Avni, D., Hariharan, N., Perry, R.P., Meyuhas, O.** 1991. Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. *Proc. Natl. Acad. Sci. U.S.A.* **88**: 3319-3323.
- Lewis, N., Liu, X., Li, Y., Nagarajan, H., Yerganian, G., O'Brien, E., Bordbar, A., Roth, A., Palsson, B.** 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. *Nat. Biotechnol.* **31**: 759-765.
- Li, H., Chen, D., Zhang, J.** 2012. Analysis of Intron Sequence Features Associated with Transcriptional Regulation in Human Genes. *PLoS One*. **7**: 1-9.

- Lim, C.Y., Santoso, B., Boulay, T., Dong, E., Ohler, U., Kadonaga, J.T.** 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.* **18**: 1606-1617.
- Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., Fu, X.D.** 2008. The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* **15**: 819-826.
- Malecová, B., Gross, P., Boyer-Guittaut, M., Yavuz, S., Oelgeschläger, T.** 2007. The Initiator Core Promoter Element Antagonizes Repression of TATA-directed Transcription by Negative Cofactor NC2. *J. Biol. Chem.* **282**: 24767-24776.
- Marbach-Bar, N., Bahat, A., Ashkenazi, S., Golan-Mashiach, M., Haimov, O., Wu, S.Y., Chiang, C.M., Puzio-Kuter, A., Hirshfield, K.M., Levine, A.J., Dikstein, R.** 2015. DTIE, a novel core promoter element that directs start site selection in TATA-less genes. *Nucleic Acids Res.* **44**:1080-94.
- O'Shea-Greenfield, A., Smale, S.T.** 1992. Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. *J. Biol. Chem.* **267**: 1391-1402.
- Parry, T.J., Theisen, J.W., Hsu, J.Y., Wang, Y.L., Corcoran, D.L., Eustice, M., Ohler, U., Kadonaga, J.T.** 2010. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev.* **24**: 2013-2018.

- Parra, G., Bradnam, K., Rose, A.B., Korf, I.** 2011. Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants. *Nucleic Acids Res.* **39**: 5328-5337.
- Pollard, H., Remy, J. S. Loussouarn, G., Demolombe, S., Behr, J.P., Escande, D.** 1998. Polyethylenimine but not cationic lipids promotes transgene delivery to the nucleus in mammalian cells. *J. Biol. Chem.* **273**: 7507-7511.
- Perry, R.** 2005. The architecture of mammalian ribosomal protein promoters. *BMC Evol. Biol.* **5**: 1-16.
- Powers, T., Walter, P.** 1999. Regulation of Ribosome Biogenesis by the Rapamycinsensitive TOR-signaling Pathway in *Saccharomyces cerevisiae*. *Mol. Biol. Cell.* **10**: 987-1000.
- Reisinger, H., Steinfeldner, W., Katinger, H., Kunert, R.** 2009. Serum-free transfection of CHO cells with chemically defined transfection systems and investigation of their potential for transient and stable transfection. *Cytotechnology.* **60**: 115-123.
- Roepcke, S., Zhi, D., Vingron, M., Arndt, P.F.** 2006. Identification of highly specific localized sequence motifs in human ribosomal protein gene promoters. *Gene.* **365**: 48-56.
- Rupp, O., Becker, J.Brinkrolf, K., Timmermann, K., Borth, N., Pühler, A., Noll, T., Goesmann, A.** 2014. Construction of a Public CHO Cell Line Transcript Database Using Versatile Bioinformatics Analysis Pipelines. *PLoS One.* **9**:1-13.

- Shandilya, J. Roberts, S.** 2012. The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochim. Biophys. Acta.* **1819**: 391-400.
- Takei, K., Haucke, V.** 2001. Clathrin-mediated endocytosis: membrane factors pull the trigger. *Trends Cell Biol.* **11**: 385-391.
- Theisen, J.W., Lim, C.Y., Kadonaga, J.T.** 2010. Three Key Subregions Contribute to the Function of the Downstream RNA Polymerase II Core Promoter. *Mol. Cell Biol.* **14**: 3471-3479.
- Thomas, M.C., Chiang, C.M.** 2006. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**: 105-178.
- Uechi, T., Tanaka, T., Kenmochi, N. A.** 2001. Complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders. *Genomics.* **72**: 223-230.
- Van der Aa, M.A., Huth, U.S., Hafele, S.Y., Schubert, R., Oosting, R.S., Mastrobattista, E., Hennink, W.E., Peschka-Suss, R., Koning, G.A., Crommelin, D.J.** 2007. Cellular uptake of cationic polymer-DNA complexes via caveolae plays a pivotal role in gene transfection in COS-7 cells. *Pharm. Res.* **24**: 1590-1598.
- Venters, B. J., Pugh, B.F.** 2009. How eukaryotic genes are transcribed. *Crit. Rev. Biochem. Mol. Biol.* **44**: 117-141.
- von Melchner, H., Ruley, H.E.** 1989. Identification of cellular promoters by using a retrovirus promoter trap. *J Virol.* **63**:3227-33.

- Wang, Y.L., Duttke, S.H., Chen, K., Johnston, J., Kassavetis, G.A., Zeitlinger, J., Kadonaga, J.T.** 2014. TRF2, but not TBP, mediates the transcription of ribosomal protein genes. *Genes Dev.* **28**: 1550-1555.
- Wattiaux, R., Jadot, M., Warnier-Pirotte, M.T., Wattiaux-De Coninck, S.** 1997. Cationic lipids destabilize lysosomal membrane in vitro. *FEBS Lett.* **417**: 199-202.
- Xu, X., Nagarajan, H., Lewis, N., Pan, S., Cai, Z., Liu, X., Chen, W., Xie, M., Wang, W., Hammond, S., Andersen, M., Neff, N., Passarelli, B., Koh, W., Fan, C., Wang, J., Gui, Y., Lee, K., Betenbaugh, M., Quake, S., Famili, I., Palsson, B., Wang, J.** 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nature Biotechnol.* **29**:735–741.
- Yang, Y., Mariati, Chusainow, J., Yap, M.G.** 2010. DNA methylation contributes to loss in productivity of monoclonal antibody-producing CHO cell lines. *J. Biotechnol.* **147**: 180-185.
- Yang, C., Bolotin, E., Jiang, T., Sladek, F.M., Martinez, E.** 2007. Prevalence of the Initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATAless core promoters. *Gene.* **389**: 52-65.
- Yudkovsky, N., Ranish, J.A., Hahn, S.** 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature.* **408**: 225-229.
- Zhou, T., Chiang, C.M.** 2001. The intronless and TATA-less human TAFII55 gene contains a functional initiator and a downstream promoter element. *J. Biol. Chem.* **276**: 25503-25511.

Zhu, J. 2012. Mammalian cell protein expression for biopharmaceutical production.

Biotechnol. Adv. **30**: 1158-1170.

Zuhorn, I.S., Bakowsky, U., Polushkin, E., Visser, W.H., Stuart, M.C., Engberts,

J.B., Hoekstra, D. 2005. Nonbilayer phase of lipoplex-membrane mixture determines endosomal escape of genetic cargo and transfection efficiency. *Mol. Ther.* **11**: 801-810.