

Electrophoretic Mobility Shift Assay (EMSA)

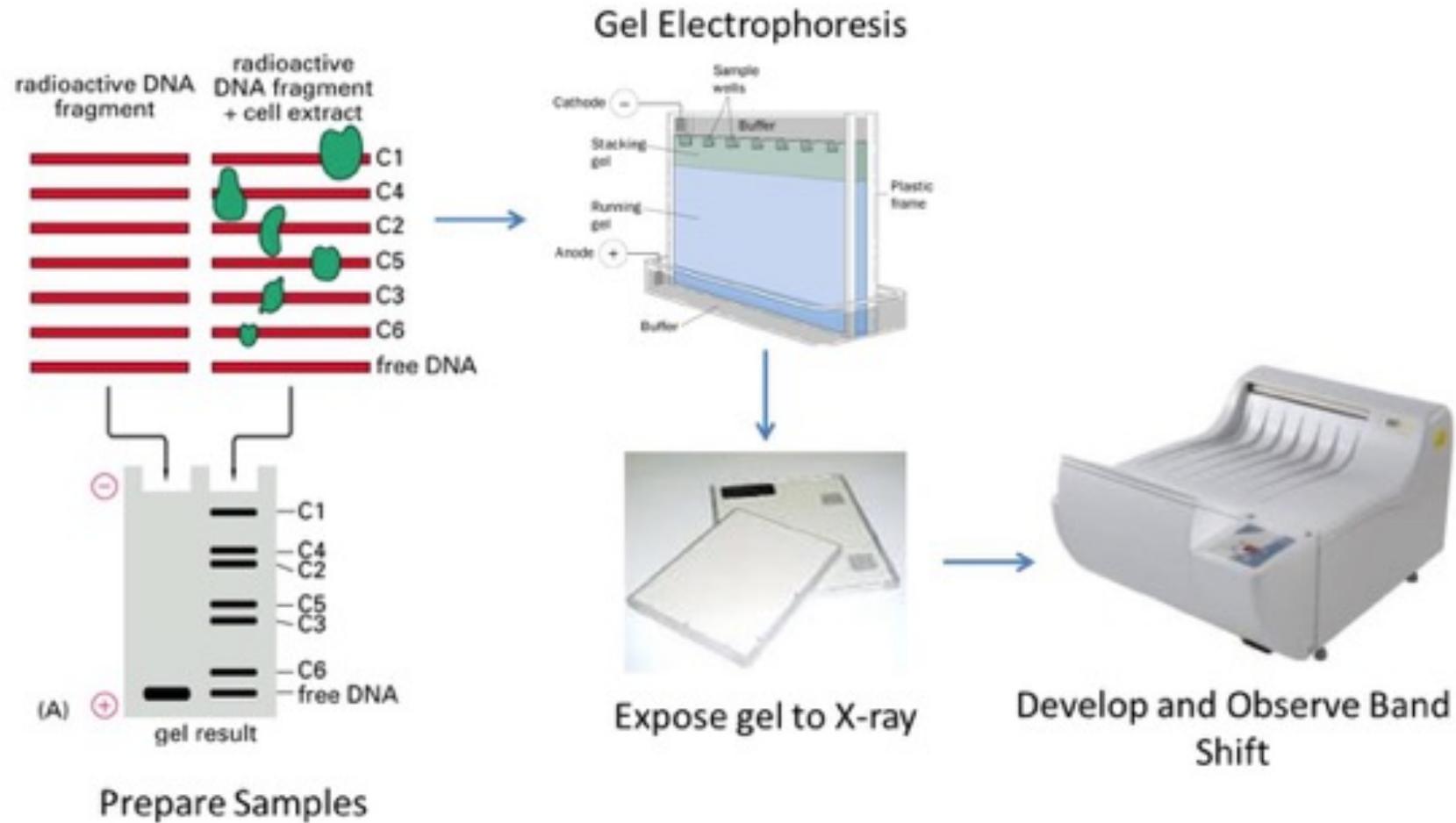
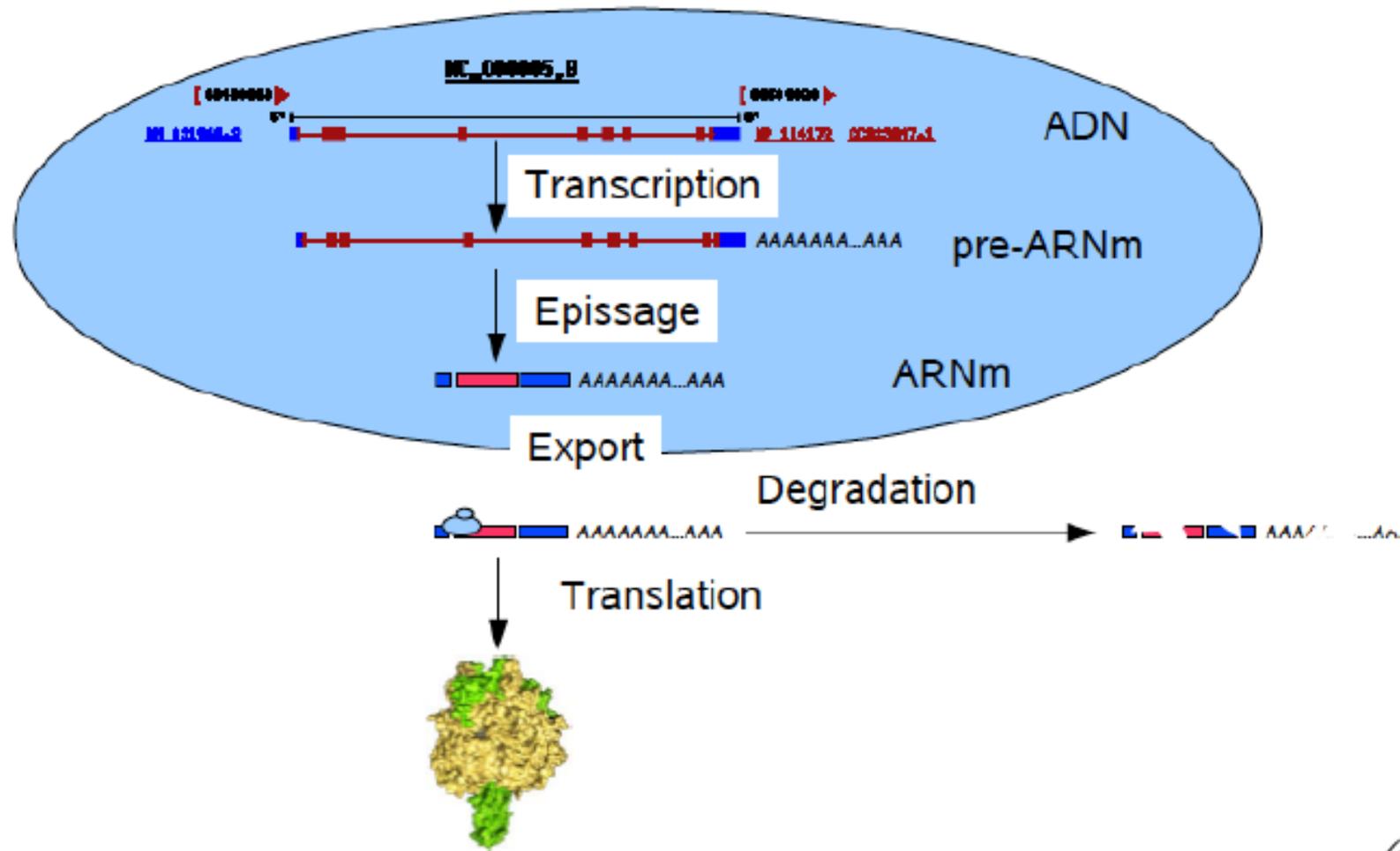
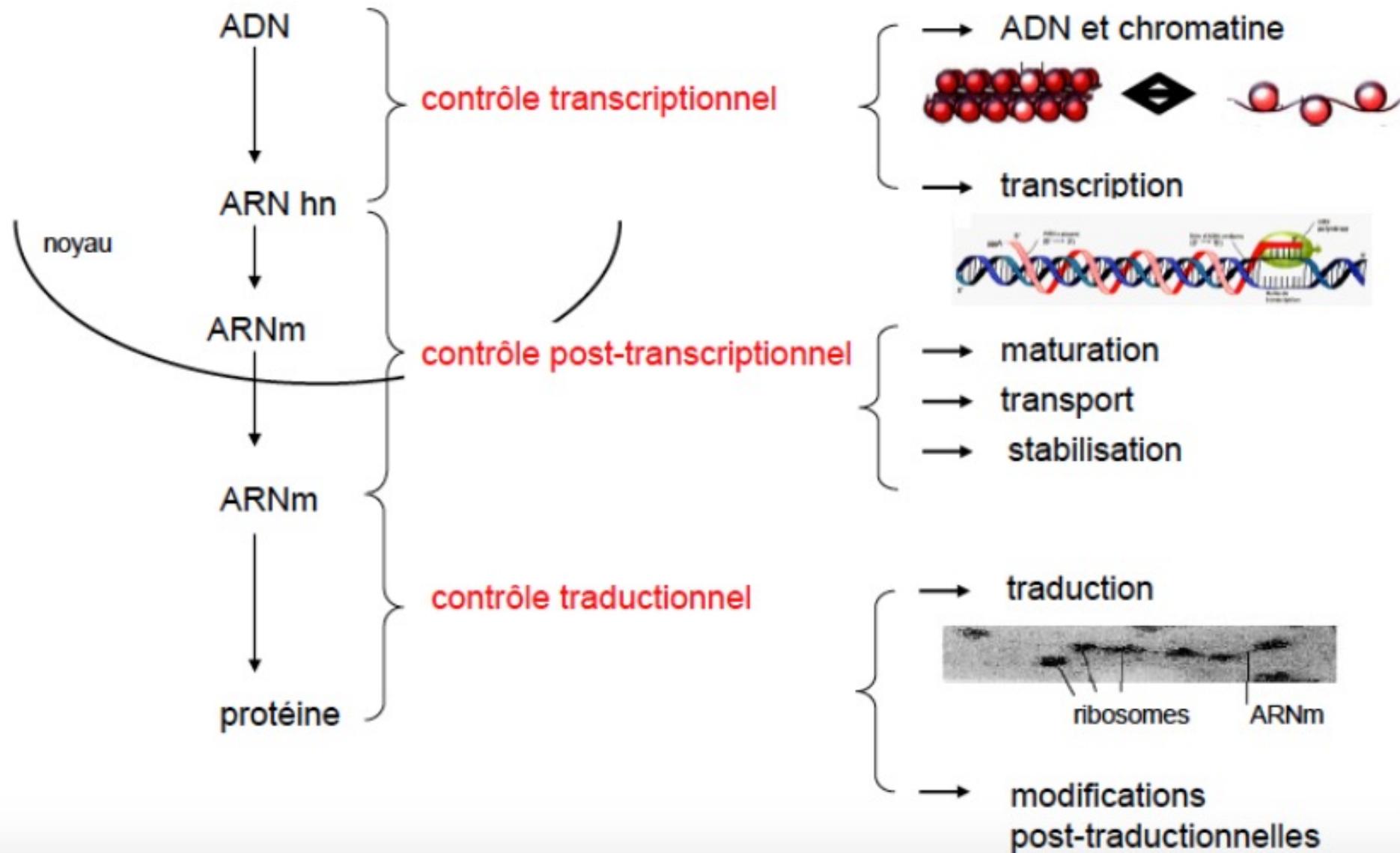


Fig1. EMSA procedure

Du GENE à la PROTEINE

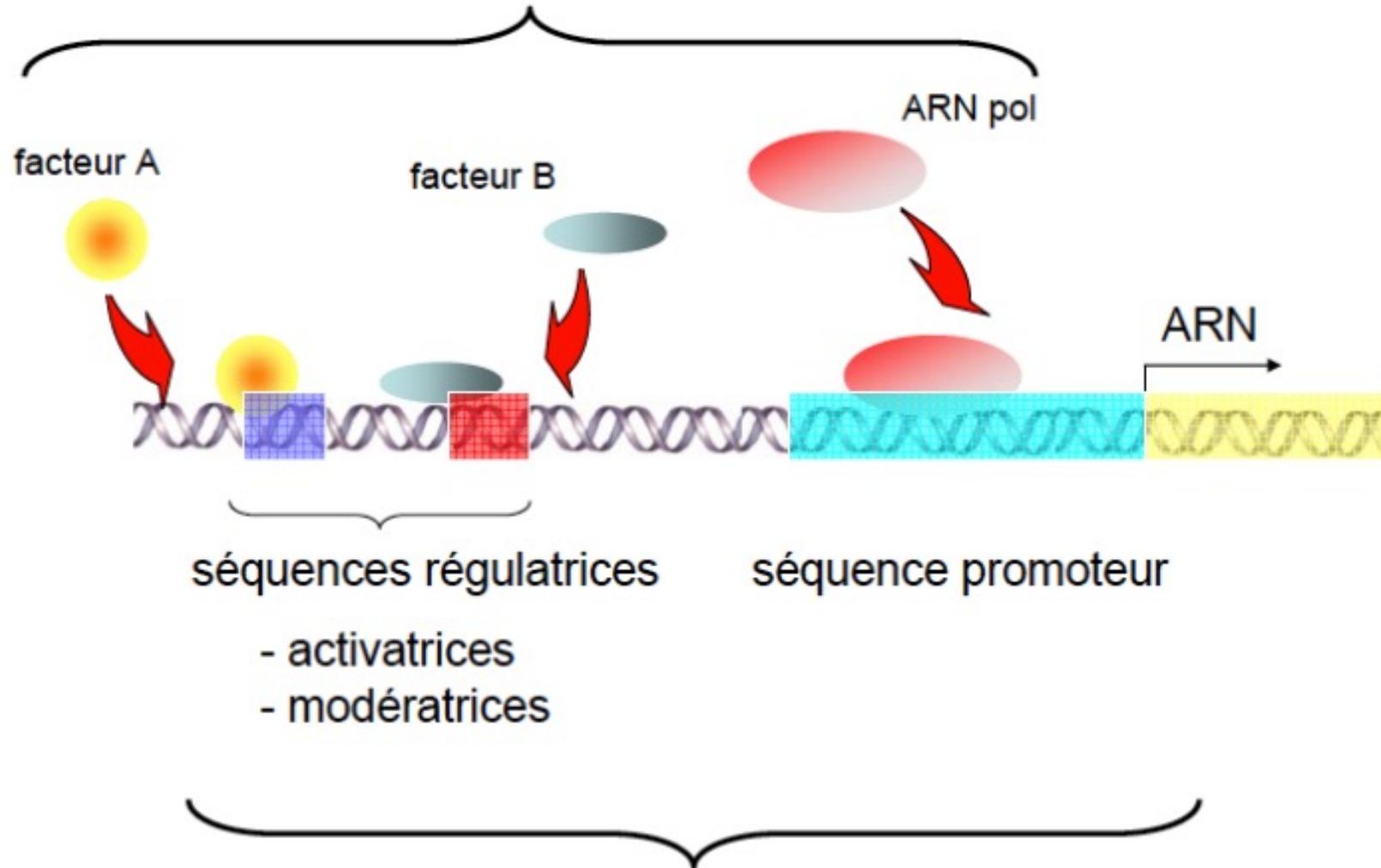


Niveaux de régulation de l'expression



Régulation transcriptionnelle

facteurs de « trans » régulation



séquences de « cis » régulation

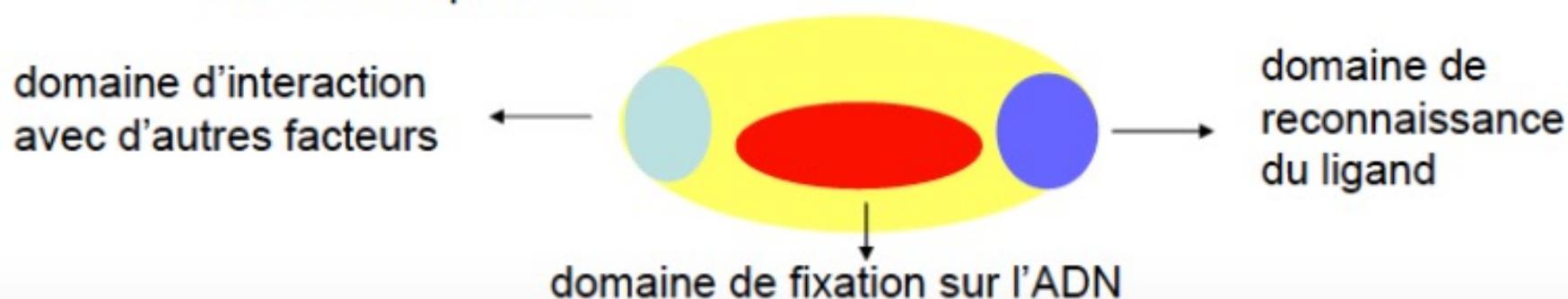
1. Séquences cis régulatrices

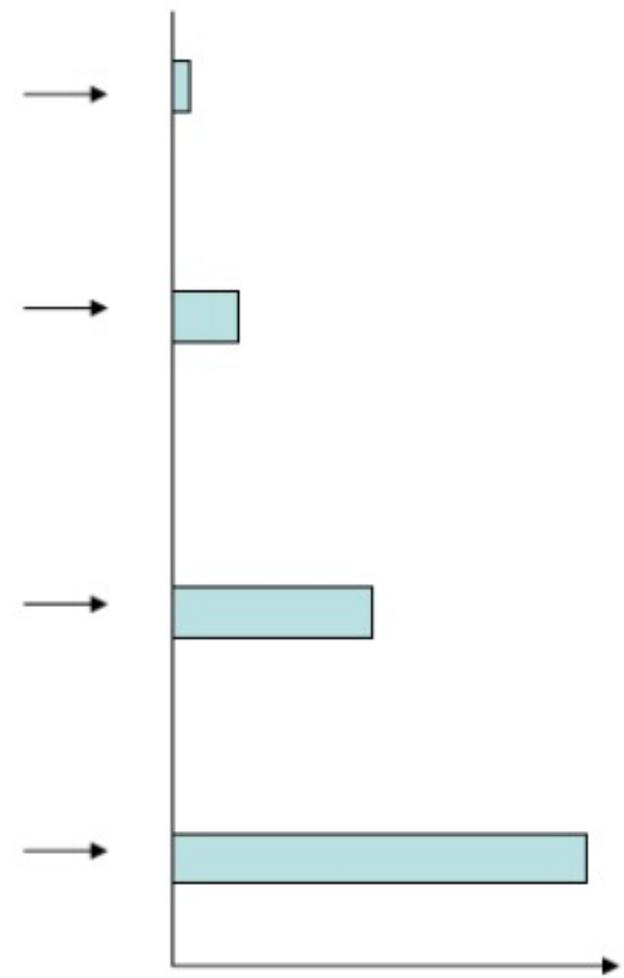
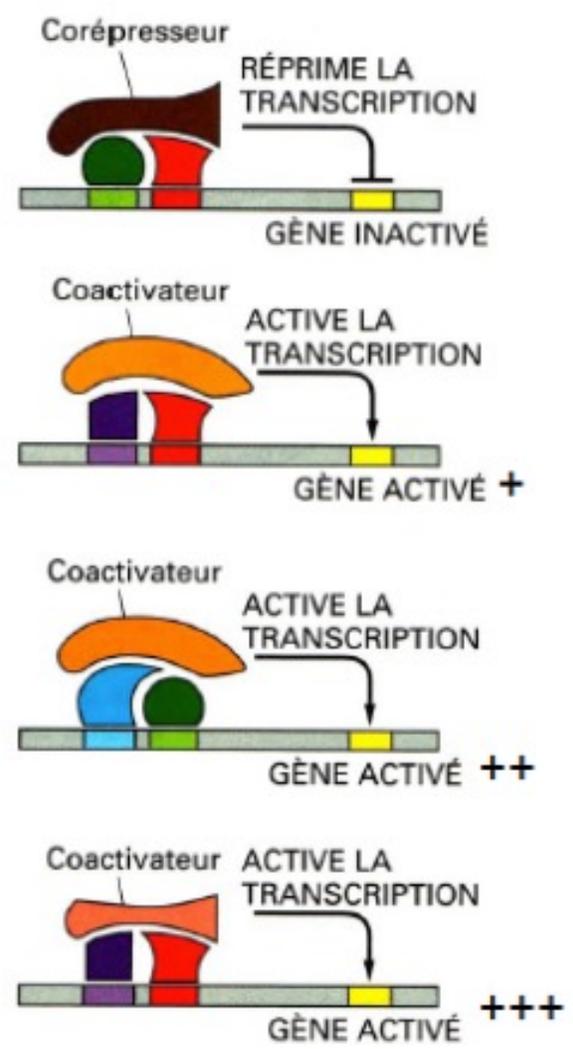
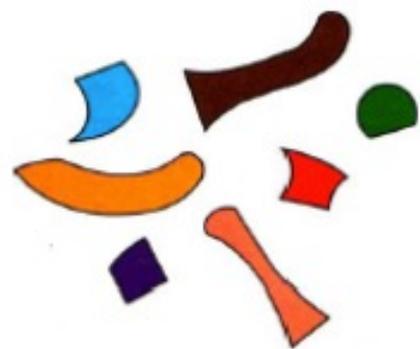
- promoteur : -1 à \approx -100 , motifs (CAAT, TATA...)
- séquences RE : GRE, CRE, IRE
- séquences activatrices ou modératrices :
 - localisation variable
 - nombreuses
- combinaisons

2. Protéines trans régulatrices

- facteurs de transcription
 - généraux
 - spécifiques (tissus, stade de développement)
 - inductibles (phosphorylation, protéolyse, ligands...)

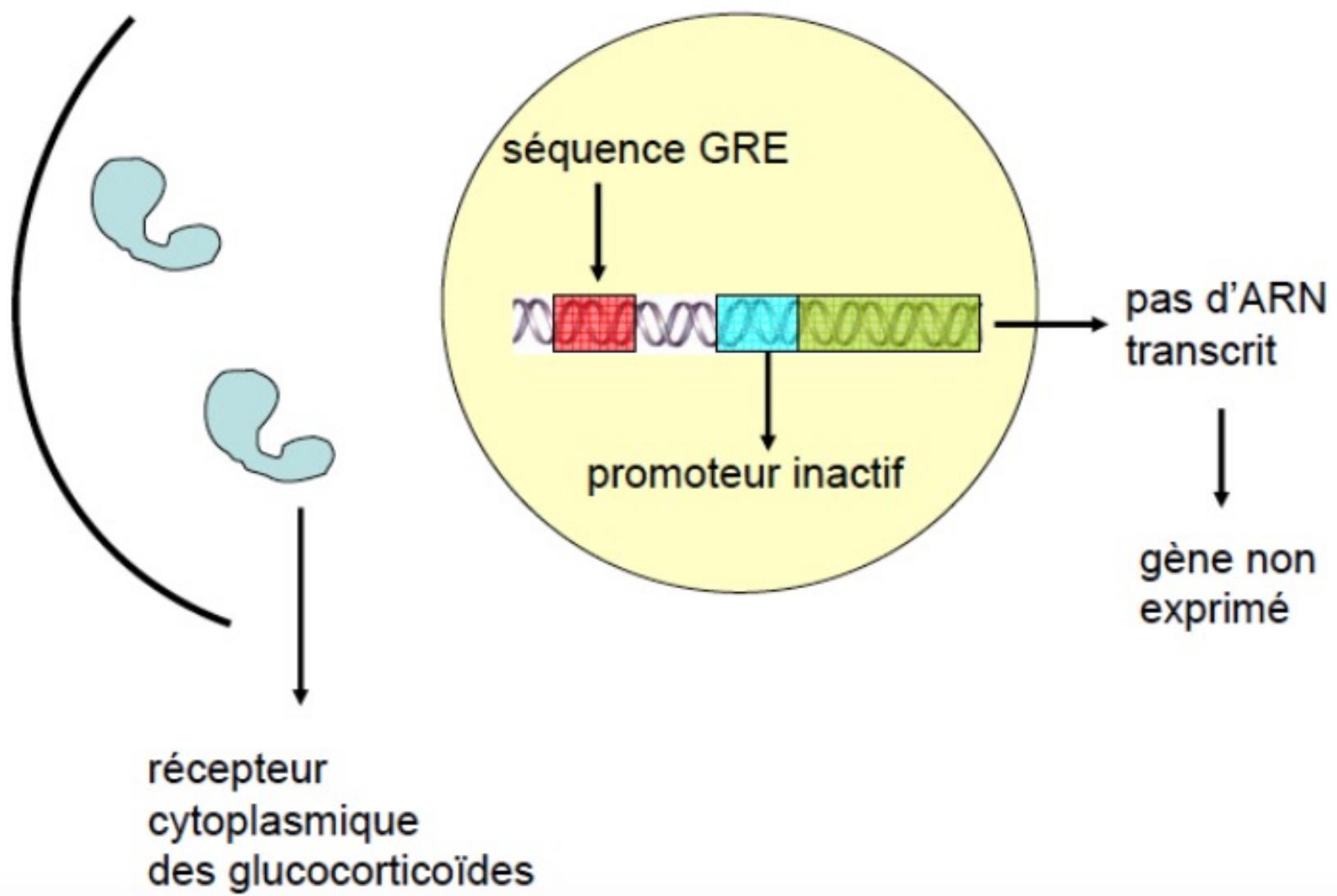
- familles de protéines



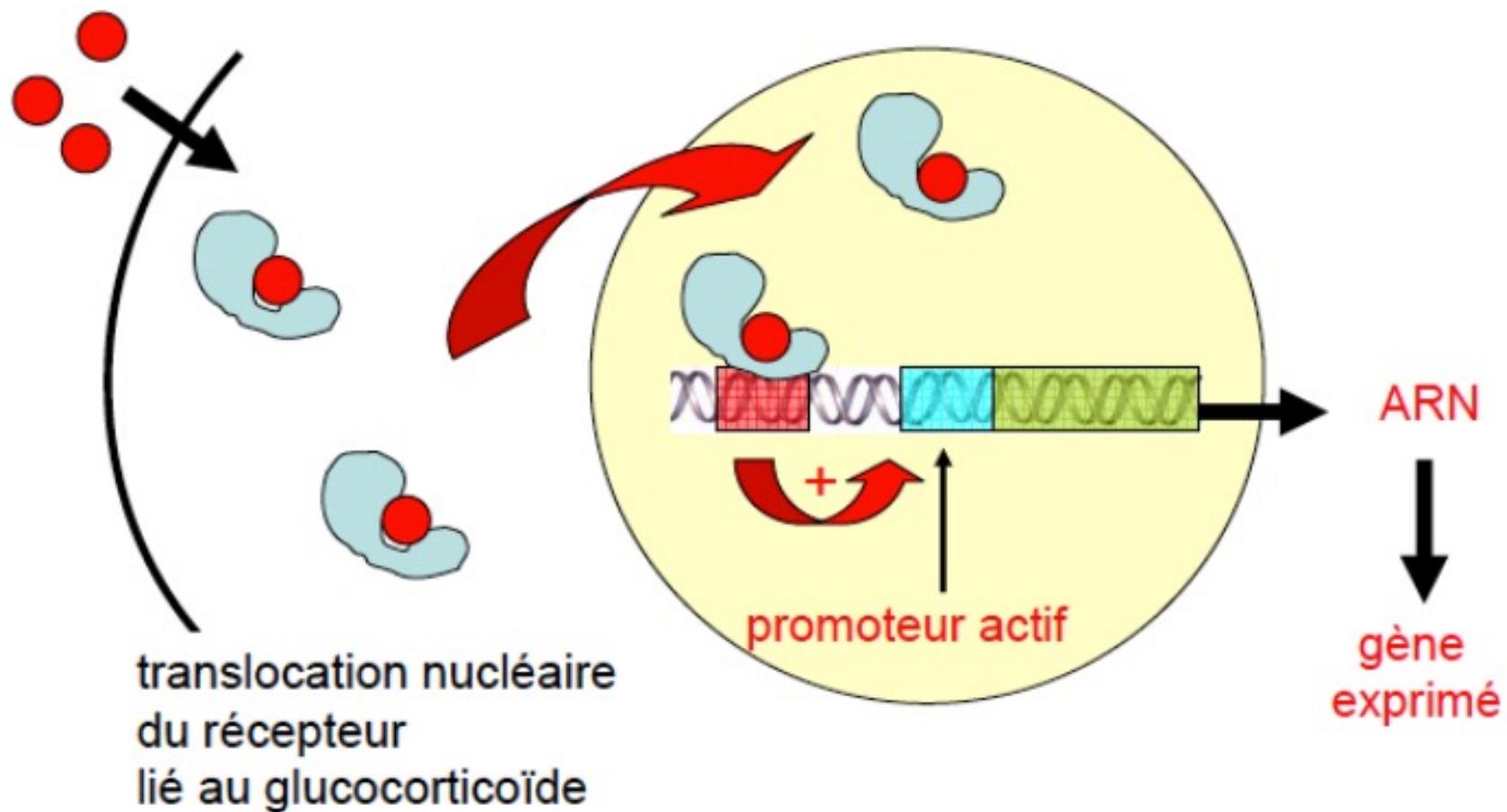


niveau d'expression

➔ exemple d'activation de la transcription de certains gènes par les hormones glucocorticoïdes



Situation en présence d'hormone glucocorticoïde





Genomic Sequence Near Gene

Get Genomic Sequence Near Gene

Note: if you would prefer to get DNA for more than one feature of this track at a time, try the [Table Browser](#) using the output format sequence.

Sequence Retrieval Region Options:

Promoter/Upstream by bases

5' UTR Exons

CDS Exons

3' UTR Exons

Introns

Downstream by bases

One FASTA record per gene.

One FASTA record per region (exon, intron, etc.) with extra bases upstream (5') and extra downstream (3')

Split UTR and CDS parts of an exon into separate FASTA records

Note: if a feature is close to the beginning or end of a chromosome and upstream/downstream bases are added, they may be truncated in order to avoid extending past the edge of the chromosome.

Sequence Formatting Options:

Exons in upper case, everything else in lower case.

CDS in upper case, UTR in lower case.

All upper case.

All lower case.

Mask repeats: to lower case to N

<http://tfbind.hgc.jp>

TFBIND

Please input your sequence in FASTA format, e.g.

> COMMENTS

ACATCTGCTATAAAATACGATGCAGTCACGT

Envoyer

Réinitialiser

or select FASTA format file

Choisir le fichier aucun fichier sél.

Compress result

upload and go

TFBIND : Software for searching transcription factor binding sites (including TATA boxes, GC boxes, CCAAT boxes, transcription start sites (TSS)). This tool uses weight matrix in transcription factor database TRANSFAC R.3.4 developed by Dr. Wingender et al, and the cut-offs originally estimated by our research.

Output format (space-separated):

Column1: Transcription factor matrix ID (from TRANSFAC R.3.4).

Column2: Transcription factor label (from TRANSFAC R.3.4). V means vertebrate.

Column3: Similarity (0.0-1.0) between a registered sequence for the transcription factor binding sites and the input sequence (at the position shown in the next column).

Column4: Position on the input sequence.

Column5: Strandness. + and - means forward and reverse strands that the transcription factor binds, respectively.

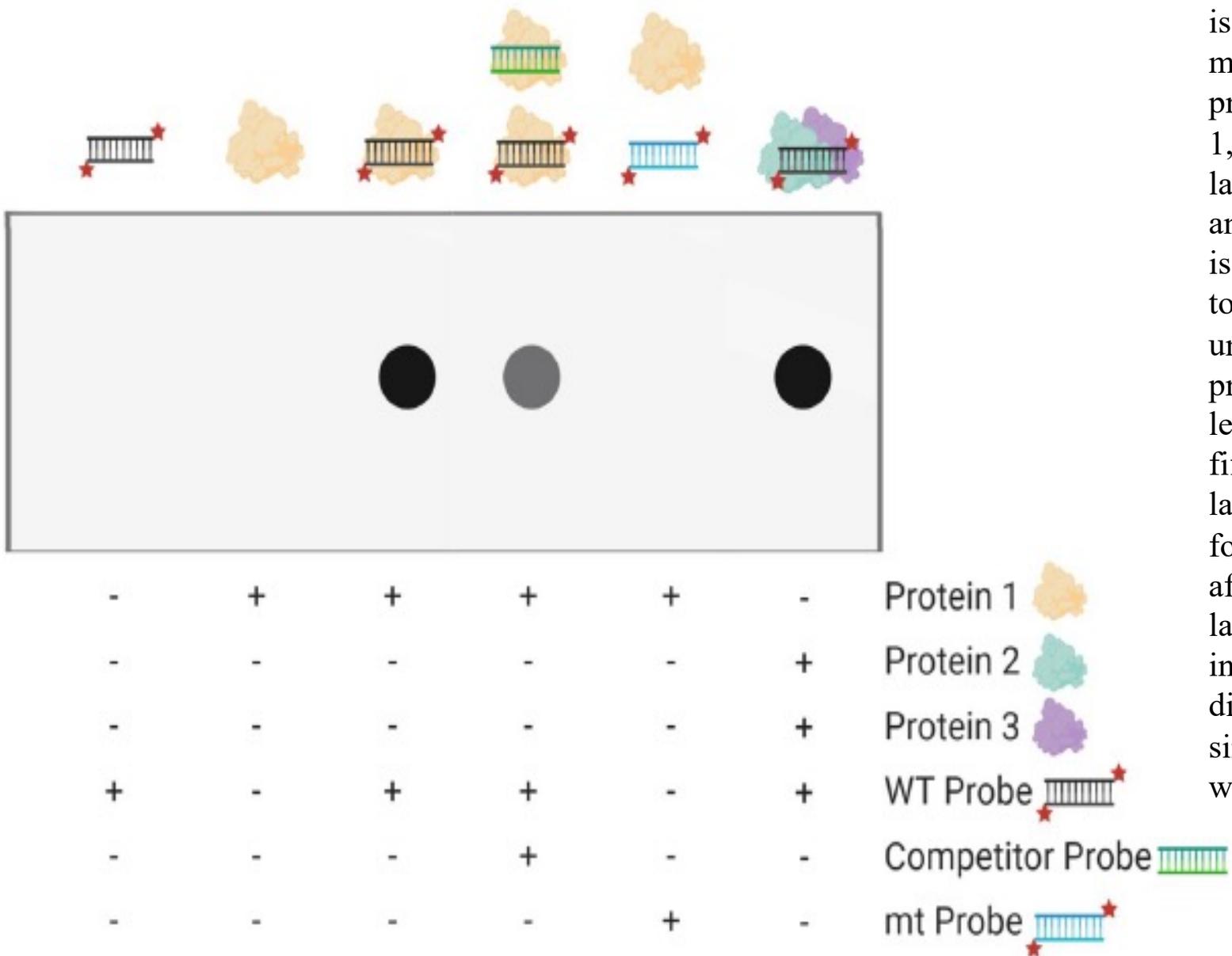
Column6: Consensus sequence (fixed) of the transcription factor binding sites. S = C or G, W = A or T, R = A or G, Y = C or T, K = G or T, M = A or C, N = any base pair.

Column7: Subsequence from the input sequence at the position - corresponding to the consensus sequence.

AC ID Score Loc. Str. Consensus Sequence Signal Sequence

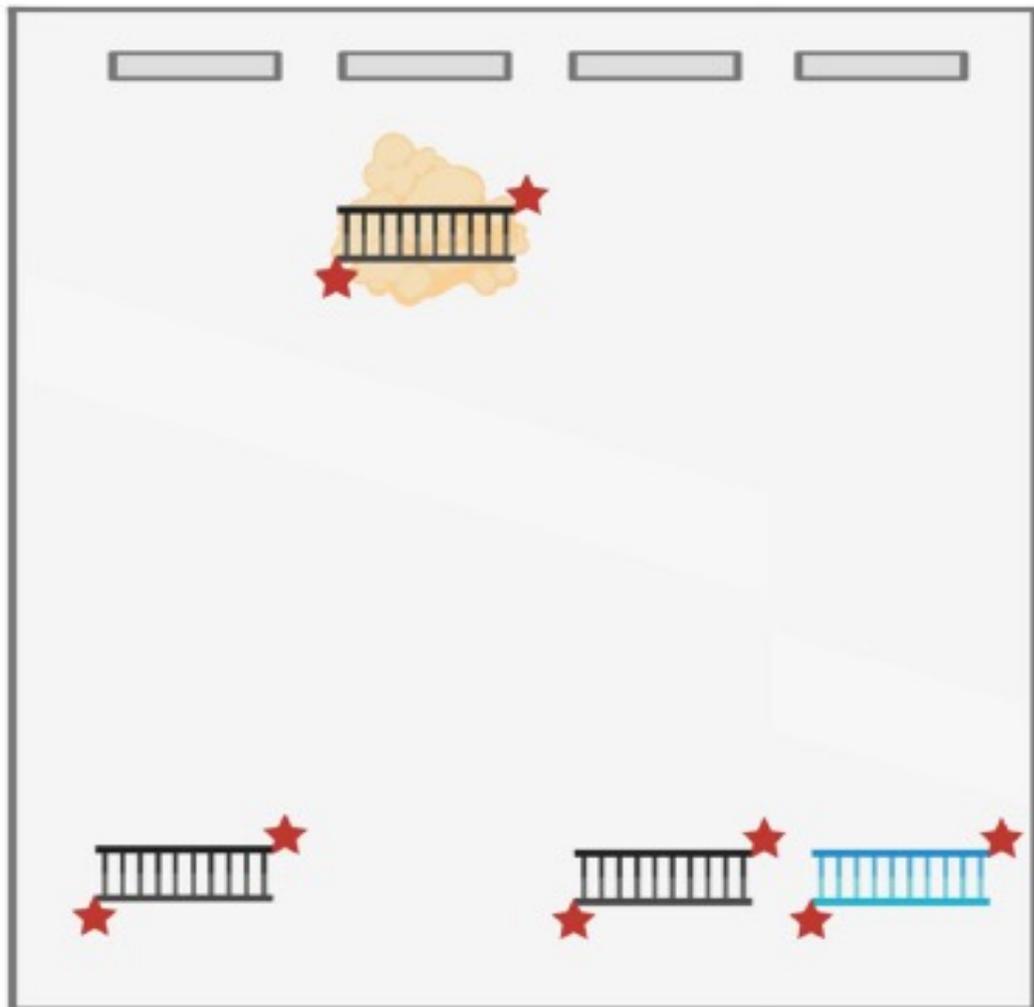
M00155 V\$ARP1_01 0.817071 1 (-) TGARCCYTTGAMCCYW AATTTTCATAGGCTCC
M00248 V\$OCT1_07 0.852873 1 (-) TNTATGNTAATT AATTTTCATAGG
M00185 V\$NFY_Q6 0.782631 5 (-) TRRCCAATSRN TTCATAGGCTC
M00253 V\$CAP_01 0.883687 6 (+) NCANNNNN TCATAGGC
M00254 V\$CAAT_01 0.820672 6 (-) NNNRRCCAATSA TCATAGGCTCCG
M00085 V\$ZID_01 0.773827 10 (+) NGGCTCYATCAYC AGGCTCCGCCCTC
M00255 V\$GC_01 0.964478 10 (-) NRGGGGCGGGGCKN AGGCTCCGCCCTCA
M00196 V\$SP1_Q6 0.914256 11 (-) NGGGGGCGGGGYN GGCTCCGCCCTCA
M00261 V\$OLF1_01 0.774080 11 (+) NNCNANTCCCYNGRGARNNKGN GGCTCCGCCCTCAAGGCCACGT
M00008 V\$SP1_01 0.883635 12 (-) GRGGCRGGGW GCTCCGCCCT
M00261 V\$OLF1_01 0.764552 12 (-) NNCNANTCCCYNGRGARNNKGN GCTCCGCCCTCAAGGCCACGTG
M00262 V\$STAF_01 0.778179 14 (+) NTTWCCCANMATGCAYYRCGNY TCCGCCCTCAAGGCCACGTGAC
M00189 V\$AP2_Q6 0.786075 16 (+) MKCCCSCNGGCG CGCCCTCAAGGC
M00139 V\$AHR_01 0.816144 19 (+) CYYCNRRSTNGCGTGASW CCTCAAGGCCACGTGACT
M00098 V\$PAX2_01 0.790741 21 (-) NNNNGTCANGNRTKANNNN TCAAGGCCACGTGACTGAT
M00191 V\$ER_Q6 0.740494 21 (+) NNARGNANNNTGACCYNN TCAAGGCCACGTGACTGAT
M00251 V\$XBP1_01 0.846581 22 (-) NNGNTGACGTGKNNNWT CAAGGCCACGTGACTGA
M00002 V\$E47_01 0.770791 23 (-) NSNGCAGGTGKNCNN AAGGCCACGTGACTG
M00211 V\$PADS_C 0.846018 23 (-) NGTGGTCTC AAGGCCACG
M00236 V\$ARNT_01 0.893901 23 (+) NNNNNCACGTGNNNNN AAGGCCACGTGACTGA
M00236 V\$ARNT_01 0.854545 23 (-) NNNNNCACGTGNNNNN AAGGCCACGTGACTGA
M00251 V\$XBP1_01 0.787821 23 (+) NNGNTGACGTGKNNNWT AAGGCCACGTGACTGAT
M00017 V\$ATF_01 0.778127 24 (-) CNSTGACGTNNNYC AGGCCACGTGACTG
M00118 V\$MYCMAX_01 0.862123 24 (+) NNACCACGTGGTNN AGGCCACGTGACTG
M00118 V\$MYCMAX_01 0.862123 24 (-) NNACCACGTGGTNN AGGCCACGTGACTG
M00119 V\$MAX_01 0.928096 24 (+) NNANCACGTGNTNN AGGCCACGTGACTG
M00119 V\$MAX_01 0.928096 24 (-) NNANCACGTGNTNN AGGCCACGTGACTG
M00121 V\$USF_01 0.949868 24 (+) NNRYCACGTGRYNN AGGCCACGTGACTG
M00121 V\$USF_01 0.949868 24 (-) NNRYCACGTGRYNN AGGCCACGTGACTG
M00122 V\$USF_02 0.947483 24 (+) NNRNCACGTGNYNN AGGCCACGTGACTG
M00122 V\$USF_02 0.947483 24 (-) NNRNCACGTGNYNN AGGCCACGTGACTG
M00001 V\$MYOD_01 0.790157 25 (-) SRACAGGTGKYG GGCCACGTGACT
M00055 V\$NMYC_01 0.904805 25 (+) NNNCACGTGNNN GGCCACGTGACT
M00055 V\$NMYC_01 0.889903 25 (-) NNNCACGTGNNN GGCCACGTGACT
M00123 V\$MYCMAX_02 0.884686 25 (+) NANCACGTGNNW GGCCACGTGACT
M00123 V\$MYCMAX_02 0.850538 25 (-) NANCACGTGNNW GGCCACGTGACT

Electrophoretic Mobility Shift Assay (EMSA)

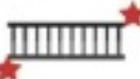


Representation of the expected results of a filter binding assay visualised in a nitrocellulose membrane. While in the first two lanes, no signal is detected, since only one of the complex molecules is present, labelled nucleic acid (WT probe, first lane) and protein of interest (protein 1, second

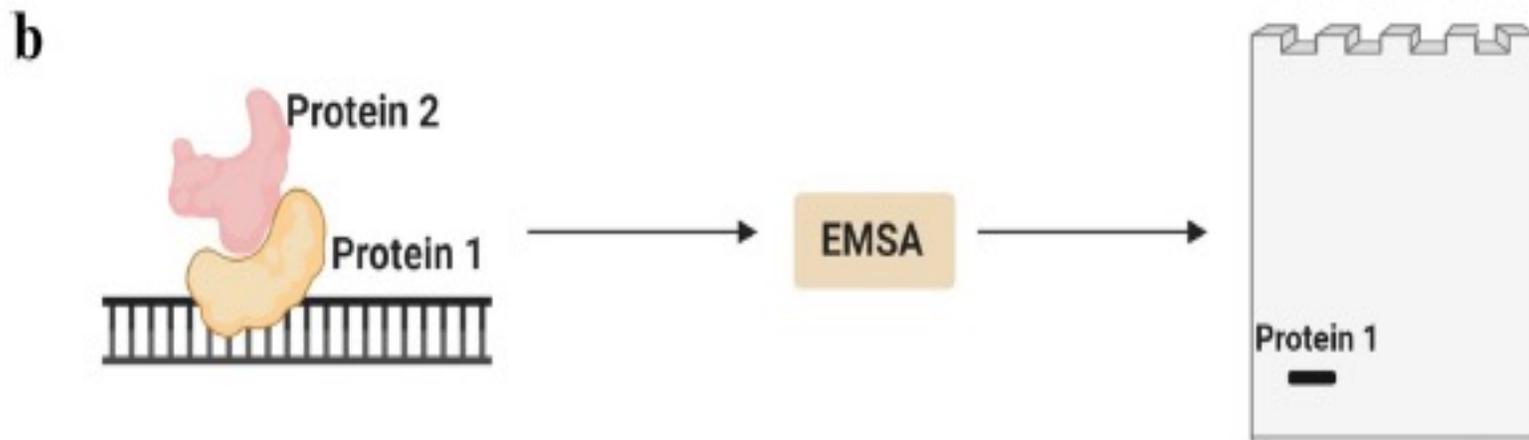
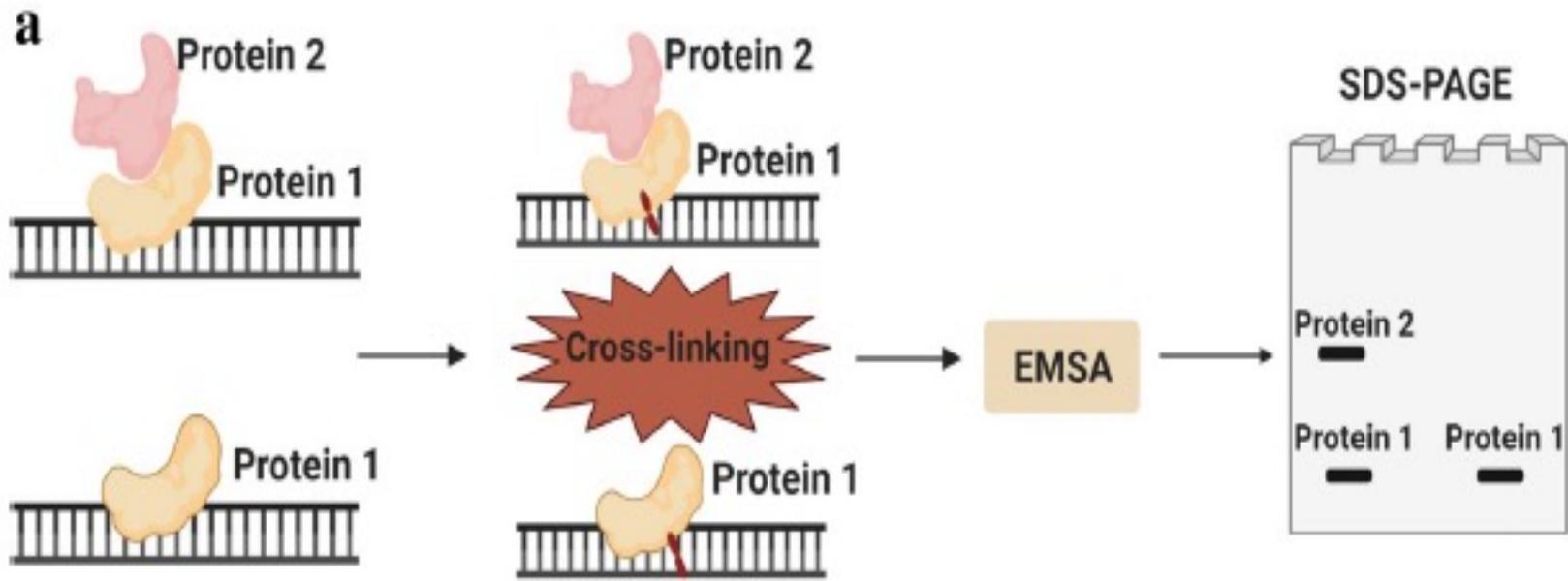
lane), in the third lane, the complex is formed and the signal is detected, since the protein, that is attached to nitrocellulose membrane, is linked to a labelled nucleic acid. In the fourth lane, an unlabelled competitor probe deviates some proteins to the formation of a different complex, leading to the weakening of the signal. In the fifth lane, the substitution of the WT probe by a labelled mutated fragment (mt probe) enables the formation of the complex, in case the mutation affects the complex formation. Finally, the last lane represents the formation of a new complex involving the WT labelled probe and two different proteins, leading to a signal quite similar to the one observed in lane 3. Created with BioRender.com



-	+	+	+
+	+	+	-
-	-	+	-
-	-	-	+

Protein 
 WT Probe 
 Competitor Probe 
 mt Probe 

Schematic representation of the expected results of an EMSA visualised in an electrophoretic gel. While the first lane presents a band with the size of the segment being tested (WT probe), in the second lane, the introduction of the binding protein and formation of a nucleic acid–protein complex delays the run of this fragment. As for the third lane, the presence of an unlabelled competitor probe leads to the formation of a complex containing this competitor probe and the protein of interest and, consequently, the band that corresponds to the free labelled WT probe appears again. Finally, in the last lane, the replacement of the WT probe by a labelled mutated fragment (mt probe) may cause a similar result to the one of the first lane, in case the mutation affects the complex formation.
 Created with BioRender.com



Comparison between the cross-linking procedure (a) and South-Western Blotting (b). In the cross-linking technique (a) a nucleic acid is cross-linked to the respective binding proteins, ensuring that all proteins present in the complex are detected in a SDS-PAGE performed after an EMSA (that isolates the complex under study), independently of being in direct contact with the nucleic acid (Protein 1) or not (Protein 2). As for South-Western Blotting (b), after an EMSA, the complex is subjected to a SDS-PAGE without being exposed to a cross-linker and, consequently, only the proteins in direct contact with the nucleic acid will be detected. Created with BioRender.com

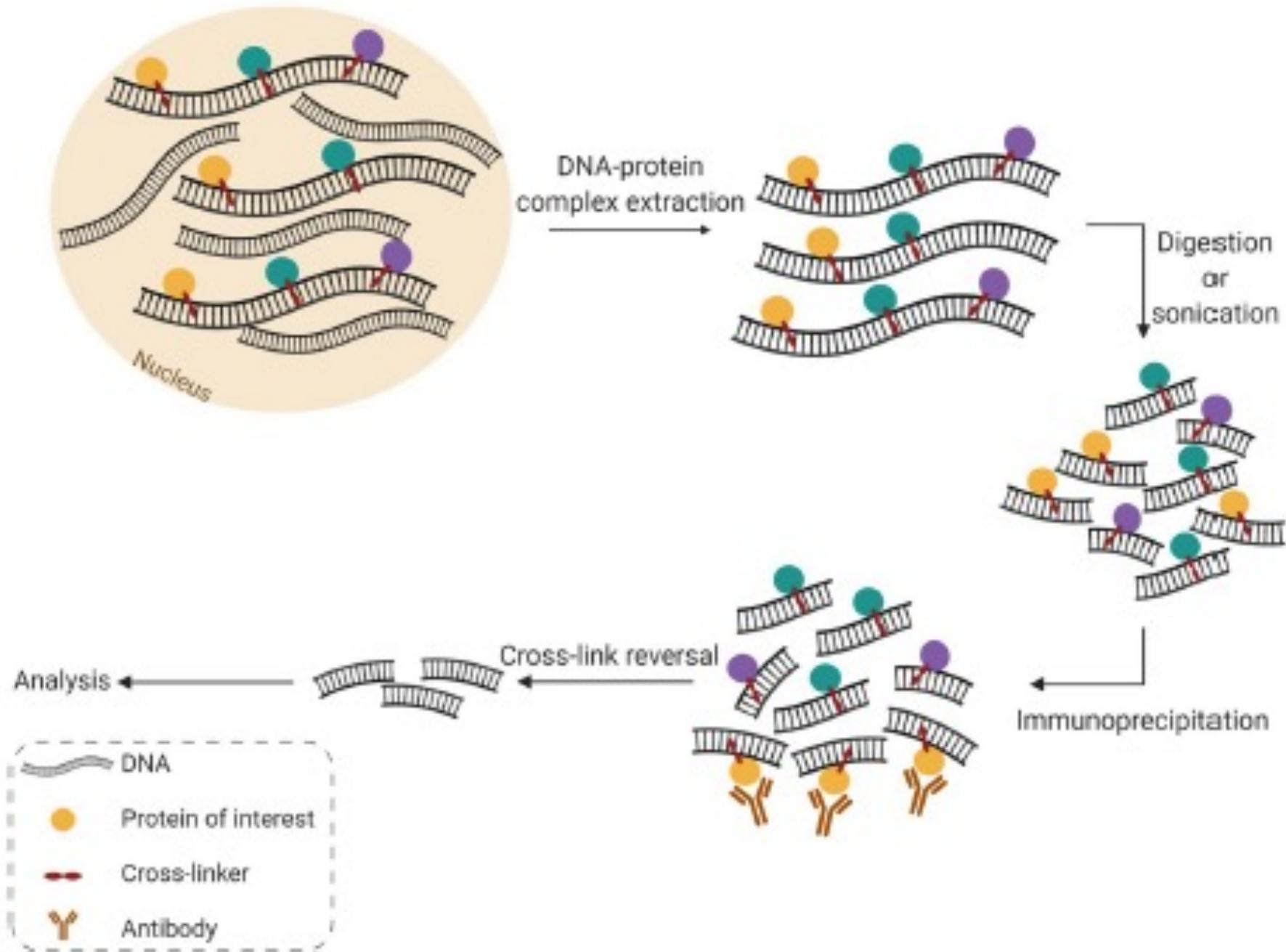
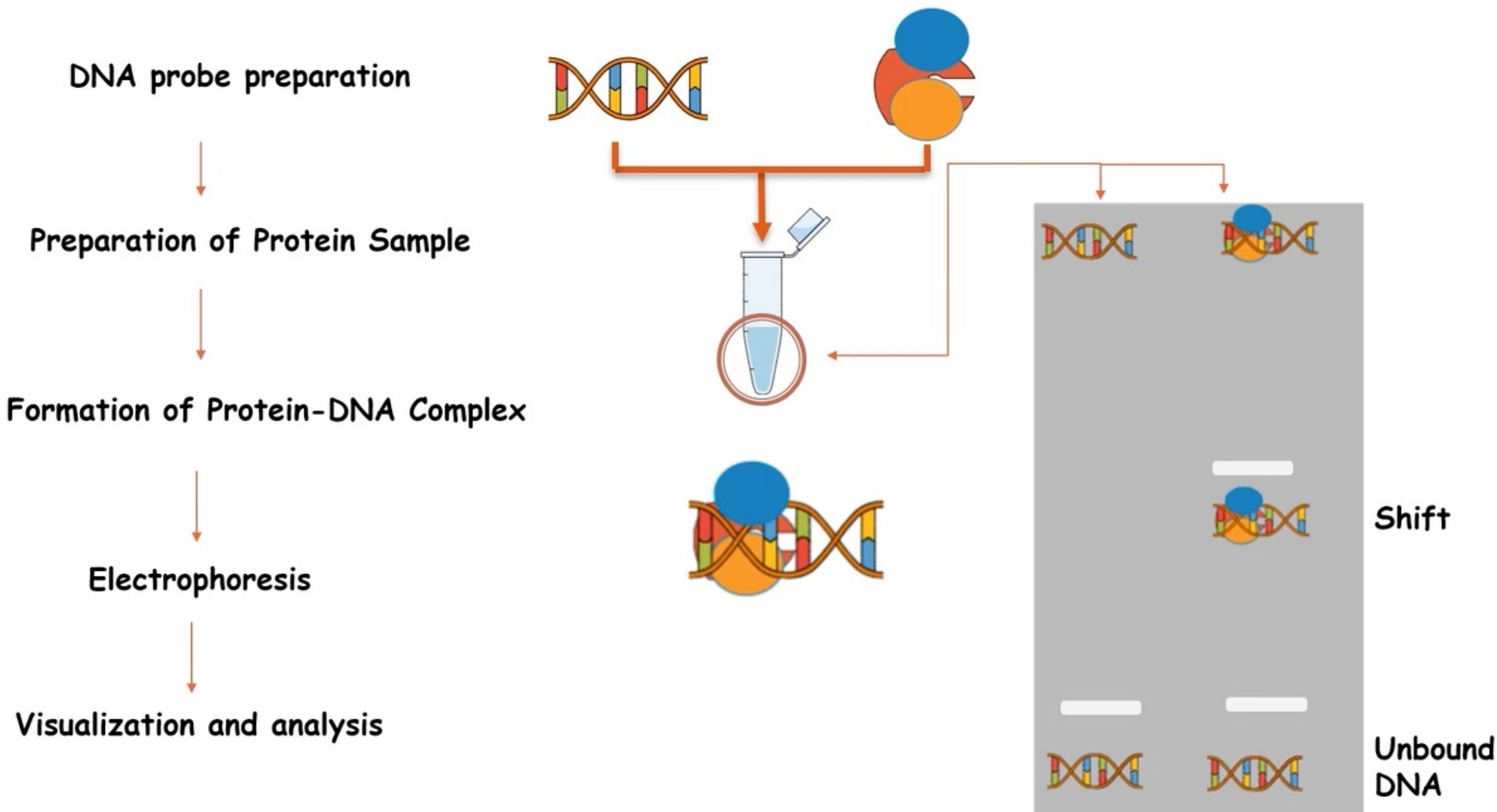
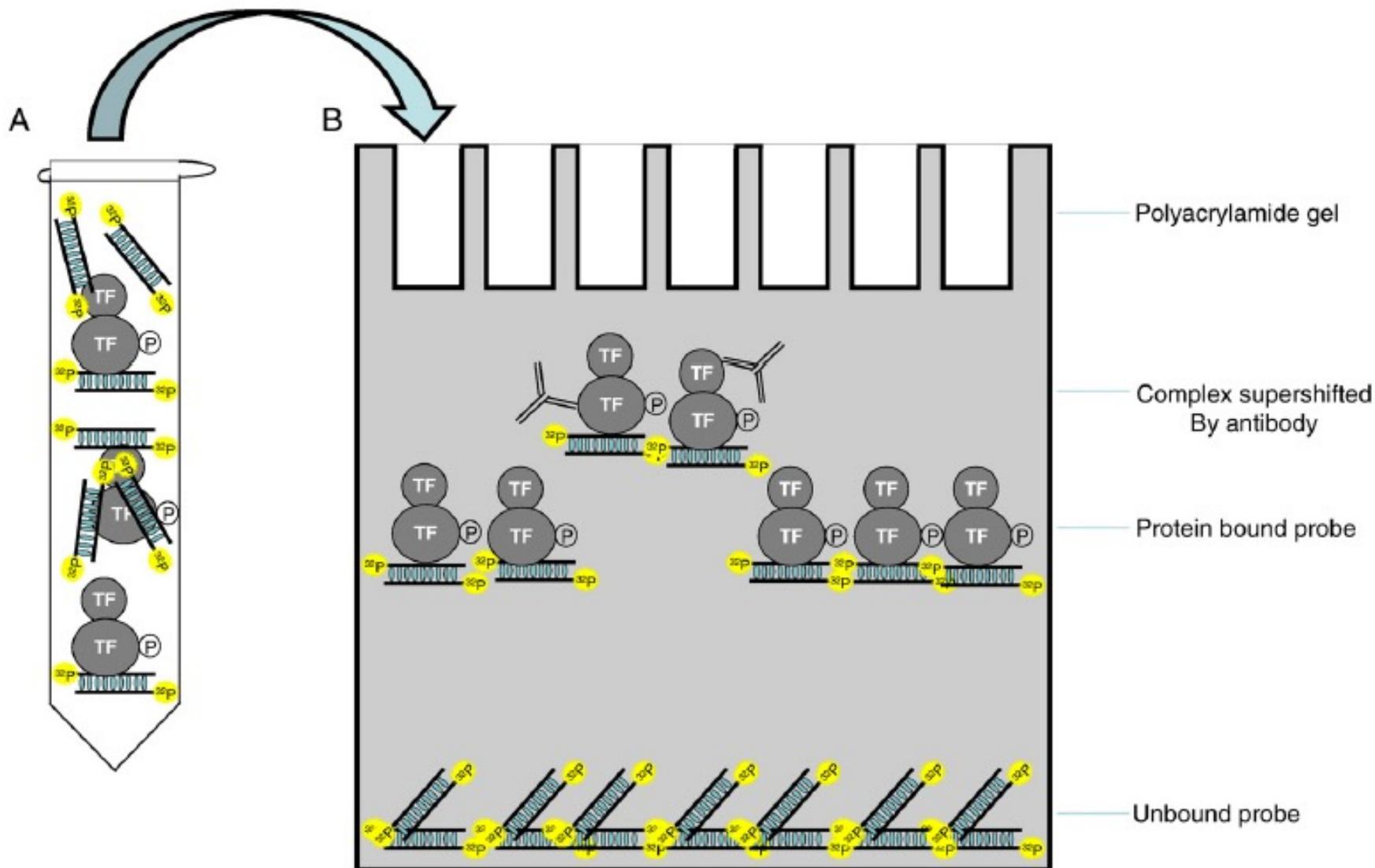


Illustration of the ChIP procedure. ChIP techniques usually start by the cross-link between DNA and proteins in a living cells using chemical cross-linkers or UV light. Subsequently, cross-linked chromatin is extracted and fragmented by digestion or sonication. The DNA-protein complexes under study are then immunoprecipitated using a specific antibody to the protein of interest and the obtained DNA is purified, released from the cross-link and analysed.





Principle of the EMSA binding reaction and polyacrylamide gel electrophoresis. A schematic diagram of A) the EMSA binding reaction, where ^{32}P labelled DNA is incubated with extracted nuclear protein. and B) the separation of protein bound probe and the unbound free probe on a non-denaturing acrylamide gel. Where TF denotes a generic dimeric transcription factor such as NF- κ B or AP-1.

I- Protéine recombinante

Les protéines recombinantes sont des protéines étrangères produites dans des hôtes d'expression. Ces protéines recombinantes sont utilisées comme réactifs de diagnostic médical dans le domaine de la santé humaine, tels que les vaccins, les médicaments ou les anticorps, ainsi que dans les analyses biochimiques.

Le processus de **production de protéines recombinantes** dépend fortement du processus d'insertion du segment d'ADN dans le génome de l'hôte. Deux méthodes sont connues pour produire de l'ADN recombinant, à savoir le **clonage moléculaire** et la **PCR**.

★ Le clonage moléculaire est un processus utilisé pour insérer un ADN recombinant dans un vecteur (véhicule utilisé pour transférer du matériel génétique) qui répliquera le segment à l'intérieur de l'organisme hôte. Le segment d'ADN peut être isolé d'un procaryote ou d'un eucaryote portant le gène ou la protéine d'intérêt. Après l'isolement du segment d'ADN, il y a sa coupe (avec des enzymes de restriction) et son insertion dans le vecteur plasmidique par ligature. Le produit résultant est maintenant appelé un plasmide recombinant et est maintenant prêt à être inséré dans l'hôte (ex. : la bactérie E. coli). Le clonage moléculaire fonctionne essentiellement sur le principe que toute séquence d'ADN insérée dans le plasmide sera répliquée avec le reste de l'ADN plasmidique.

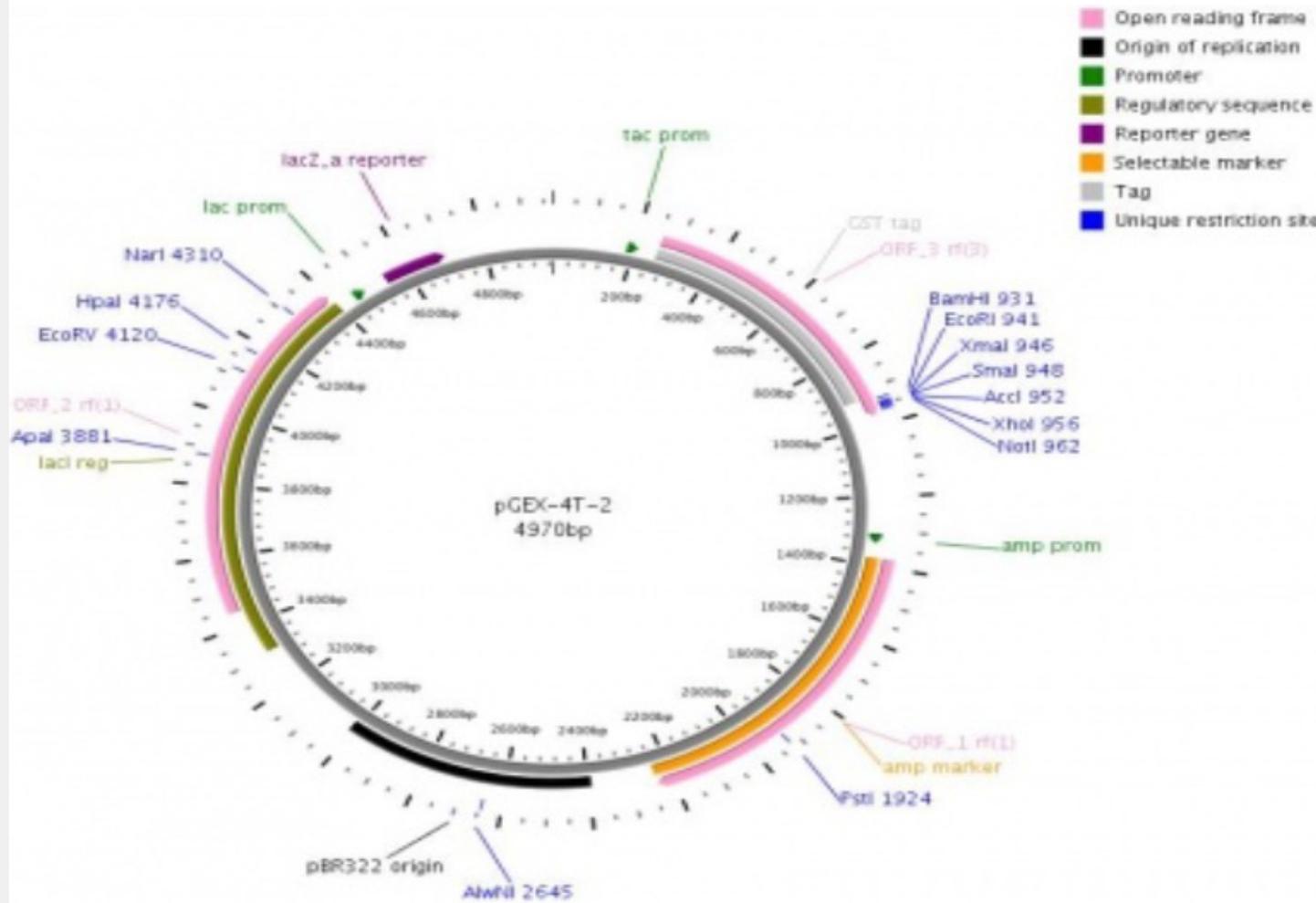
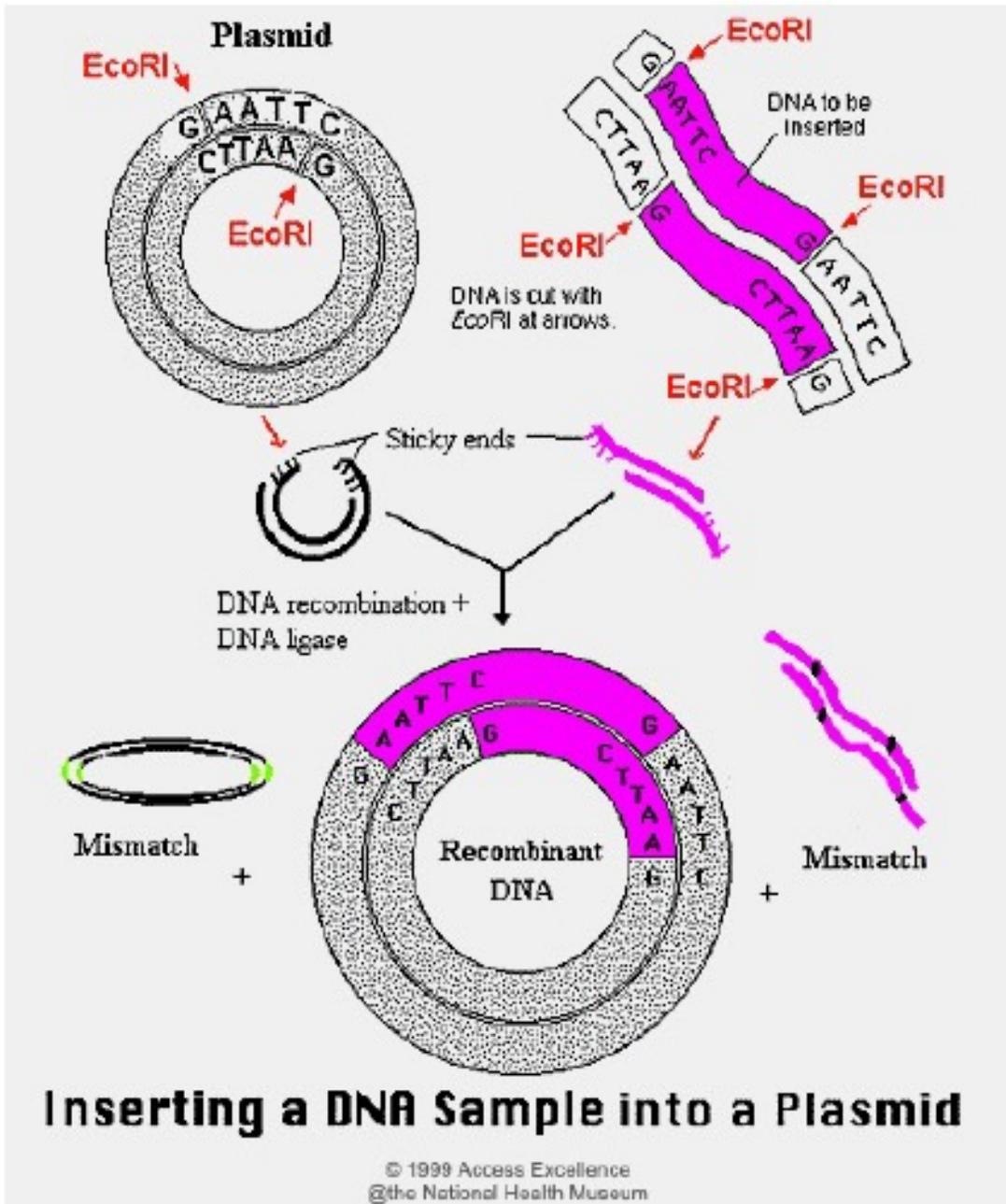
Une protéine recombinante est une protéine produite en utilisant la technologie de l'ADN recombinant. L'ADN recombinant est une technique de biologie moléculaire in vitro qui permet de contrôler à l'acide aminé près la séquence nucléotidique de la protéine que l'on veut produire.

On peut également combiner des séquences d'ADN provenant de différentes sources, provenant potentiellement d'espèces différentes, pour créer de nouveaux gènes composites, dits chimériques et des protéines correspondantes.

La production de protéines recombinantes implique plusieurs étapes. Tout d'abord, le gène codant pour la protéine ou le peptide d'intérêt est isolé à partir de l'organisme source et cloné dans un vecteur d'ADN recombinant. Ce nom de « vecteur » vient du fait qu'il est une sorte de véhicule nucléotidique permettant de transporter l'information génétique à l'intérieur d'une cellule « usine » hôte. Le vecteur est généralement un plasmide, de l'ADN circulaire qui peut être répliqué dans une cellule hôte. Ensuite, le vecteur d'ADN recombinant contenant le gène d'intérêt est introduit dans une cellule hôte, telle que des bactéries, des levures ou des cellules de mammifères. Une fois à l'intérieur de la cellule hôte, le vecteur est répliqué et transcrit en ARN messager (ARNm).

Une protéine recombinante est une protéine produite en utilisant la technologie de l'ADN recombinant. L'ADN recombinant est une technique de biologie moléculaire *in vitro* qui permet de contrôler à l'acide aminé près la séquence nucléotidique de la protéine que l'on veut produire. On peut également combiner des séquences d'ADN provenant de différentes sources, provenant potentiellement d'espèces différentes, pour créer de nouveaux gènes composites, dits chimériques et des protéines correspondantes.

La production de protéines recombinantes implique plusieurs étapes. Tout d'abord, le gène codant pour la protéine ou le peptide d'intérêt est isolé à partir de l'organisme source et cloné dans un vecteur d'ADN recombinant. Ce nom de « vecteur » vient du fait qu'il est une sorte de véhicule nucléotidique permettant de transporter l'information génétique à l'intérieur d'une cellule « usine » hôte. Le vecteur est généralement un plasmide, de l'ADN circulaire qui peut être répliqué dans une cellule hôte. Ensuite, le vecteur d'ADN recombinant contenant le gène d'intérêt est introduit dans une cellule hôte, telle que des bactéries, des levures ou des cellules de mammifères. Une fois à l'intérieur de la cellule hôte, le vecteur est répliqué et transcrit en ARN messenger (ARNm).



ENZYME	RECOGNITION SITE	ENZYME	RECOGNITION SITE	ENZYME	RECOGNITION SITE
<i>Aat II</i>	GAOGI▼C	<i>BsrBR I</i>	GATNN▼NNATC	<i>Fok I</i>	GGATG(9/13)
<i>Acc I</i>	GT▼(A/T)(T/G)AC	<i>BsrS I</i>	ACTGGN▼	<i>Hae II</i>	(A/G)GCGC▼ (T/C)
<i>AccII</i>	T▼OCGGA	<i>BssH II</i>	G▼OGCGC	<i>HaeIII</i>	GG▼OC
<i>Acc65 I</i>	G▼GTAOC	<i>Bst71 I</i>	GCAGC(8/12)	<i>Hha I</i>	GOG▼C
<i>AccB7 I</i>	OCANNNN▼NTGG	<i>Bst98 I</i>	C▼TTAAG	<i>Hinc II</i>	GT(T/C)▼(A/G)AC
<i>Acy I</i>	G(A/G)▼OG(T/C)C	<i>Bst E II</i>	G▼GTNAOC	<i>Hind III</i>	A▼AGCTT
<i>Age I</i>	A▼OCGGT	<i>Bst O I</i>	OC▼(A/T)GG	<i>Hinf I</i>	G▼ANTC
<i>Alu I</i>	AG▼CT	<i>Bst X I</i>	OCANNNN▼ NTGG	<i>Hpa I</i>	GTT▼AAC
<i>A/w26 I</i>	G▼TCTC(1/5)	<i>Bst Z I</i>	C▼GGOCG	<i>Hpa II</i>	C▼OGG
<i>A/w44 I</i>	G▼TGCAC	<i>Bsu36 I</i>	OC▼TNAGG	<i>Hsp92 I</i>	G(A/G)▼ OG(T/C)C
<i>Apa I</i>	GGGOC▼C	<i>Cfo I</i>	GOG▼C	<i>Hsp92 II</i>	CATG▼
<i>Ava I</i>	C▼(T/C)CG(A/G)G	<i>Cla I</i>	AT▼OGAT	<i>I-Ppo I</i>	CTCTCTAA▼ GGTAGC
<i>Ava II</i>	G▼G(A/T)OC	<i>Csp I</i>	OG▼G(A/T)OCG	<i>Kpn I</i>	GGTAC▼C
<i>Ba I</i>	TGG▼OCA	<i>Csp 45 I</i>	TT▼OGAA	<i>Mbo I</i>	▼GATC
<i>BamH I</i>	G▼GATOC	<i>Dde I</i>	C▼TNAG	<i>Mbo II</i>	GAAGA(8/7)
<i>Ban I</i>	G▼G(T/C)(A/G)OC	<i>Dpn I</i>	G▼A▼TC	<i>Mku I</i>	A▼OGCGT
<i>Ban II</i>	G(A/G)GC(T/C)▼C	<i>Dra I</i>	TTT▼AAA	<i>Msp I</i>	C▼OGG
<i>Bbu I</i>	GCA TG▼C	<i>EdHK I</i>	GACNNN▼NNGTC	<i>MspA I</i>	C(A/C)G▼ C(G/T)G
<i>Bcl I</i>	T▼GATCA	<i>Eco47 III</i>	AOG▼GCT	<i>Nae I</i>	GOC▼GGC
<i>Bgl I</i>	GOCNNN▼NGGC	<i>Eco52 I</i>	C▼GGOCG	<i>Nar</i>	GG▼OGCC
<i>Bgl II</i>	A▼GATCT	<i>Eco72 I</i>	CAC▼GTG	<i>Nd I</i>	OC▼(G/C)GG
<i>BsaM I</i>	GATTGCN▼	<i>EcoI CR I</i>	GAG▼CJC	<i>Nco I</i>	C▼CATGG
<i>BsaO I</i>	OG(A/G)(T/C)▼OG	<i>EcoR I</i>	G▼AATTC	<i>Nde I</i>	CA▼TATG
<i>Bsp1286 I</i>	G(G/A/T)GC(C/A/T)▼ C	<i>EcoR V</i>	GAT▼ATC	<i>NgoM I</i>	G▼OCGGC

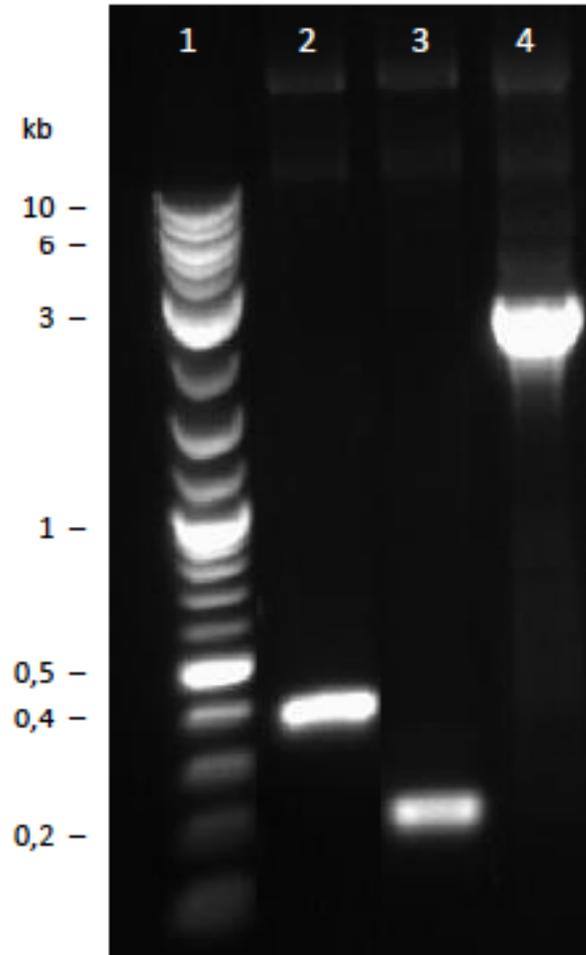


Figure : Electrophorèse sur gel d'agarose 1%. Contrôle de l'intégrité du plasmide recombinant

Séquençage de l'ADN

Le séquençage de l'ADN est réalisé en utilisant la méthode de Sanger. Les résultats de séquençage sont analysés à l'aide du logiciel Chromaspro113 et comparés à des bases de données (NCBI Blast).

Décongeler 200 μ L de bactéries compétentes.

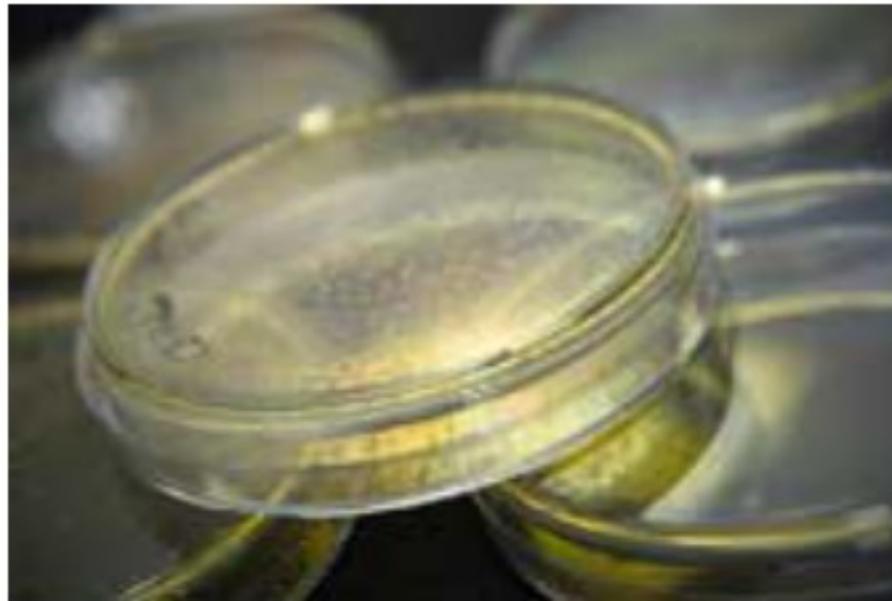
Souches bactériennes

- E. coli BL21 (DE3),
- Rosetta 2 (DE3)
- Rosetta-gami 2 (DE3) sont utilisées pour sur-exprimer la protéine

- Ajouter 10 ng de plasmide.
2. Laisser incuber dans la glace pendant 1 h.
3. Incuber à 42°C pendant 2 min.
4. Ajouter 1 mL de milieu de culture stérile, mélanger par aller-retour. Incuber 1 h à 37°C.
5. Centrifuger à 3 000 g pendant 3 min à température ambiante.
6. Eliminer 1 mL de surnageant. Reprendre délicatement le culot avec les 200 µL restants.
7. Inoculer 2 boîtes de cultures avec 100 µL de bactéries chacune. Les boîtes de cultures doivent contenir du milieu de culture et l'antibiotique de sélection.

Laisser incuber 15 min à 37°C,

Les bactéries ayant intégré le plasmide recombinant sont inoculées dans 10 ml de milieu LB (Luria Broth, 1% de tryptone, 0,5% d'extrait de levure et 0,5% de chlorure de sodium) contenant de l'ampicilline à une concentration de 100 $\mu\text{g}/\text{ml}$. Les cultures sont incubées sous agitation (200 rpm) toute la nuit à 37°C. Ces précultures servent à ensemencher 1 l de milieu de culture supplémenté de l'antibiotique.



* Les cultures sont incubées à 37°C sous agitation. Lorsque la densité optique à 600 nm atteint 0,8-1, la température est abaissée à 30°C et l'expression de la protéine est induite par l'ajout d'IPTG (isopropyl-B-D –thiogalactopyranoside) à une concentration de 1 mM.

* Après 3 h d'incubation à 30°C, les bactéries sont sédimentées par centrifugation (8200 x g pendant 15 min à 4°C), resuspendues dans du tampon 50 mM HEPES, 0,5 M NaCl, pH 7,4, et lysées par sonication.

* Les cultures sont incubées à 37°C sous agitation. Lorsque la densité optique à 600 nm atteint 0,8-1, la température est abaissée à 30°C et l'expression de la protéine est induite par l'ajout d'IPTG (isopropyl-B-D –thiogalactopyranoside) à une concentration de 1 mM.

* Après 3 h d'incubation à 30°C, les bactéries sont sédimentées par centrifugation (8200 x g pendant 15 min à 4°C), resuspendues dans du tampon 50 mM HEPES, 0,5 M NaCl, pH 7,4, et lysées par sonication.

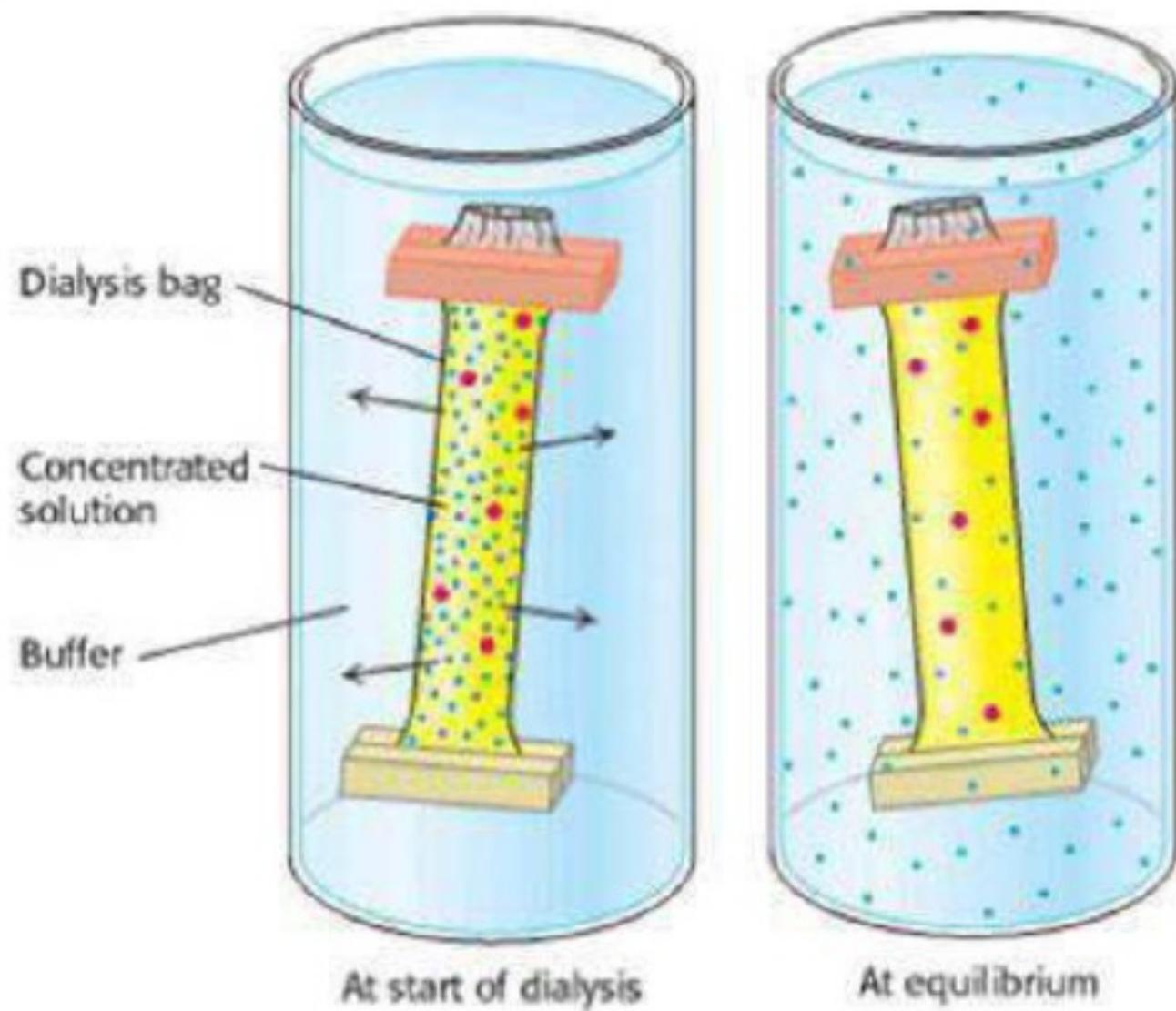
Les protéines fixées sont ensuite éluées sur 30 fois le volume de la colonne grâce au passage d'un gradient linéaire d'Imidazole (de 0 à 110 mM) à travers la résine. Les fractions contenant la protéine d'intérêt sont réunies et dialysées dans du tampon 20 mM Tris-HCl pH 7,5.

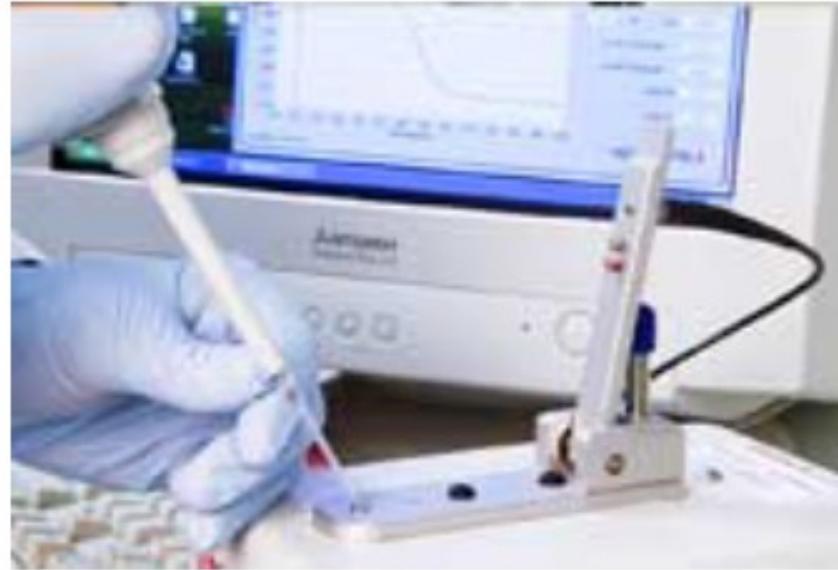
Dans

DIALYSE

Il se produit souvent qu'une préparation de macromolécules contienne différents produits dont on veut se débarrasser. Ces produits, sels, glucides, détergents ou autres petites molécules, étaient présents dans la préparation initiale ou ont été introduits lors d'une étape de la purification. Une façon simple d'éliminer ces petites molécules est la dialyse. Les molécules de petite taille qui traversent la membrane du sac à dialyse sont qualifiées de diffusibles. Celles qui sont trop grosses sont dites non diffusibles.

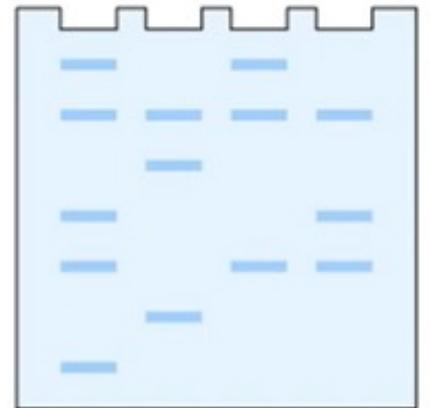
1- Dialyse





Advantages

- Sensitivity
- Specificity
- Quantification
- Flexibility
- Speed
- Minimal sample requirements
- No need for radioactive labeling
- Direct visualization
- Supershift analysis



Applications

- Transcription Factor Binding Studies
- DNA-Protein Interaction Characterization
- Protein-Nucleic Acid Interaction Studies
- Study of Protein Complexes
- Screening for Regulatory Elements
- Assessment of DNA Mutations and Variants
- Drug Development and Screening

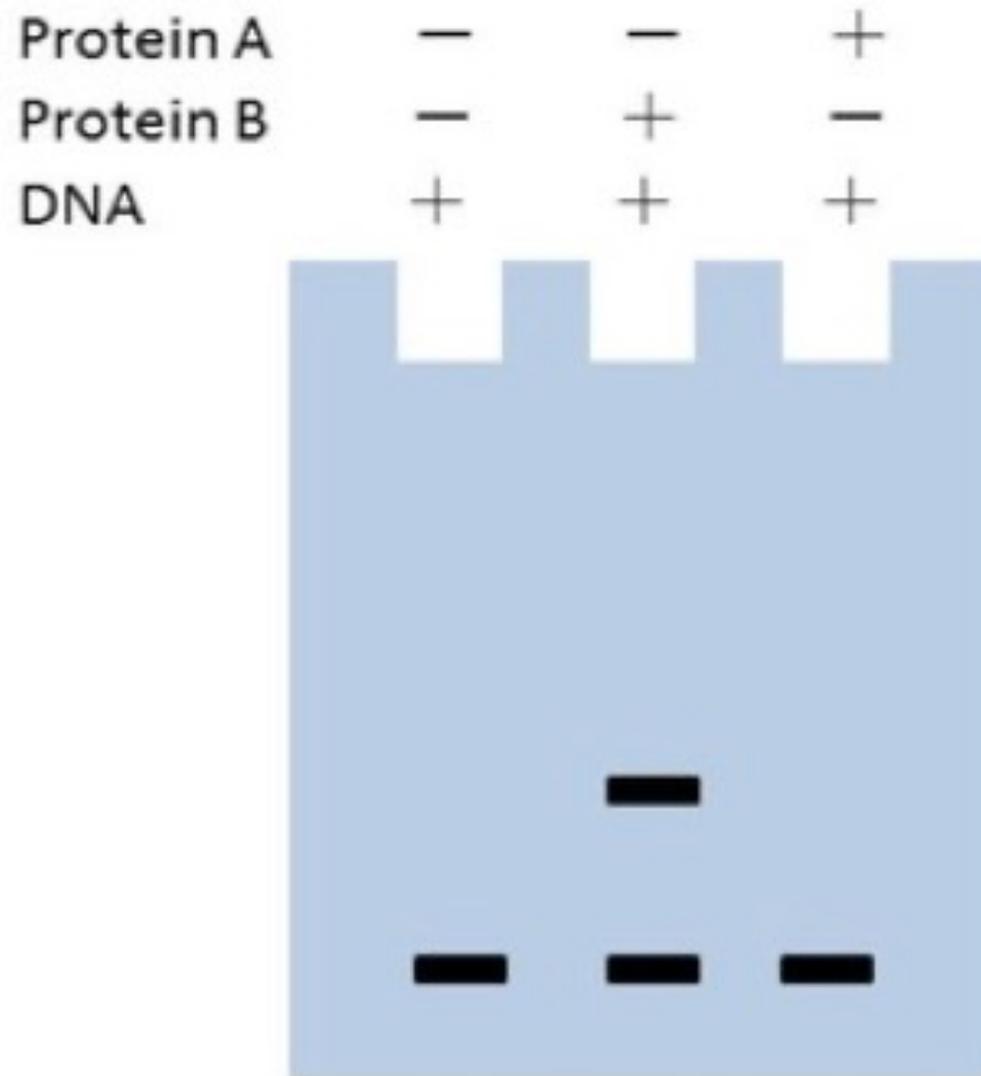


Fig2. EMSA Result 1

Analysis: Protein B binds to DNA and protein A does not interact with DNA

Protein A	-	+	-	+
Protein B	-	-	+	+
DNA	+	+	+	+

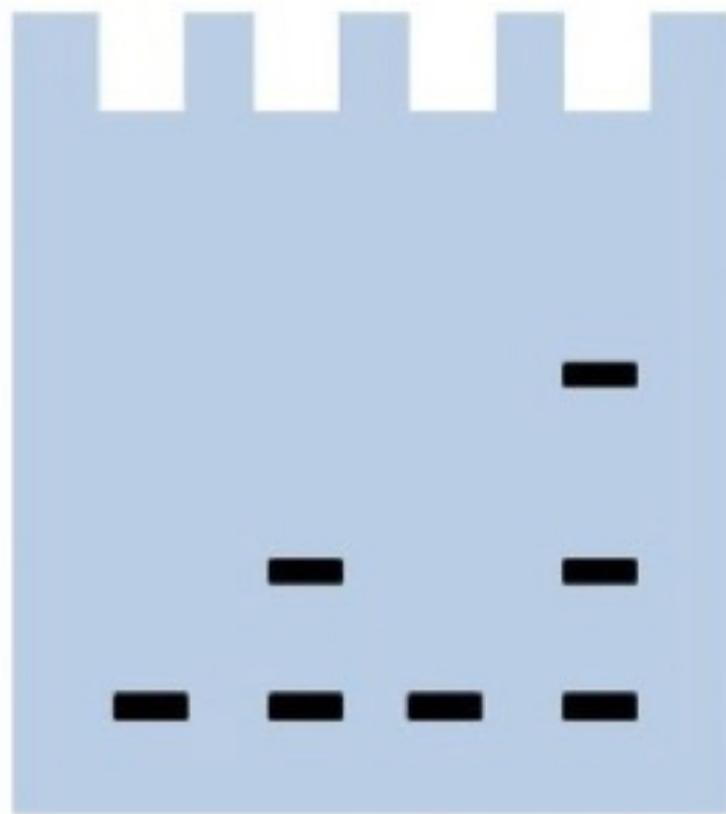
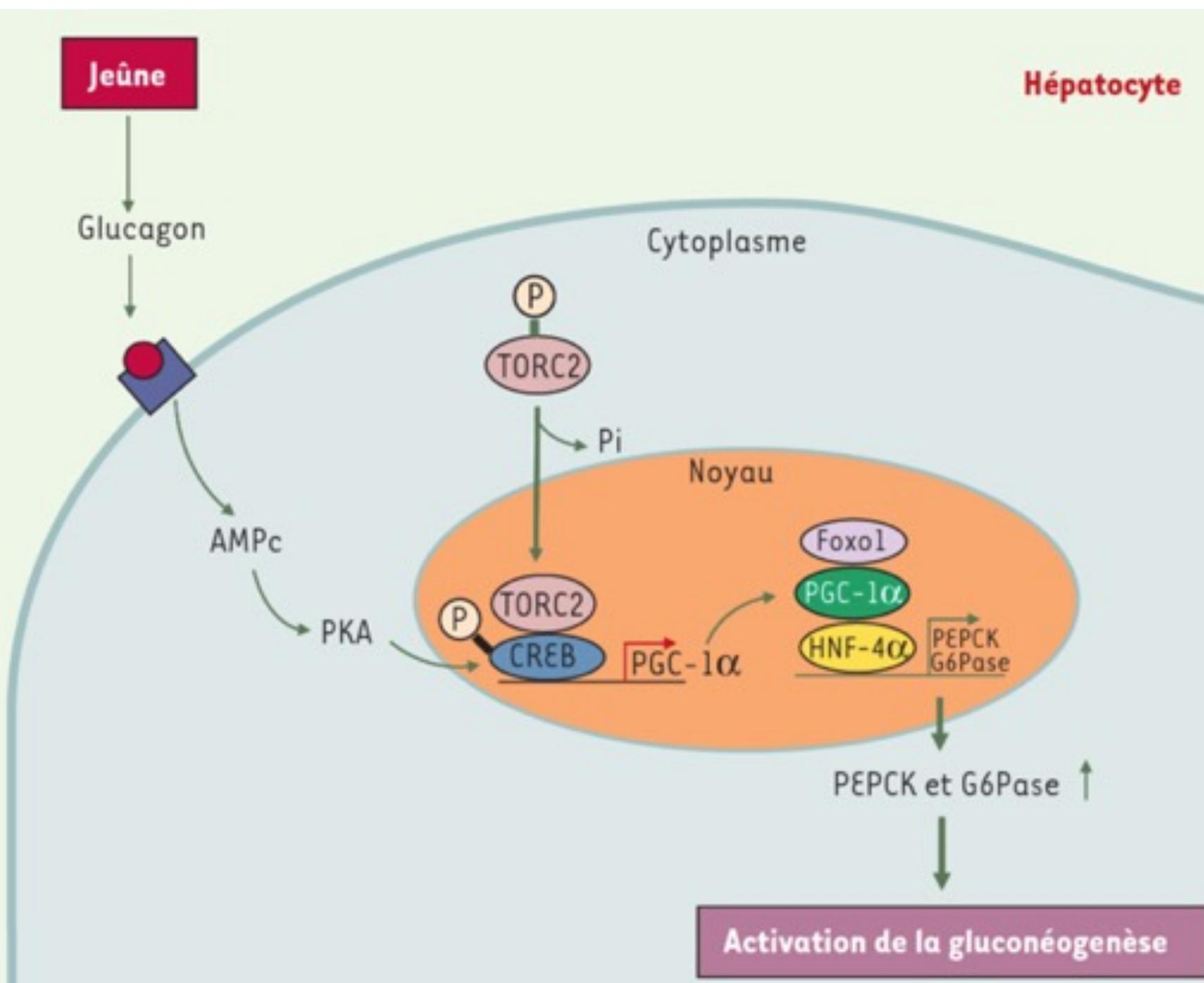


Fig3. EMSA Result 2

Analysis: Protein A binds to DNA. Protein B binds with DNA if there is protein A. The upper band in column four presents DNA+A+B. Protein B does not interact with DNA independently.



Régulation transcriptionnelle des gènes de la gluconéogenèse au cours du jeûne.

En réponse à l'activation par la PKA, le facteur de transcription CREB recrute le coactivateur TORC2 et permet l'expression du gène codant le coactivateur PGC-1a.

En association avec les facteurs de transcription HNF4a et Foxo1, PGC-1a active la transcription des gènes G6Pase et PEPCK provoquant une activation de la gluconéogenèse hépatique

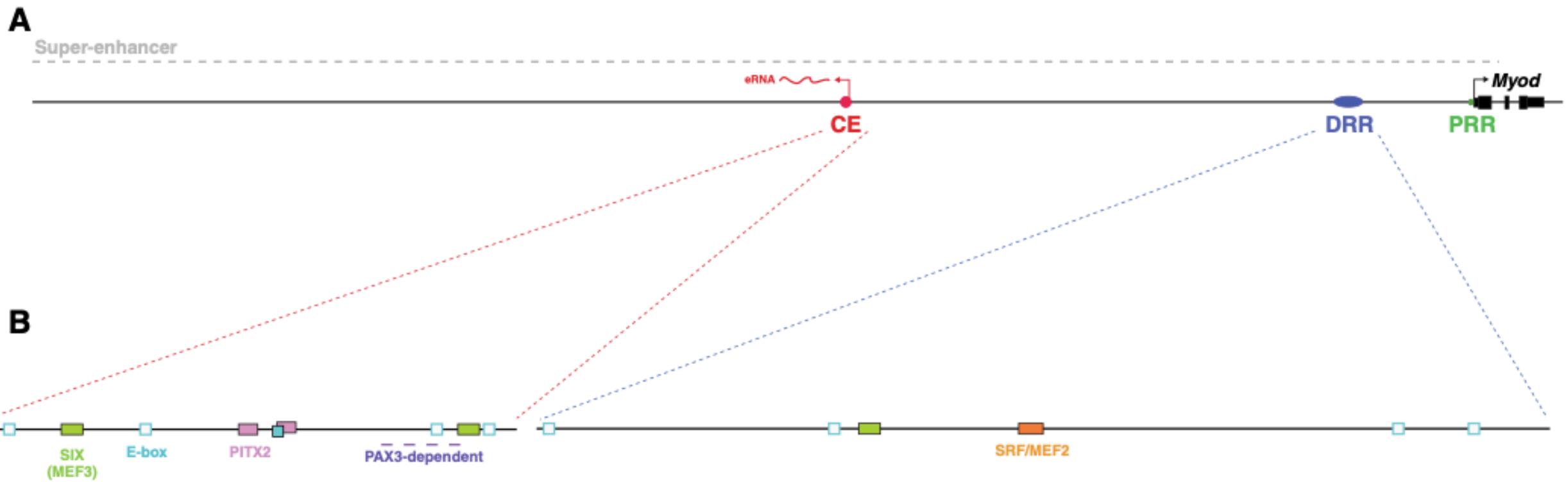


Fig.1 Sequences that regulate *Myod* expression. **a** Schematic of a region 50 kb upstream of the *Myod* gene, including the core enhancer (CE, in red), the distal regulatory region (DRR, in blue) and the proximal regulatory region (PRR, in green) immediately upstream of the *Myod* transcription start site (TSS, black arrow). This region lies within a super enhancer, associated with high levels of transcription factors occupancy, H3K27ac marks and multiple enhancer RNAs (eRNAs). One of these is an eRNA is transcribed from the CE that plays a role in regulating *Myod* expression through promoting chromatin accessibility at the *Myod* promoter. **b** Transcription factor binding sites for bHLH factors (E-box) including MYOD, MYF5, CLOCK, BMAL2, SIM2, MSC and TCF21 are shown in blue, for SIX and EYA factors in green (MEF3 site), for PITX2 in pink (Paired binding site) and for SRF/MEF2 in orange (CaRG box).