

50376
1991
298

66239

50376
1991
298

UNIVERSITE DES SCIENCES ET TECHNIQUES
DE LILLE FLANDRES-ARTOIS

Année 1991

N° d'ordre: 802

THESE DE DOCTORAT D'UNIVERSITE

(nouveau régime)

présentée à l'Université de Lille I

pour l'obtention du grade de

DOCTEUR DES SCIENCES DE LA VIE ET DE LA SANTE

par

François HAMY



ELEMENTS PEPTIDIQUES DE RECONNAISSANCE

DANS LA STRUCTURE DE RECEPTEURS NUCLEAIRES

présentée le 15 Novembre 1991 devant la Commission d'Examen

JURY

Président: Pr. André VERBERT
Rapporteurs: Pr. Jean-Pierre BOHIN
Pr. Pierre FORMSTECHE
Membres: Pr. Michael J. WARING
Dr. Jean-Pierre HENICHART

Je tiens à remercier,

Monsieur le Professeur André VERBERT. Je suis très sensible à l'honneur que vous me faites de présider ce Jury.

Monsieur le Professeur Jean-Pierre BOHIN et Monsieur le Professeur Pierre FORMSTECHEUR pour avoir accepté d'être les rapporteurs de ce travail.

Monsieur le Professeur Michael J. WARING de me faire l'honneur de juger ce travail.

Monsieur Jean-Pierre HENICHART, Directeur de Recherche à l'Institut National de la Santé et de la Recherche Médicale. Je tiens à lui adresser très sincèrement le témoignage de ma reconnaissance pour le travail qu'il m'a confié, sa passion dynamique qu'il m'a fait partager et à le remercier pour sa disponibilité permanente et les conseils qu'il m'a prodigués.

Je ne saurais trop faire part de ma gratitude toute particulière à Nicole HELBECQUE, Directeur de Recherche à l'Institut National de la Santé et la Recherche Médicale. Qu'il me soit permis de vous adresser mes respectueux remerciements pour l'attention constante, l'infinie patience et la disponibilité totale que vous m'avez données tout au long de ce travail.

Je remercie Michèle LOHEZ, Christian BAILLY, Fabrice BAILLY et Jean-François GOOSSENS qui m'ont toujours fait part de leur bonne humeur, de leur amicale contribution et de leur précieuse expérience.

Enfin, je sais gré au Ministère de la Recherche et de la Technologie d'avoir allégé mes soucis pécuniaires au cours de cette période.

*"A la Nature,
à tous ceux qu'elle émerveille."*

Hubert Reeves.

INTRODUCTION

INTRODUCTION

Un centre d'intérêt majeur de la biologie est à l'heure actuelle l'élucidation des mécanismes cellulaires et moléculaires qui régissent la transmission d'un signal et son accomplissement avec un attrait tout particulier pour les organismes supérieurs. Un aspect fondamental de la réponse de la cellule à des signaux de son environnement est la modulation de l'expression génique. De nombreuses voies trouvent leur origine dans la membrane de la cellule ou dans des compartiments intracellulaires pour se terminer dans le noyau de la cellule où elles déterminent le mode d'expression de gènes spécifiques ou modulent leur intensité.

C'est en cela que la compréhension des mécanismes par lesquels les cellules spécifient la petite fraction de leur information génétique à exprimer dans un contexte particulier, ainsi que du contrôle de cette expression à un temps donné, sont des tâches essentielles de la biologie moléculaire, voire de ses éventuelles applications thérapeutiques.

En termes d'expression génique, il doit y avoir deux types de décisions cellulaires : des décisions irréversibles conduisant à une activation ou une inhibition permanente de certains programmes génétiques, et des décisions réversibles qui dépendent de signaux transitoires. On peut s'attendre à ce que le premier type de décisions joue un rôle essentiel au cours du développement de l'organisme par l'établissement d'un mode caractéristique d'expression tissu-spécifique, alors que la régulation réversible serait prédominante dans la détermination des réponses transitoires des cellules différenciées aux divers signaux environnementaux ou métaboliques. On ne sait pas à l'heure actuelle si les mécanismes qui mènent à ces deux types de choix sont similaires ou non; cependant il est presque certain que les deux impliquent des changements dans l'organisation de l'ADN génomique au sein de la chromatine et/ou la liaison de protéines régulatrices à des séquences spécifiques de l'ADN.

Le mécanisme moléculaire de la régulation génique réversible par les hormones stéroïdiennes représente un modèle de choix de signaux agissant par l'intermédiaire d'un récepteur nucléaire. Les hormones stéroïdiennes et thyroïdiennes agissent dans la coordination d'évènements complexes impliqués dans le développement, la différenciation et la réponse physiologique à divers stimuli. Ces molécules, par leur liaison à un récepteur intracellulaire spécifique, organisent toute une batterie de processus physiologiques et comportementaux par la régulation d'un réseau de gènes.

HISTORIQUE

Des pathologies maintenant connues pour être associées à des désordres fonctionnels relatifs aux hormones stéroïdiennes et thyroïdiennes n'ont été identifiées que relativement tard dans l'histoire de la médecine : ce n'est seulement que depuis le début de ce siècle que les fondements d'une étude physiologique furent fournis par l'isolement et l'analyse structurale de ces hormones. Il était connu depuis les travaux de Huxley et d'autres, que des extraits de glande thyroïde étaient capables de contrôler la métamorphose d'amphibiens, mais ce n'est pas avant 1915 que Kendall (1915) a pu cristalliser la molécule responsable et montrer qu'elle était composée de deux résidus de tyrosine iodés (Huxley *et al.*, 1929; Gudernatsch, 1929).

Dix ans plus tard, à la fois Kendall et Reichstein achevèrent l'analyse structurale du cortisol purifié à partir du cortex adrénal, ce qui conduisit à la réalisation qu'il était (comme toutes les autres hormones stéroïdiennes) dérivé du cholestérol (Gaunt, 1975; Bentley *et al.*, 1986). Alors que beaucoup considèrent que cela constituait un aboutissement de l'endocrinologie moderne, on réalisa plus modestement que des alchimistes chinois (Needham, 1983), pour des motivations médicales, avaient développé des méthodes empiriques du X^{ème} au XVI^{ème} siècle pour purifier des stéroïdes proches de l'homogénéité.

Depuis les années 1900 jusqu'à nos jours, il s'est produit une formidable expansion dans nos connaissances des organes endocrines et des diverses physiologies qu'ils coordonnent (Gaunt, 1975; Krieger, 1981). Sur la base d'expériences biologiques, trois classes majeures d'hormones stéroïdiennes ont été décrites : les stéroïdes adrénaux (comprenant cortisol et aldostérone), les stéroïdes sexuels (progestérone, œstrogène et testostérone), et la vitamine D₃. On a démontré que ces molécules étaient cruciales pour un bon développement et une physiologie correcte chez les vertébrés et, par conséquent, chacune d'elle est devenue un pôle de recherches cliniques et biologiques (Baxter et Tyrell, 1981; Troen et Oshima, 1981; Broackus, 1981).

Les stéroïdes surrenaliens influencent largement l'homéostasie, contrôlent le métabolisme minéral et celui du glycogène et sont médiateurs dans la réponse au stress. Ils ont des effets considérables sur les systèmes nerveux et immunitaire et influencent la croissance et la différenciation de cellules en culture.

Les stéroïdes sexuels provoquent le développement et la détermination du système reproducteur embryonnaire, masculinisent ou féminisent le cerveau à la naissance, contrôlent la reproduction, les caractères sexuels secondaires et le comportement sexuel chez l'adulte.

La vitamine D est nécessaire à un bon développement osseux, elle joue un rôle critique dans le métabolisme du calcium et dans la différenciation osseuse.

Une production aberrante de ces hormones a été associée à un large spectre de pathologies incluant le cancer. Les hormones stéroïdiennes et thyroïdiennes sont d'égale importance dans la métamorphose. Un têtard thyroïdectomisé ne se développera pas pour donner une grenouille, mais l'addition de thyroxine dans l'eau induit tous les changements conduisant au développement d'un adulte terrestre (Huxley *et al.*, 1929). Similairement, les ecdystéroïdes sont les hormones de la métamorphose chez les insectes. Il a pu être montré que l'on peut associer directement l'action de

l'ecdysone à des changements dans la structure des chromosomes (Ashbrenner et al., 1980) au cours du phénomène de "puffing" induit par l'ecdysone, suggérant le lien entre les hormones stéroïdiennes et l'activation de l'expression des gènes.

Comment de si petites et relativement simples molécules peuvent-elles induire une telle diversité de réponses complexes? Le premier indice fut fourni par l'identification de récepteurs aux hormones stéroïdiennes et thyroïdiennes au début des années 1970, grâce à l'utilisation de ligands marqués radioactivement (Jensen et De Sombre, 1972; Tata et al., 1972; Tompkins et al., 1974). Dans chaque cas, l'hormone une fois liée à son récepteur induit des changements tels que le complexe devient hautement affine pour la chromatine. Pour certains récepteurs, ces changements coïncident aussi avec leur migration du cytoplasme vers le compartiment nucléaire. A son tour, ceci conduit à l'activation ou à la répression d'un nombre limité de gènes (50 à 100 par cellule) (Ivarie et O'farrel, 1978).

La sélectivité est obtenue en partie par l'expression réduite des différents récepteurs dans certains types de cellules et de tissus. L'organisation dans la structure de la chromatine étant différente pour chaque type cellulaire, l'accessibilité à certains gènes pour le complexe hormone-récepteur doit être elle aussi différente.

Des tentatives pour purifier les récepteurs aux hormones stéroïdiennes et thyroïdiennes furent entreprises en dépit du fait que ces molécules ne sont présentes qu'à l'état de traces (10^3 à 10^4 par cellule) et que cela requiert donc des enrichissements de l'ordre de 10^5 à 10^6 pour atteindre l'homogénéité. Le développement d'analogues stéroïdiens/thyroïdiens synthétiques de haute affinité permit de venir à bout des vicissitudes dans l'isolement des récepteurs et révolutionna ainsi les études cliniques et biochimiques (Krieger, 1981; Baxter et Tyrrell, 1981; Troen et Oshima, 1981; Broackus, 1981).

Au début des années 1980, la plupart des récepteurs (hormis ceux des androgènes, des minéralocorticoïdes et de l'hormone thyroïdienne) étaient purifiés (Yamamoto, 1985; Ringold, 1985) et l'on pouvait déjà proposer quelques traits concernant leur mécanisme d'action. Chaque récepteur subit une altération structurale ou "transformation" consécutivement à la liaison à l'hormone et devient alors nucléaire et capable de se lier à l'ADN. L'analyse du récepteur purifié des glucocorticoïdes a alors révélé que les propriétés de liaison à l'hormone et de liaison à l'ADN, bien que portées par une seule molécule, pouvaient être fonctionnellement séparées après protéolyse ménagée. Ceci fut la première suggestion d'une structure du récepteur organisée en domaines (Parks et al., 1974; Wrangle et Gustaffson, 1978; Wrangle et al., 1979; Carlstedt-Duke et al. 1982; Dellwegge et al., 1982; Okret et al., 1982; Wrangle et al., 1984;).

La purification et la caractérisation biochimique du récepteur des glucocorticoïdes furent accompagnées de l'identification d'une variété de gènes de réponse aux glucocorticoïdes (Yamamoto, 1985; Ringold, 1985; Parks et al., 1974; Ringold et al., 1975; Kurtz et Feigelson 1977; Karin et al., 1980; Hager et Palmiter, 1981; Spindler et al. 1982; Evans et al., 1982). La plupart de ces gènes ont été isolés et il a été montré qu'ils étaient transcriptionnellement régulés par les glucocorticoïdes. Des

études de transfert de gènes, en particulier avec le promoteur du virus de la tumeur mammaire de la souris ou le promoteur du gène humain de la métallothionéine IIa, ont permis l'identification de petites séquences (env. 20 paires de bases), actives en cis qui sont requises pour l'activation hormonale de la transcription (Govindan et al., 1982; Pfahl et al., 1982; payvar et al., 1983; scheidereit et al., 1983; Ostrowski et al., 1984). L'adjonction de ces éléments à un quelconque autre gène non hormonodépendant rend ce gène dépendant de la régulation par les hormones. Ces séquences, appelées éléments de réponse hormonale (HREs), fonctionnent indépendamment de leur position et de leur orientation et se comportent donc à la manière d'activateurs transcriptionnels (Benoist et Chambon, 1981; Laimonis et al., 1982; Banerji et al., 1983; Chandler et al., 1983). Contrairement aux autres activateurs transcriptionnels, ceux-ci ont leur activité sous la dépendance de la présence ou de l'absence d'une hormone. Ces études suggèrent que la régulation transcriptionnelle fait suite à la liaison du complexe hormone-récepteur à l'ADN. D'ores et déjà, on pouvait donc schématiser le mécanisme d'action des hormones stéroïdiennes et thyroïdiennes de la façon décrite à la figure 1.

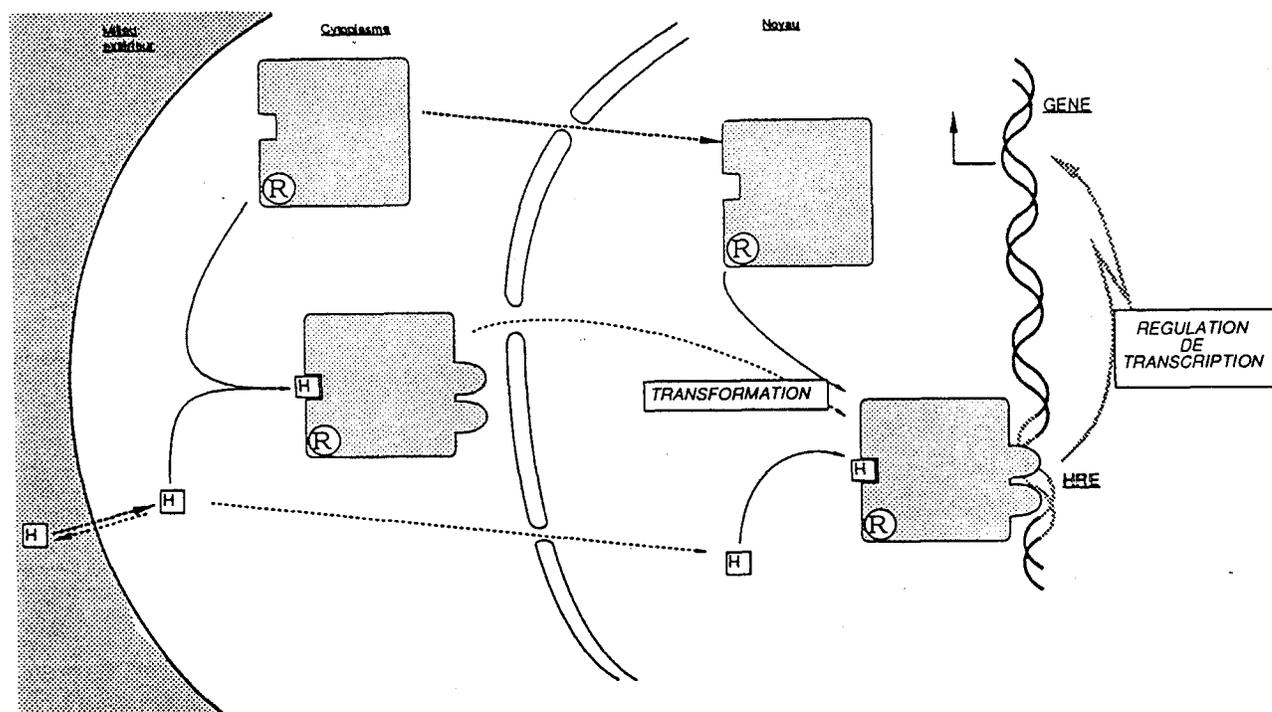


Figure 1. Mode d'action schématisé des récepteurs nucléaires: L'hormone (H), à majorité de caractère lipophile, pénètre dans la cellule par diffusion facilitée et se lie à son récepteur (R) avant ou après être entrée dans le noyau. La complexation du ligand à son récepteur initie chez ce dernier des changements conformationnels qui vont lui permettre (pour certains) d'être adressé vers le noyau et d'interagir spécifiquement avec des sites d'ADN (HRE). Il s'ensuit une activation ou une inhibition de la machinerie transcriptionnelle aboutissant à la modification de l'expression de gènes spécifiques.

Cette interprétation put être vérifiée par des études d'empreintes in-vitro révélant que les récepteurs purifiés des glucocorticoïdes, des œstrogènes, de la progestérone et des hormones thyroïdiennes se lient à des séquences en amont des gènes de réponse, sur des sites correspondant aux HREs identifiés génétiquement (Compton et al., 1983; Chambon et al. 1984, von der Ahe et al.,

1985; Klein-Hitpaß et al., 1986; Druege et al., 1986; Jantzen et al., 1987; Glass et al., 1987; Waterman et al., 1988). L'apparente symétrie de ces éléments suggère qu'ils interagissent avec un dimère de récepteur.

Anatomie comparée.

L'étude des récepteurs hormonaux est essentielle pour comprendre à la fois les origines de systèmes régulateurs complexes et leur contribution au fonctionnement de l'organisme. L'isolement des ADN complémentaires (ADNc) des récepteurs stéroïdiens a identifié une famille de gènes apparentés qui codent pour des récepteurs liant des ligands d'une remarquable diversité. L'interaction entre les gènes des récepteurs stéroïdiens, les circuits génétiques qu'ils contrôlent et leur contribution à l'organisation spatiale chez l'embryon et à la physiologie des organes chez l'adulte peuvent maintenant être élucidés.

Le clonage et l'expression du récepteur humain des glucocorticoïdes ont fourni la première structure complète d'un récepteur stéroïdien et ont révélé un segment d'ADN partageant une étonnante parenté avec l'oncogène viral erbA (Debuire *et al.*, 1984; Hollenberg *et al.*, 1985; Weinberger *et al.*, 1985). Cette relation entre l'oncogène v-erbA et les récepteurs hormonaux fut indépendamment confirmée par le clonage des récepteurs humains aux œstrogènes, progestérone, aldostérone et vitamine D (Green *et al.*, 1986; Greene *et al.*, 1986; Krust *et al.*, 1986; Loosfelt *et al.*, 1986; Weiler *et al.*, 1987; Misrahi *et al.*, 1987; Conneally *et al.*, 1987; Gronemeyer *et al.*, 1987; Mac Donnell *et al.*, 1987; Arrizza *et al.*, 1987). Deux équipes furent à l'origine de la caractérisation du produit du protooncogène erbA qui conduisit à l'identification inattendue d'un récepteur de l'hormone thyroïdienne (Weinberger *et al.*, 1986; Sap *et al.*, 1986). Il se présenta alors l'hypothèse unificatrice de la structure des récepteurs et de l'action des hormones.

Bien que les hormones stéroïdiennes et les hormones thyroïdiennes ne partagent aucune parenté ni structurale ni biosynthétique, l'existence d'une structure commune à leurs récepteurs conforte la proposition qu'il existe une grande superfamille de gènes dont les produits sont des facteurs de transcription ligand-induit.

La présence d'une séquence très conservée a initié la recherche de tels gènes occultes. Grâce aux techniques d'hybridation à faible stringence, au moins cinq nouveaux produits de gènes ont été identifiés. Deux d'entre eux, appelés gènes apparentés aux récepteurs des œstrogènes (ERR1 et ERR2), sont plus proches des récepteurs des hormones stéroïdiennes que de ceux des hormones thyroïdiennes; cependant ils ne lient aucun ligand stéroïde connu (Giguere *et al.*, 1988). Les autres sont plus proches du récepteur de l'hormone thyroïdienne. En fait, l'un d'entre eux est un deuxième récepteur de l'hormone thyroïdienne (Thompson *et al.*, 1987; Benbrook *et al.*, 1987). Un autre est un récepteur de l'acide rétinolique (Giguere *et al.*, 1987; Petkovich *et al.*, 1987; Robertson *et al.*, 1987), un métabolite de la vitamine A qui joue un rôle très important dans les phénomènes de différenciation et de morphogénèse. Le troisième partage beaucoup d'homologies avec le récepteur de l'acide rétinolique précédemment identifié (Dejean *et al.*, 1986; DeThe *et al.*, 1987). Bien que son ligand soit

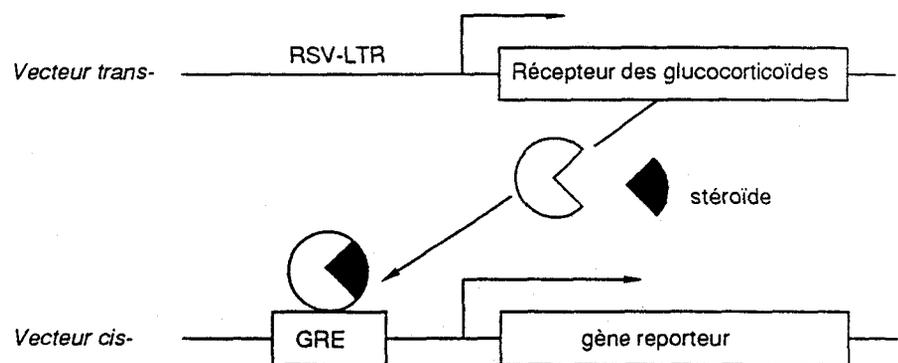
Figure 2: Comparaison des séquences en acides aminés des membres de la superfamille des récepteurs hormonaux. Les séquences primaires sont alignées sur la base de l'homologie maximale en acides aminés entre les régions. Le nombre d'acides aminés de chaque récepteur correspond à la forme humaine du récepteur à l'exception de *v-erbA* et *E75*.
 Abréviations: GR, récepteur des glucocorticoïdes; MR, récepteur des minéralocorticoïdes; PR, récepteur de la progestérone; ER, récepteur des œstrogènes; ERR1 et ERR2, récepteurs apparentés au récepteur des œstrogènes; VDR, récepteur de la vitamine D3; T3R α et T3R β , récepteurs de l'hormone thyroïdienne; RAR, récepteur de l'acide rétinoïque.
 (D'après Evans, 1985)

Domaines fonctionnels.

Le mode d'action des hormones stéroïdiennes/ thyroïdiennes classiquement admis propose que la liaison de l'hormone au récepteur induise un changement allostérique qui permet au complexe hormone-récepteur d'être adressé vers le noyau et de se lier à son élément de réponse hormonale situé dans la région promoteur d'un gène cible. C'est cette liaison qui conduit à la régulation de l'expression du gène. Le clonage de l'ADNc de récepteurs hormonaux a fourni la première opportunité de disséquer les bases moléculaires de l'action des stéroïdes.

L'identification de domaines fonctionnels pour la liaison du récepteur à l'hormone, à l'ADN et pour la transactivation fut facilitée par la technique de "screening" utilisant des cellules en culture transfectées par deux vecteurs d'expression (figure 3): le vecteur *trans*- donnant lieu à une production efficace de récepteur dans des cellules qui n'expriment pas le gène en temps normal. Le vecteur *cis*- contient un gène reporteur de luciférase (ou tout autre permettant une quantification), couplé à un promoteur contenant un élément de réponse hormonale. L'application de l'hormone ou d'un agoniste expérimental active le gène de luciférase et la quantification de cet effet est directement proportionnelle à l'efficacité du complexe hormone-récepteur à activer l'expression génomique.

Figure 3: Expériences de cotransfection
 Des cellules en culture sont cotransfectées avec l'ADNc d'un récepteur cloné dans un vecteur d'expression (vecteur *trans*-). L'efficacité de ce facteur de transcription peut être suivie par l'activité d'un gène reporteur sous la dépendance d'un promoteur contenant l'élément de réponse hormonale correspondant. Dans ce cas, le promoteur est issu du virus murin de tumeur mammaire. Il contient une séquence activatrice GRE.



Dans le cas du récepteur des glucocorticoïdes, le vecteur *cis*- contient le promoteur du virus de tumeur mammaire qui possède un élément de réponse hormonale aux glucocorticoïdes (GRE) bien

caractérisé. Dans les essais de cotransfection, l'expression du vecteur *cis-* est augmentée d'un facteur 500 et ceci de façon hormonodépendante. Par le biais de ces expériences, il est possible de suivre les effets de mutations *in-vitro* sur l'activité des récepteurs.

Les comparaisons de séquences, associées à des études fonctionnelles ont donné le jour au concept de structure, organisée en domaines fonctionnels, commune à tous les produits des gènes de la superfamille (figure 4). De même, la suggestion a été faite qu'ils pourraient tous être des dérivés évolutifs d'un gène commun ancestral présent avant la divergence entre vertébrés et invertébrés.

L'organisation en domaines fonctionnels des protéines de la superfamille des récepteurs aux hormones stéroïdiennes/thyroïdiennes est schématisée sur la figure 4. Les récepteurs sont organisés en six régions d'homologie variable (notées de A à F sur la figure 4). Les deux régions les plus conservées sont les régions C et E. Il est intéressant de constater que si les séquences d'un récepteur sont très conservées d'une espèce à l'autre, seules ces régions C et E sont conservées à travers les divers membres de la superfamille des récepteurs nucléaires, suggérant ainsi que ces régions sont importantes pour les fonctions communes aux récepteurs.

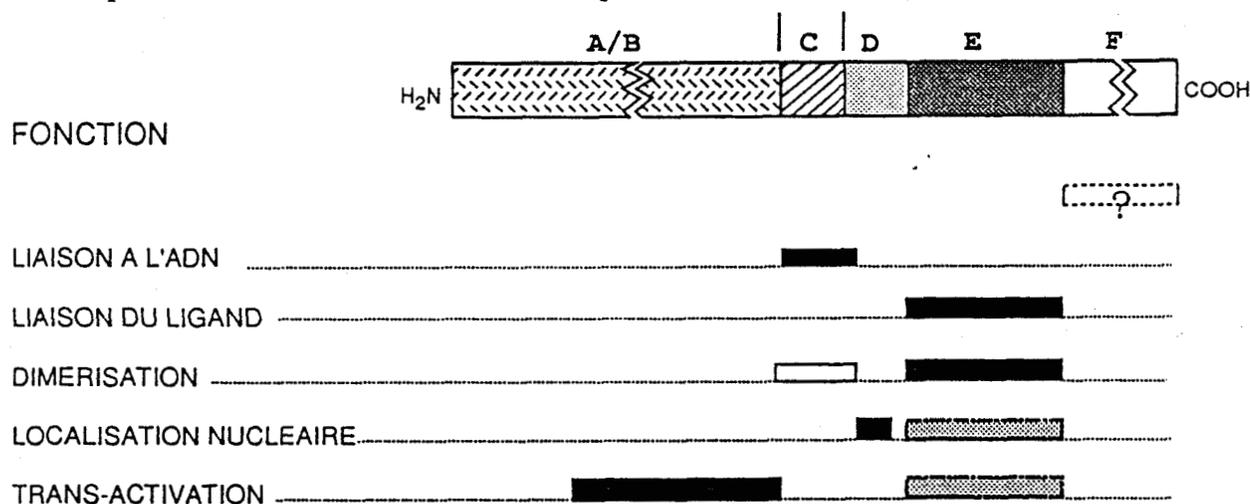


Figure 4: Organisation en domaines fonctionnels des protéines de la superfamille des récepteurs aux hormones stéroïdiennes/thyroïdiennes.

Le domaine de liaison à l'hormone a une taille approximative de 25 kDa et contient beaucoup d'acides aminés hydrophobes. Des études portant sur des récepteurs mutants générés *in-vitro* du GR, de l'ER et du PR indiquent que l'intégrité de ce grand domaine est importante pour la liaison de l'hormone. En outre, l'hormone se lie avec une affinité correcte à des récepteurs dont la région A/B ou C est absente. Le repliement correct du domaine de liaison à l'hormone est donc indépendant du domaine de liaison à l'ADN. Il semble que les acides aminés conservés de la région E soient importants pour la formation d'une poche hydrophobe globale pouvant accueillir le squelette hydrophobe d'une hormone stéroïdienne alors que les acides aminés non-conservés seraient plutôt impliqués dans des interactions spécifiques avec le ligand. Ce domaine de liaison de l'hormone fait actuellement l'objet de beaucoup de recherches dans une finalité de conception rationnelle de nouvelles molécules agonistes ou antagonistes des hormones stéroïdiennes/thyroïdiennes. Toutefois, la région E possède d'autres fonctions que celle de liaison de l'hormone : il a pu être démontré que ce

domaine joue un rôle dans la dimérisation de l'ER ainsi que dans la localisation nucléaire du GR. De plus, dans ces deux récepteurs, le domaine E contient aussi une partie de la fonction de *trans*-activation du récepteur.

Plusieurs choses laissent penser que la région centrale hautement conservée (notée C) contenait le domaine de liaison à l'ADN: (i) la présence d'un groupe de résidus basiques susceptibles d'interagir avec l'ADN, (ii) la présence d'un motif riche en résidus cystéine et (iii) la forte homologie de cette partie entre les différents récepteurs.

Afin de vérifier cette supposition, plusieurs études de mutation furent menées sur cette région (Giguere *et al.*, 1986; Kumar *et al.*, 1986; Hollenberg *et al.*, 1987; Godowski *et al.*, 1987). Les mutants continuent de lier l'hormone indiquant que la structure de la protéine est intacte; cependant ils perdent la propriété de liaison à l'ADN. Une preuve directe de la fonction a été apportée par la substitution de ce domaine putatif de liaison à l'ADN du récepteur humain de l'œstrogène par celui du récepteur des glucocorticoïdes. Le récepteur hybride résultant a, comme prévu, sa spécificité initiale substituée dans le même sens. Cela suggéra une stratégie générale appelée "échange de domaine" (finger-swap; figure 5) largement exploitée avec succès pour caractériser de nouveaux récepteurs hormonaux.

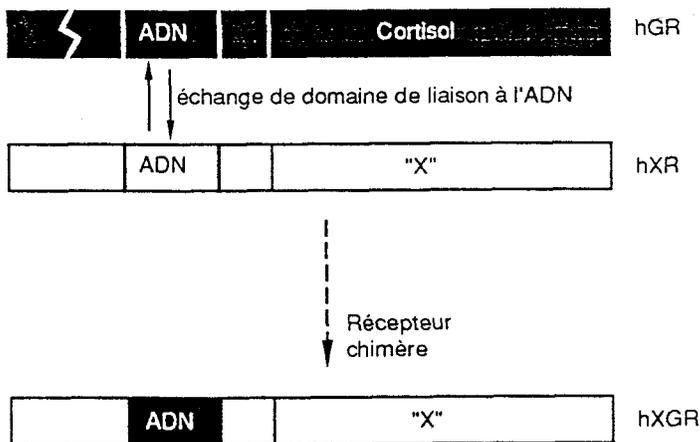


Figure 5: La technique du "finger-swap"

La structure modulaire des récepteurs hormonaux permet l'échange d'un domaine pour un autre pour générer un récepteur hybride fonctionnel. Ex: si le domaine de liaison à l'ADN d'un récepteur de l'hormone "X" est substitué par le domaine de liaison à l'ADN du récepteur des glucocorticoïdes, le récepteur chimère résultant stimule un promoteur contenant le GRE sous la dépendance de l'hormone "X".

Ces aboutissements lèvent la question de savoir s'il existe des aspects structuraux de ce domaine liant l'ADN qui permettent d'expliquer ses propriétés.

HYPOTHESES DE TRAVAIL

Les études rapportées dans ce mémoire trouvent leur motivation première dans une volonté de disséquer des phénomènes de reconnaissance moléculaire situés à des points clés d'effets biologiques. Comprendre ces phénomènes et les mimer par le biais de molécules synthétiques nous laisse l'espoir d'utiliser ces outils dans différentes disciplines de la biologie ou pour l'élaboration de composés biologiquement actifs, modèles pour la conception rationnelle d'agents thérapeutiques.

Nous avons choisi le modèle des protéines de la superfamille des récepteurs aux hormones stéroïdiennes et thyroïdiennes pour plusieurs raisons inhérentes à leurs propriétés exposées lors du premier chapitre: (i) le mécanisme d'action de ces récepteurs est divisible en plusieurs étapes où la reconnaissance moléculaire est clairement impliquée (ex: liaison à l'hormone, localisation nucléaire, reconnaissance d'un HRE, dimérisation, activation de la machinerie transcriptionnelle); (ii) la structure de ces protéines est divisible en entités fonctionnelles (i.e. en domaines à chacun desquels est attribuée une fonction plus ou moins bien définie du récepteur). L'hypothèse qui en découle alors est qu'il existe des séquences en acides aminés relativement courtes responsables du déroulement des différentes étapes dont la résultante est l'effet d'une hormone stéroïdienne/thyroïdienne sur un organisme. La relation entre une de ces séquences et sa fonction est la reconnaissance moléculaire avec d'autres constituants de la cellule.

Malgré l'intérêt que représente chacune des étapes du mécanisme d'action des récepteurs nucléaires, il serait prétentieux de les vouloir étudier toutes et nous avons donc restreint notre champ d'investigation à deux de ces étapes qui intéressent directement la thématique principale du laboratoire de Jean-Pierre Hénichart où ces recherches ont été effectuées. En effet le laboratoire s'intéresse depuis de nombreuses années à l'élaboration de substances capables d'interagir avec l'ADN, notamment dans le domaine des antitumoraux (Helbecque *et al.*, 1987; Kenani *et al.*, 1987; Bailly *et al.*, 1988; Bailly *et al.*, 1989a; Bailly *et al.*, 1989b; Bernier *et al.*, 1989; Morier-Teissier *et al.*, 1989; Plouvier *et al.*, 1991). Les molécules à action antitumorale ont en majorité l'ADN pour cible et leur toxicité subséquente les rend d'un usage délicat en thérapeutique. Les agents antitumoraux actuellement en usage clinique montrent une certaine sélectivité d'action envers les cellules ou les tissus cancéreux, néanmoins aucune cause rationnelle ne vient expliquer cette sélectivité. Une spécificité paraissant raisonnable pour ces molécules agissant sur l'ADN consiste en la fixation spécifique à une séquence oligonucléotidique donnée. Les recherches se sont donc orientées vers la conception d'hybrides entre une substance antitumorale (intercalant, alkylant,...), en général très affine pour l'ADN, et une partie dotée de reconnaissance spécifique de l'ADN, en majorité des antibiotiques naturels (nétropsine, distamycine) (pour revue Bailly et Henichart, 1991). Les molécules conçues jusqu'à présent ont un bon tropisme pour des séquences données d'ADN, mais la reconnaissance se fait par le petit sillon de la double hélice et par conséquent la séquence reconnue est de taille limitée (maximum 10 paires de bases). L'interaction des facteurs de transcription se fait par des motifs protéiques interagissant avec le grand sillon de l'ADN ouvrant ainsi de plus larges possibilités de reconnaissance (20-30 paires de bases) sur des séquences qui ont un rôle important dans le contrôle de la transcription.

D'autre part, les substances ayant une action sur l'ADN doivent remplir une condition impérative au moins chez les eucaryotes: elles doivent être localisées dans le compartiment nucléaire de la cellule. En effet, certaines substances très actives sur l'ADN in-vitro, sont exclues du noyau malgré leur petite taille d'où un intérêt certain pour des facteurs pouvant provoquer l'adressage nucléaire.

C'est pourquoi notre choix s'est porté sur deux étapes clés du mécanisme d'action des récepteurs nucléaires et surtout sur les éléments peptidiques responsables de ces fonctions: la localisation nucléaire des récepteurs et leur liaison spécifique à l'ADN.

PEPTIDES SIGNAL
DE
LOCALISATION NUCLEAIRE

Un modèle de transport sélectif pour l'import des protéines nucléaires.

La diversité des réactions biochimiques qui s'effectuent dans le noyau nécessite le mouvement de grandes quantités de macromolécules. Par exemple, pour des cellules en cours de développement, les histones doivent être importées en quantité suffisante pour doubler le contenu chromatinien à chaque division cellulaire. Cela suppose que chaque pore doit transporter de l'ordre de 100 à 500 molécules d'histones par minute en faisant abstraction de toutes les protéines nucléaires nécessaires.

Un modèle tentant d'expliquer la localisation préférentielle (voire exclusive) de certaines protéines dans le noyau de la cellule commence à émerger de différentes études. Selon ce modèle, les protéines migrent vers le noyau et se lient au complexe du pore nucléaire. Il a été montré l'existence de protéines qui reconnaissent un signal de localisation nucléaire et qui délivreraient les protéines possédant ce signal au pore nucléaire. Consécutivement à la liaison au complexe du pore, ce dernier s'ouvre et la protéine le traverse. On sait que ce passage de l'enveloppe nucléaire par les protéines nécessite la consommation de molécules d'ATP. Il n'est pas établi avec certitude à l'heure actuelle si les protéines sont transportées en association ou non avec les protéines cytoplasmique reconnaissant le signal de localisation nucléaire. Dans le cas d'un cotransport, une fois à l'intérieur du noyau, la protéine importée se dissocierait de la protéine liant le signal qui pourrait ainsi être recyclée dans le cytoplasme. Des aspects variés de ce modèle sont actuellement discutés dans la littérature au vu de résultats sans cesse renouvelés.

De courtes séquences en acides aminés dirigent la protéine vers le noyau.

Notre vision actuelle du processus de localisation nucléaire dérive de la découverte de signaux de localisation nucléaire premièrement suggérés par De Robertis *et al.* (1978). Plusieurs critères définissent un signal de localisation nucléaire : la délétion ou une mutation dans la séquence correspondant au signal provoque l'accumulation dans le cytoplasme d'une protéine normalement nucléaire, et lorsqu'il est fusionné à une protéine non nucléaire (par voie chimique ou génétique), le signal de localisation nucléaire vectorise la protéine dans le compartiment nucléaire. La première démonstration que le signal de localisation nucléaire pouvait se restreindre à une courte séquence d'acides aminés contigus est venue des études sur la nucléoplasmine, une protéine majeure de l'œuf du xénope. Dingwall et ses collaborateurs (1982) ont montré que la nucléoplasmine à laquelle on a ôté la partie C-terminale n'entre pas dans le noyau après sa micro-injection dans le cytoplasme; cependant, elle reste nucléaire si elle a été micro-injectée dans le noyau. Cette expérience démontre que la portion C-terminale de la nucléoplasmine contient une information nécessaire au passage de la protéine entière vers le noyau et indique qu'un groupe d'acides aminés inclus dans cette portion doit être déterminant. Ce résultat écarte aussi l'hypothèse de la rétention sélective d'une protéine par des éléments du noyau comme seul mécanisme de sa localisation nucléaire.

Des études ultérieures sont venues confirmer la notion d'un signal de localisation nucléaire intrinsèque à la protéine transportée (Lanford et Butel, 1984; Kalderon *et al.*, 1984). Par exemple, la substitution d'un seul acide aminé dans la séquence de l'antigène T de SV40 (Lys₁₂₈ → Thr ou Asn)

donne lieu à une protéine non nucléaire. Cette découverte a conduit à la suggestion que la séquence Pro₁₂₇-Lys-Lys-Lys-Arg-Lys-Val serait capable de diriger une protéine non nucléaire vers le noyau. Comme prévu, ce peptide s'est montré suffisant pour diriger la pyruvate kinase, normalement cytoplasmique, vers le compartiment nucléaire (Kalderon *et al.*, 1984).

De nombreuses séquences similaires aux signaux de localisation nucléaire ont été identifiées dans d'autres protéines. La comparaison entre ces différents signaux putatifs conduit à la proposition du signal de localisation nucléaire de l'antigène T de SV40 comme prototype. En fait tous les signaux putatifs ou démontrés fonctionnels ne sont pas exactement du type de celui de SV40 (Hall *et al.*, 1984; Silver et Goodson, 1988); cependant ils partagent tous une caractéristique commune : ils sont formés en très large majorité d'acides aminés à caractère basique (Arg, Lys; chaîne latérale protonée à pH physiologique).

Ceci reflète l'état des connaissances sur le signal de localisation nucléaire au moment où nous avons commencé l'étude sur les récepteurs nucléaires. La localisation cellulaire des récepteurs antérieurement à la fixation de l'hormone est l'objet de controverses. En résumé, il semble que certains soient cytoplasmiques en absence d'hormone, d'autres sont en équilibre de localisation cytoplasmique et nucléaire, d'autres encore sont exclusivement nucléaires alors que certains seraient déjà fixés à l'ADN, sur leur élément de réponse (la liaison à l'hormone n'induirait alors que l'activation de la machinerie transcriptionnelle). Quoi qu'il en soit, toutes les protéines de la superfamille sont (bien-entendu) synthétisées dans le cytoplasme et leur taille est supérieure à la limite d'exclusion stérique des pores de l'enveloppe nucléaire. Elles doivent donc comporter un élément dans leur séquence ayant fonction de signal d'adressage nucléaire. Il a été identifié dans la formule des membres de la superfamille une courte séquence comportant une large majorité de résidus basiques située dans la région D, charnière entre le domaine C de liaison à l'ADN et le domaine E de liaison à l'hormone. La nature de cette séquence homologue du signal d'adressage nucléaire de l'antigène T de SV40, ainsi que sa position très conservée au sein de la superfamille a suggéré l'hypothèse qu'elle pourrait être responsable d'une fonction commune à tous les récepteurs hormonaux : la localisation nucléaire.

Nous avons débuté cette étude par le signal putatif d'adressage nucléaire des récepteurs de l'acide rétinoïque. L'acide rétinoïque nous intéresse par ailleurs au laboratoire pour son pouvoir de puissant différenciateur sur certaines lignées tumorales (HL60, SK-NSH) (Nervi *et al.*, 1989; Menino *et al.*, 1988). Cette activité passe nécessairement par le(s) récepteur(s) de l'acide rétinoïque et si l'action d'un récepteur n'est pas uniquement régulée par sa localisation subcellulaire, elle doit au moins y contribuer dans certains cas. En effet, il y a actuellement des tentatives pour relier directement l'activité de protéines impliquées dans les phénomènes de différenciation/transformation (Jun, Fos, cdc-46,...) et leur localisation dans le cytoplasme ou le noyau (C. Dingwall, communication personnelle). On suppose qu'il pourrait exister des cas où cette localisation correspond à un "interrupteur" d'activité.

Après synthèse chimique du peptide Arg-Asn-Lys-Lys-Lys-Lys-Glu correspondant à la séquence 457-463 du récepteur α de l'acide rétinoïque (Giguere *et al.*, 1987), 457-463 du récepteur β (De The *et al.*, 1990), 457-463 du récepteur γ (Sanchez *et al.*, 1985), nous l'avons conjugué

chimiquement à la sérumalbumine bovine (normalement exclue du noyau), puis nous avons observé la localisation subcellulaire du complexe par immunofluorescence indirecte. Le complexe était introduit dans le cytoplasme de cellules en culture au moyen de deux méthodes : la technique classique de micro-injection et une technique originale de co-intériorisation virale.

Cette étude a fait l'objet d'un article intitulé "*Nuclear targeting of a viral co-internalized protein by a short peptide from hRAR.*" paru dans la revue "Bioconjugate Chemistry" (2, 375-378, 1991).

Reprinted from BIOCONJUGATE CHEMISTRY, 1991, 2.
Copyright © 1991 by the American Chemical Society and reprinted by permission of the copyright owner.

Nuclear Targeting of a Viral-Cointernalized Protein by a Short Signal Sequence from Human Retinoic Acid Receptors

François Hamy,[†] Philippe Verwaerde,[‡] Nicole Helbecque,[†] Pierre Formstecher,[§] and Jean-Pierre Hénichart^{*†}

INSERM U16, INSERM U233, Laboratoire de Biochimie Structurale, Faculté de Médecine, Place de Verdun, 59045 Lille Cedex, France. Received December 14, 1990

The synthesis of a putative nuclear localization signal from the human retinoic acid receptor is described. Its ability to target a carrier protein to the nucleus is demonstrated following microinjection or a new technique of cell internalization in the presence of adenovirus 3.

INTRODUCTION

Eucaryotic cells contain different sets of proteins, either nuclear or cytoplasmic, which are all synthesized in the cytoplasm. Mature polypeptides destined to the nuclear compartment of the cell must subsequently migrate to the nucleus (1). This involves mechanism(s) of transport, implying passage through pore complexes of the nuclear envelope. The phenomenon occurs in a fundamentally different manner from transport of proteins into mitochondria or endoplasmic reticulum due to the nature of nuclear pores which act as large aqueous channels ca. 9 nm in diameter (2). Two mechanisms of nuclear localization are commonly considered (3). Small proteins (MW < 50 kDa) freely diffuse into the nucleus with subsequent retention of nuclear proteins by binding to a nondiffusible substrate. On the other hand, access of the majority of large proteins to the nucleus requires that they participate in a specific interaction with a cellular component involved in transport of proteins to the nucleus.

A short peptidic sequence of the protein has been shown to be implied in the nuclear transfer. Experimental evidence for this was first provided by intracellular injection experiments with nucleoplasmin (4). The precise signal sequences of a number of nuclear proteins have been determined, the best characterized one being that of simian virus 40 large T-antigen.

In the present study we describe the synthesis of a short basic amino acid tract occurring in the sequences of the various forms of the retinoic acid receptor, recently deduced from their cloned cDNAs isolated from human (5-10) and mouse (11). The peptide (Asp-Arg-Asn-Lys₄-Glu) was chemically cross-linked to bovine serum albumin (BSA). This sequence, distinct from SV40 T-antigen was shown to be a new nuclear localization signal (NLS) sequence since it can direct a normally cytoplasmic BSA to the nucleus.

The comparison between two techniques (microinjection/viral cointernalization) for introducing the complex in the cytoplasm of cultured cells demonstrates the practical utility for membrane transcytosis using the original technique of viral cointernalization. This later

technique could be a very important contribution to the study of intracellular biochemistry.

EXPERIMENTAL PROCEDURES

Peptide Synthesis. The peptide Cys-Gly-Tyr-Gly-Asp-Arg-Asn-Lys-Lys-Lys-Lys-Glu was synthesized on a LKB Biolynx peptide synthesizer using Fmoc-amino acid chemistry (12) on a Wang resin to give a free carboxyl at the C-terminus (13). After cleavage from the resin and deprotection using aqueous trifluoroacetic acid, the peptide was purified by high-performance liquid chromatography on a C18 column eluted by 20% aqueous acetonitrile. The major peak (82%) was collected and lyophilized. Purity was assessed by the presence of a single peak by analytical reversed-phase high-pressure liquid chromatography, the correct amino acid composition (Gly, 2.06; Cys, 1.68; Tyr, 1.035; Asx, 1.88; Arg, 1.02; Lys, 4.02; Glu, 0.96), and the presence of the predicted parent ion by fast atom bombardment mass spectroscopy ($M^+ = 1425.6$).

Peptide-Protein Conjugation. The synthetic peptide was conjugated to bovine serum albumin (BSA) with a bifunctional cross-linking reagent, *m*-maleimidobenzoic acid succinimido ester (MBS), according to the method of Lerner (14). For activating BSA, 5.0 mg of BSA in 0.05 M phosphate buffer, pH 7.0, was reacted with MBS dissolved in tetrahydrofuran at a molar ratio of 1:40 BSA:MBS. The solution was stirred for 30 min at room temperature. Unreacted MBS was removed by dialysis against distilled water for one night at 4 °C. Purified peptide (2 mg) was added to 2 mg of MBS-activated BSA in 2 mL of 0.05 M phosphate buffer, pH 7.0. The solution was stirred for 3 h at room temperature and the final product was isolated by dialysis against water for 48 h at 4 °C.

A sample was examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (5-30%) to visualize an increase in molecular weight due to the addition of peptide to the protein. A more accurate estimate of the average degree of conjugation was determined by amino acid analysis (three measurements). The increase in the ratios of residues contained in the peptide to those that were absent was used to calculate the ratio of the peptide to BSA in the conjugated sample.

Microinjection. MCF 7 and HeLa S3 cells were grown in Dulbecco medium (GIBCO Laboratories, Grand Island, NY) on No. 2 cover slips which had been premarked to identify and localize injection sites. About 150 cells of each kind were microinjected per assay in the growth

* Author to whom correspondence should be addressed.

[†] INSERM U16.

[‡] INSERM U233.

[§] Laboratoire de Biochimie Structurale.

medium at 23 °C by the method of Graessmann and Graessmann (15) with a Leitz micromanipulator. Glass needles were femtotips (Eppendorf, Hamburg, Germany), the internal diameter of which was estimated to be 0.5 μm . Approximately 100 cells were injected. After injection cells were returned to a CO₂ incubator for 1 and 2 h.

Viral Cointernalization. The adenovirus serotype 3 (subgroup B) was produced in suspension-grown KB cells. Adenovirions were extracted and purified in CsCl gradients, as previously described (16) and the virus titers were determined and expressed as fluorescent focus units (FFU). MCF 7 and HeLa S3 cell monolayers were infected with adenovirus 3 suspensions (300 FFU/cell) in Dulbecco medium including BSA or peptide-BSA at 1 mg/mL. After several incubation times at 37 °C (20, 40, 60, 120 min), cells were prepared for immunofluorescence.

Immunofluorescence. Soon after incubation cells were harvested for immunofluorescence according to the method of Langranger et al. (17). Cells were washed three times in phosphate-buffered saline (PBS), pH 7.4, fixed in cold methanol (95% in PBS) (4 °C) for 5 min and washed again in four changes of PBS (pH 7.4). Cells were then fixed in paraformaldehyde (3%, final) and glutaraldehyde (0.05%, final) in PBS (pH 7.4) for 20 min, washed in solution A (5 mM Pipes, 5 mM MgCl₂, 2 mM ETGA, in PBS, pH 6.1) and incubated for 10 min in 0.1 M glycine in solution A at room temperature. Subsequently cells were permeabilized with Triton X100, 5% in solution A for 5 min, washed in solution A, and incubated in normal rabbit IgG (0.5% in PBS, pH 7.4) for 16 h at 4 °C. BSA was detected by indirect immunofluorescence using goat anti BSA IgG (Sigma, St. Louis, MO), internalized fluorescent macromolecules were examined with an Olympus BH-2 microscope (excitation wavelength at 530–560 nm, edge filter at 580 nm). Photomicrographs were taken with a Kodak citachrome film (400 Asa).

RESULTS

Conjugation of the Synthetic Peptide to a Carrier Protein. A dodecamer synthetic peptide was synthesized which contains the eight amino acid core sequence Asp-Arg-Asn-Lys-Lys-Lys-Lys-Glu from hRAR (human retinoic acid receptor) that is supposed to act as nuclear localization signal. The sequence Cys-Gly-Tyr-Gly was added to the amino terminal as a spacer containing a cysteinyl residue for ease of coupling to a carrier protein (18). It was coupled to BSA chosen as carrier protein because of its widespread use in similar studies. Coupling was made by using the MBS procedure (14). The coupling ratios were estimated by two methods: amino acid analysis and estimation of the increase in apparent molecular weight of BSA (67 000 \rightarrow 72–74 000) by retardation of mobility by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and coomassie staining. Both data gave an average ratio of 4.2 peptides per carrier molecule.

Cell Internalization of the Peptide-BSA Conjugate. Microinjection of BSA alone (Figure 1B) shows immunofluorescence in the cytoplasm and a relatively dark nucleus: very little if any nuclear transport without the signal peptide attached. With the signal attached, cells exhibit bright nuclei, but the cytoplasm was never completely depleted, even 2 h after microinjection. These data suggest that this peptide from hRARs is a nuclear localization signal but not as strong as those observed in the literature. The same results were observed for the same time (2 h) and shorter times (40 and 50 min) following cointernalization (data not shown). Though it seems clear that viral cointernalization works to get the fluorescence

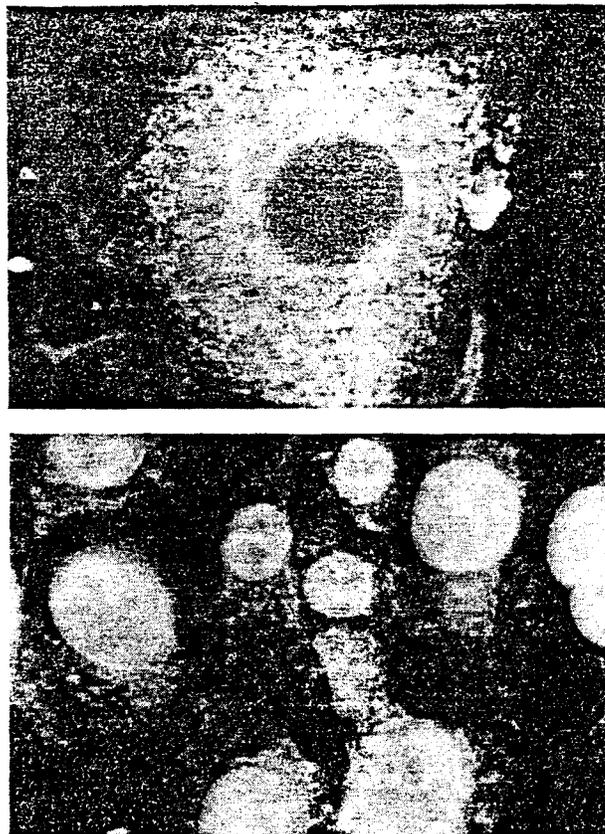


Figure 1. Subcellular localization of microinjected BSA (A, top) and peptide-BSA conjugate (B, bottom) in HeLa S3 cells. Proteins were visualized by indirect immunofluorescence staining as described in the Experimental Procedures. (The results shown are representative of more than 50 cells positive for nuclear localization.)

into the cytoplasm, results were not as clear, essentially because of leakage of control BSA on the surface of the nucleus probably due to the mechanism of action of adenovirus 3.

The choice of adenovirus 3 (Ad3) was directed by the fact that this DNA virus has a slow infection process (early viral RNA synthesis occurs about 3 h after infection (19)). Moreover, events such as membrane and cytoskeleton reorganization, which normally arise at early times after adenovirus infection, are less important with adenovirus 3 (e.g. processing of vimentin is hardly, if at all, induced by Ad3 serotype) (20). Therefore biochemical activities of the virus do not a priori interfere with nuclear translocation that arises between 20 and 30 min after the protein enters the cytoplasm (kinetics data from ref 21). Secondly endocytosis vesicles including adenovirus 3 develop until the formation of an endosome, the content of which is poured in the vicinity of nuclear pores.

DISCUSSION

The problem of the existence of specific targeting signals for the nuclear import of proteins has been the subject of many studies during the last decade (22). The demonstration of the presence in retinoic acid receptors of such a signal sequence led us to bring into play different methodologies. The first of them concerned the introduction of the protein, bearing several copies of the basic sequence, into the cytoplasm. Among the numerous ways of introducing proteins into the cytoplasm of cultured cells (23), we first selected microinjection because of its widespread use in this kind of study, though the traumatism caused by the glass needle immediately triggers

Table I. Comparison of Amino Acid Sequences of Putative Nuclear Localization Signals in the Steroid/Thyroid Hormone Receptor Superfamily, SV40 Large T-Antigen, and *Xenopus laevis* Nucleoplasmin and Lamin L1^a

protein	sequence and position of its first amino acid	receptor domain	experimental arguments for nuclear localization signal activity
SV 40 large T-antigen	<u>P</u> <u>K</u> <u>K</u> <u>K</u> <u>R</u> <u>K</u> <u>V</u>		+
<i>X. laevis</i> nucleoplasmin	<u>Q</u> <u>A</u> <u>K</u> <u>K</u> <u>K</u> <u>K</u> <u>L</u>		+
<i>X. laevis</i> L1	<u>K</u> <u>G</u> <u>K</u> <u>R</u> <u>K</u> <u>R</u> <u>I</u>		+
retinoic acid receptors	(α) 162 (β) 155 <u>R</u> <u>N</u> <u>K</u> <u>K</u> <u>K</u> <u>E</u> (γ) 164	hinge (+14)	+
thyroid hormone receptors	(α) 142 <u>R</u> <u>E</u> <u>R</u> <u>R</u> <u>R</u> <u>K</u> <u>E</u> (β) 191 <u>R</u> <u>E</u> <u>K</u> <u>R</u> <u>R</u> <u>E</u> (α) 130 <u>K</u> <u>R</u> <u>V</u> <u>A</u> <u>K</u> <u>R</u> <u>K</u> <u>L</u> (β) 179 <u>K</u> <u>R</u> <u>L</u> <u>A</u> <u>K</u> <u>R</u> <u>K</u> <u>L</u>	hinge (+27) hinge (+27) hinge (+15) hinge (+15)	
vitamin D receptor	102 <u>R</u> <u>K</u> <u>R</u> <u>E</u> <u>M</u> <u>I</u> <u>L</u> <u>K</u> <u>R</u> <u>K</u>	hinge (+18)	
glucocorticoid receptor	491 <u>R</u> <u>K</u> <u>T</u> <u>K</u> <u>K</u> <u>K</u> <u>I</u> <u>K</u>	hinge (+10)	+
progesterone receptor	637 <u>R</u> <u>K</u> <u>F</u> <u>K</u> <u>K</u> <u>F</u> <u>N</u> <u>K</u>	hinge (+10)	+
estrogen receptor	256 <u>R</u> <u>K</u> <u>D</u> <u>R</u> <u>R</u> <u>G</u> <u>R</u>	hinge (+11)	

^a The sequences are reported for human T3 receptors α (31) and β (32), vitamin D receptor (33), glucocorticoid receptor (34), progesterone receptor (35), and estrogen receptor (36, 37). These sequences are highly conserved among species for a given receptor (ref 6 and references cited therein). Numbers in parentheses indicate the distance of the sequence from the last conserved cysteine in the DNA binding domain (38). Basic amino acids are underlined. Signal sequences of *X. laevis* nucleoplasmin and lamin L1 were found in ref 27.

different sorts of metabolic reactions to repair damages to the integrity of the membrane. Therefore, to make an attempt at being more physiological, we used viral coinfection, a phenomenon based on the early permeabilization of cells to proteins by animal viruses. Evidence for coinfection was first described by Otero (24). Cell infection by a virus first involves fixation of a viral particle on the external side of the host-cell membrane. Subsequently the viral particle internalization occurs, accompanied by an entry of extracellular medium among which are macromolecules such as proteins. This method presents many advantages since (i) it is more physiological (for short times) than others like microinjection, (ii) close to the totality of cells are infected, and (iii) it is easy to carry out even for kinetic studies and for suspended cells.

In the case of the steroid/thyroid hormone receptor superfamily, experiments conducted either on the glucocorticoid (25) or on the progesterone (26) receptors suggest the existence of at least two nuclear localization signals. One of them is responsible for hormone-dependent nuclear uptake. The other, constitutively active and located at the C-terminus of the DNA-binding region, consists of a small basic sequence similar to other nuclear localization signals characterized in many proteins. Such motifs are present in the hinge region of all steroid receptor superfamily (Table I). However, none of the consensus proposed for the nuclear localization signals, either in the case of steroid hormone receptors (26) or of various nuclear proteins (27), can account for the diversity of the potential nuclear localization signals found in the steroid/thyroid receptors superfamily. Our demonstration that the particular sequence found in the retinoic acid receptors is effectively able to target to the nucleus a nonnuclear protein is the first one in this field. However, the peptide corresponding to this sequence is a relative weak nuclear location signal (no total nuclear immunofluorescence was observed for 4.2 peptides/BSA after 2 h) and further work will be required to understand how this signal works in its natural context. Mutagenesis experiments might address this question. In addition, it could be thought that other sequences of the protein are also involved in its nuclear localization. In particular, a recent study (28) has pointed out that, in several but not in all nuclear proteins containing a SV40-like motif, there are two conserved basic amino acids 10 residues before the signal. The spacer can sometimes be longer (e.g. 11 residues in the estrogen

receptor). In the case of hRARs, 12 residues separate a highly conserved lysyl residue in the superfamily from the NLS.

Though the initial site for signal recognition has not been defined, recent results (29, 30) suggest that nuclear import may be a receptor-mediated process. If this hypothesis reflects a step of the actual mechanism of protein import into the nucleus, the nuclear localization signal must then be implicated in protein-protein interactions. In this case its conformational flexibility might play a critical role too. The presumed nuclear localization signal sequence from hRARs corresponds to a sequence localized in a region of the receptors which has a predicted helical conformation. In fact, the association basic amino acid sequence/helix-predicted region is conserved in the hinge region of each member from the steroid/thyroid receptors superfamily, and preliminary results indicate that the corresponding synthetic peptides share a similar surrounding-induced behavior with those from hRAR (work in progress). A generalization in that field of sequence/conformational behavior should lead to the conception of a simplified nuclear localization vector agent providing a tool for nuclear vectorization of diverse molecules.

ACKNOWLEDGMENT

This work was financially supported by INSERM and the Association pour la Recherche contre le Cancer (No. 6932). F.H. and P.V. are allocataires of the Ministère de la Recherche et de la Technologie.

LITERATURE CITED

- (1) Dingwall, C., and Laskey, R. A. (1986) Protein import into the cell nucleus. *Annu. Rev. Cell. Biol.* 2, 367-370.
- (2) Gerace, L., and Burke, B. (1988) Functional organization of the nuclear envelope. *Annu. Rev. Cell. Biol.* 4, 335-375.
- (3) Newmeyer, D. D., and Forbes, D. J. (1988) Nuclear import can be separated into distinct steps: in vitro nuclear pore binding and translocation. *Cell* 52, 641-653.
- (4) Dingwall, C., Schoolnick, S. V., and Laskey, R. A. (1982) A polypeptide domain that specifies migration of nucleoplasmin into the nucleus. *Cell* 30, 449-458.
- (5) Benbrook, D., Lenhardt, E., and Pfahl, M. (1988) A new retinoic acid receptor identified from a hepatocellular carcinoma. *Nature* 333, 669-672.

- (6) Brand, N., Petkovich, M., Krust, A., Chambon, P., De Thé, H., Marchio, A., Tiollais, P., and Dejean, A. (1988) Identification of a second retinoic acid receptor. *Nature* 332, 850-853.
- (7) De Thé, H., Marchio, A., Tiollais, P., and Dejean, A. (1989) Differential expression and ligand regulation of the retinoic acid receptor α and β genes. *EMBO J.* 8, 429-433.
- (8) Giguère, V., Ong, E. S., Segui, P., and Evans, R. M. (1987) Identification of a receptor for the morphogen retinoic acid. *Nature* 330, 624-629.
- (9) Krust, A., Kastner, P., Petkovich, M., Zelent, A., and Chambon, P. (1989) A third human retinoic acid receptor: hRAR. *Proc. Natl. Acad. Sci. U.S.A.* 86, 5310-5314.
- (10) Petkovich, M., Brand, N. J., Krust, A., and Chambon, P. (1987) A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature* 330, 444-450.
- (11) Zelent, A., Krust, A., Petkovich, M., Kastner, P., and Chambon, P. (1989) Cloning of murine α and β retinoic acid receptors and a novel receptor γ predominantly expressed in skin. *Nature* 339, 714-717.
- (12) Sabatier, J. M., Teissier-Rochat, M., Granier, C., Van Ritschoten, J., Pedroso, E., Grandas, A., Albericio, F., and Giralt, E. (1987) Convergent solid phase synthesis VI: Synthesis by the Fmoc-procedure with a modified protocol of two protected segments, sequence 15-17 and 18-31 of the neurotoxin II of the scorpion *Androctonus australis* Hector. *Tetrahedron* 43, 5973-5980.
- (13) Wang, S. S. (1973) pAlkoxy alcohol resin for solid phase synthesis of protected peptide fragments. *J. Am. Chem. Soc.* 95, 1328-1333.
- (14) Lerner, R. A., Green, N., Alexander, H., Liu, F. T., Sutcliffe, J. G., and Shinnick, T. M. (1981) Chemically synthesized peptides predicted from the nucleotide sequence of the hepatitis B virus genome elicit antibodies reactive with the native envelope protein of dane particle. *Proc. Natl. Acad. Sci. U.S.A.* 78, 3403-3407.
- (15) Graessmann, M., and Graessmann, A. (1976) "Early" simian-virus 40 specific RNA contains information for tumor antigen formation and chromatin replication. *Proc. Natl. Acad. Sci. U.S.A.* 73, 366-370.
- (16) D'Halluin, J. C., Milleville, M., Boulanger, P. A., and Martin, G. R. (1978) Temperature-sensitive mutant of adenovirus type 2 blocked in virion assembly: accumulation of light intermediate particles. *J. Virol.* 26, 344-356.
- (17) Langranger, G., De Mey, J., Moeemans, M., Daneels, G., De Brabander, M., and Small, J. V. (1984) Ultrastructural localization of α -actinin and filamin in cultured cells with the immunogold staining (IGS) method. *J. Cell. Biol.* 99, 1324-1334.
- (18) Lanford, R. D., Kanda, P., and Kennedy, R. C. (1986) Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. *Cell* 46, 575-582.
- (19) Flint, B. J. (1986) Regulation of adenovirus mRNA formation. *Adv. Virus Res.* 31, 169-228.
- (20) Belin, M. T., and Boulanger, P. (1987) Processing of vimentin occurs during the early stages of adenovirus infection. *J. Virol.* 61, 2559-2566.
- (21) Richardson, W. D., Mills, A. D., Dilworth, S. M., Laskey, R. A., and Dingwall, C. (1988) Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. *Cell* 52, 655-664.
- (22) Silver, P. A. (1991) How proteins enter the nucleus. *Cell* 64, 489-497.
- (23) Richardson, W. D. (1988) Introducing proteins into cultured animal cells. *J. Cell. Sci.* 91, 319-322.
- (24) Otero, M. J., and Carrasco, L. (1987) Proteins are co-internalized with virion particles during early infection. *Virology* 160, 75-80.
- (25) Picard, D., and Yamamoto, K. R. (1987) Two signals mediate hormone dependent nuclear localization of glucocorticoid receptor. *EMBO J.* 6, 3333-3340.
- (26) Guiochon-Mantel, A., Loosfelt, H., Lescop, P., Sar, S., Atger, M., Perrot-Applanat, M., and Milgrom, E. (1989) Mechanisms of nuclear localization of the progesterone receptor: evidence for interaction between dimers. *Cell* 57, 1147-1154.
- (27) Chelsky, D., Ralph, R., and Jonak, G. (1989) Sequence requirements for synthetic peptides mediated translocation to the nucleus. *Mol. Cell. Biol.* 9, 2487-2492.
- (28) Robbins, J., Dilworth, S. M., Laskey, R. A., and Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmic nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence. *Cell* 64, 615-623.
- (29) Adam, S. A., Lobl, T. J., Mitchell, M. A., and Gerace, L. (1989) Identification of specific binding proteins for a nuclear location sequence. *Nature* 337, 276-279.
- (30) Yamasaki, L., Kanida, P., and Lanford, R. E. (1989) Identification of four nuclear transport-binding proteins that interact with diverse transport signals. *Mol. Cell. Biol.* 9, 3028-3036.
- (31) Benbrook, D. (1987) A novel thyroid hormone receptor encoded by a cDNA clone from a human testis library. *Science* 238, 788-791.
- (32) Weinberger, C., Thompson, C. C., Ong, E. S., Lebo, R., Gruol, D. J., and Evans, R. M. (1986) The c-erb-A gene encodes a thyroid hormone receptor. *Nature* 324, 641-646.
- (33) Walters, M. R., Hungiker, W., and Norman, A. W. (1981) 1,25-Dihydroxyvitamin D₃ receptors. *Trends Biochem. Sci.* 6, 268-271.
- (34) Hollenberg, S. M., Weinberger, C., Ong, E. S., Cerelli, G., Oro, A., Lebo, R., Thompson, E. B., Rosenfeld, M. G., and Evans, R. M. (1985) Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* 318, 635-641.
- (35) Misrahi, M., Atger, M., D'Auriol, L., Loosfelt, H., Meriel, C., Friedlansky, F., Guiochon-Mantel, A., Galibert, F., and Milgrom, E. (1987) Complete amino acid sequence of the human progesterone receptor deduced from cloned cDNA. *Biochem. Biophys. Res. Commun.* 143, 740-748.
- (36) Green, S., Walter, P., Kumar, V., Krust, A., Bornet, J. M., Argos, P., and Chambon, P. (1986) Human oestrogen cDNA: sequence, expression and homology to v-erb-A. *Nature* 320, 134-139.
- (37) Greene, G. L., Gilna, P., Waterfield, M., Baker, A., Hort, Y., and Shine, J. (1986) Sequence and expression of human estrogen receptor complementary DNA. *Science* 231, 1150-1154.
- (38) Evans, R. M. (1988) The steroid and thyroid hormone receptor superfamily. *Science* 240, 889-895.

Registry No. Cys-Gly-Tyr-Gly-Asp-Arg-(Lys)₄-Glu, 135929-31-6; retinoic acid, 302-79-4.

ELEMENTS PEPTIDIQUES
DE
RECONNAISSANCE
D'UNE
SEQUENCE D'ADN

Il ressort donc de cette étude que la séquence Arg-Asn-Lys-Lys-Lys-Lys-Glu des récepteurs de l'acide rétinoïque est suffisante pour l'adressage nucléaire d'une protéine normalement exclue du noyau. Dans une seconde étape, nous avons voulu généraliser cette relation séquence-fonction pour d'autres signaux putatifs des membres de la superfamille des récepteurs hormonaux nucléaires. Nous avons pour ce faire utilisé la même approche qui consiste en la synthèse de séquences présumées signal de localisation nucléaire d'autres récepteurs et en la démonstration de leur capacité d'adressage nucléaire de la sérumalbumine bovine. Les séquences signal putatives ont été sélectionnées essentiellement selon deux critères : elles devaient être représentatives des différents sous-groupes de la superfamille et être assez différentes en structure primaire pour déceler d'éventuelles différences d'efficacité pouvant in-fine aider à la définition d'une séquence consensus. Nous avons eu aussi le souci de synthétiser le peptide signal de l'antigène T de SV40 afin de posséder un témoin positif de localisation nucléaire et de pouvoir comparer les efficacités des signaux des récepteurs par rapport à ce prototype souvent étudié dans la littérature.

D'autre part, la technique de cointernalisation virale nous a permis de mettre au point un protocole de compétition entre un complexe signal-sérumalbumine et le signal seul. Cette partie de l'étude a fait l'objet de la rédaction d'un article "*Comparison between synthetic nuclear localization signal peptides from the steroid/thyroid hormone receptors superfamily.*" paru dans la revue "Biophysical Biochemical Research Communications".

COMPARISON BETWEEN SYNTHETIC NUCLEAR LOCALIZATION SIGNAL PEPTIDES FROM THE STEROID/THYROID HORMONE RECEPTORS SUPERFAMILY.

François HAMY, Nicole HELBECQUE and Jean-Pierre HENICHART.

INSERM Unité N°16
1, Place de Verdun 59045 LILLE CEDEX - FRANCE.

Summary: The main objective of the study is to demonstrate that short basic peptides from the steroid/thyroid hormone receptors superfamily act as Nuclear Localization Signals out of receptors context. Such synthesized peptides, chemically coupled to Bovine Serum Albumin were shown to enable the corresponding BSA-conjugate to be transported to the nucleus. A second objective is to demonstrate the utility of viral cointernalization as a good method for rapid quantitation, comparison and competition in nuclear entry.

It becomes now well-established that large nuclear proteins, which are synthesized in the cytoplasm, use a molecular recognition step involving an intrinsic Nuclear Localization Signal (NLS) to enter the nucleus via nuclear pores. Though the identification of a "receptor" for such signals is an active field of research, it has not yet been precisely defined. The intrinsic signal, however, has been well-characterized in several nuclear proteins, the first one being the C-terminal region of nucleoplasmin (1). Other studies have defined NLS as a short stretch of basic aminoacids such as the sequence Pro-Lys-Lys-Lys-Arg-Lys-Val of SV40 large T antigen (2-5) though not all NLSs are of this type (6).

For about ten years, SV40-like NLSs have been characterized in many nuclear proteins but attempts to find a consensus, at least for sequence, have failed. In a way to make a good comparison in efficiency for nuclear targeting of several presumed SV40-like NLSs, we focussed the present study on representative members from a single nuclear protein family: the steroid/thyroid hormone receptors superfamily. Actually, since the question "*A nuclear localization signal in steroid hormone receptors ?*" was asked by Wolff *et al.* in 1987 (7), an answer is just beginning to raise from several studies (8,9). The problem of nuclear localization of members from the steroid/thyroid hormone receptor superfamily seems to be more complex than in the case of SV40 large T-antigen. Although all members share strong relatedness in structure and activity, it seems that the nuclear localization process depending on the considered receptor: some, like the glucocorticoid receptor are cytoplasmic in the absence of ligand and become nuclear in its presence (10); others, like the estrogen receptor (11) or the retinoic acid receptors (12) are nuclear, even when unliganded.

Many studies at the moment are aimed to explain these differences in nuclear localization regulation. Results agree to define a primary SV40-related NLS in all receptors between the DNA- and hormone-binding regions. However, different regulated transports put forward the presence

of a second NLS (9) except for the oestrogen receptor (13) or dimerization (7) and more recently, the presence of an intrinsic "inactivation function" in certain cases (12).

Indeed, problems of nuclear transport regulation of these hormone-inducible transcription factors are quite essential to elucidate for a good understanding of signal transduction by steroid hormones. By the present study, we would like to settle up with the primary SV40-related NLSs of receptors which are representative of the superfamily as to bring a contribution to further works concerning the regulation(s).

There are different ways to demonstrate a peptidic sequence to be a nuclear localization signal: the presumed NLS region of the nuclear protein can be deleted or mutated in as to show its essential nature; other assays demonstrating that the NLS is sufficient for nuclear targeting use either molecular biology techniques to obtain fusion proteins or chemical crosslinking of the synthetic NLS to a nonnuclear protein. In this study, we used this latter technique as to be able to compare different receptor NLSs out of the receptor context and consequent regulation. We crosslinked Bovine Serum Albumin (BSA) with peptides corresponding to the presumed NLSs from human Retinoic Acid Receptors (hRARs), human Glucocorticoid Receptor (hGR), human Estrogen Receptor (hER), human Androgen Receptor (hAR) and SV40 large T-antigen as a positive control. For introducing the NLS-BSA complexes in the cytoplasm of cultured cells, we chose to make use of viral coinfection, an original method which we have recently described to be useful for this purpose (14).

MATERIALS AND METHODS

Peptide synthesis: The peptides were synthesized on a LKB Biolynx peptide synthesizer using Fmoc-amino acids chemistry (15). After cleavage from the resin and deprotection using aqueous trifluoroacetic acid, peptides were purified by high performance liquid chromatography on a C18 column eluted by 20% aqueous acetonitrile (Yield 79-91%). Purity was assessed by analytical reversed-phase high-pressure liquid chromatography, amino acid analysis, and fast atom bombardment-mass spectroscopy (see Table I).

Peptide-protein conjugation: The synthetic peptides were conjugated to bovine serum albumin (BSA) with a bifunctional cross-linking reagent, m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) according to Lerner (16).

A sample was examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (5-30%) to visualize an increase in molecular weight due to the addition of peptide to the protein. A more accurate estimate of the average degree of conjugation was determined by amino-acid analysis (3 measurements).

Viral co-internalization: The adenovirus serotype 3 (subgroup B) was produced in suspension grown KB cells. Adenovirions were extracted and purified in CsCl gradients, as previously described (17) and the virus titers were determined and expressed as fluorescent focus units (FFU). MCF 7 and HeLa S3 cells monolayers were infected with adenovirus 3 suspensions (300FFU/cell) in Dulbecco medium including BSA or peptide-BSA at 1 mg/ml. After 45 minutes incubation at 37°C (kinetic determined in (13)), cells were prepared for immunofluorescence.

Immunofluorescence: Soon after incubation cells were harvested for immunofluorescence according to Langranger et al. (18). BSA was detected by indirect immunofluorescence using goat anti BSA IgG (Sigma, St Louis, USA). Internalized fluorescent macromolecules were examined in an Olympus BH-2 microscope (excitation wavelength at 530-560nm, edge filter at 580nm). Photomicrographs were taken using a Kodak citachrome film (400 Asa).

RESULTS

Conjugation of the synthetic peptides to a carrier protein: Synthetic peptides were synthesized which contain a 7-9 amino acids core sequence supposed to act as nuclear localization signal and the sequence Cys-Gly-Tyr-Gly was added to the amino terminal as a spacer containing a cysteinyl residue for ease of coupling to a carrier protein (5). They were coupled to bovine serum albumin (BSA) chosen as carrier protein because of its widespread use in similar studies. Coupling was made using the m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) procedure (16). The coupling ratios were estimated by two methods : estimation of the increase in apparent molecular weight of BSA by mobility retardation by sodium dodecyl sulfate polyacrylamide gel electrophoresis and amino acid analysis. The increase in the ratios of residues contained in each complex are reported in table I.

Comparison of nuclear targeting efficiencies: Soon after Adenovirus infection in the presence of BSA or peptide-BSA, cells were returned to a CO₂ incubator for 45 minutes and then immediately harvested for immunofluorescence. Microphotographs were taken with the concern of keeping the same conditions: microscope enlargement, film, aperture and exposure time. Negatives were scanned using a Joyce-Loeble micro-densitometer. For each assay 20 cells were scanned horizontally following the longest line passing through the nucleus, area under the optic density curve was integrated and a nuclear immunofluorescence/cytoplasmic fluorescence value was averaged. The values for each complex were reduced to the corresponding coupling ratios and the nuclear targeting efficiency of SV40 fixed to 1 as to compare with literature data. The other values were then calculated to give for each one a nuclear targeting efficiency/SV40 reported on histogram in fig 1. Results reported indicate that:(i) all presumed NLSs are efficient for nuclear transport, (ii) efficiency values are in the range of the one for SV40 large T antigen, and (iii) there are sensible differences for these values between receptors (e.g 1:2 for hRAR and hGR).

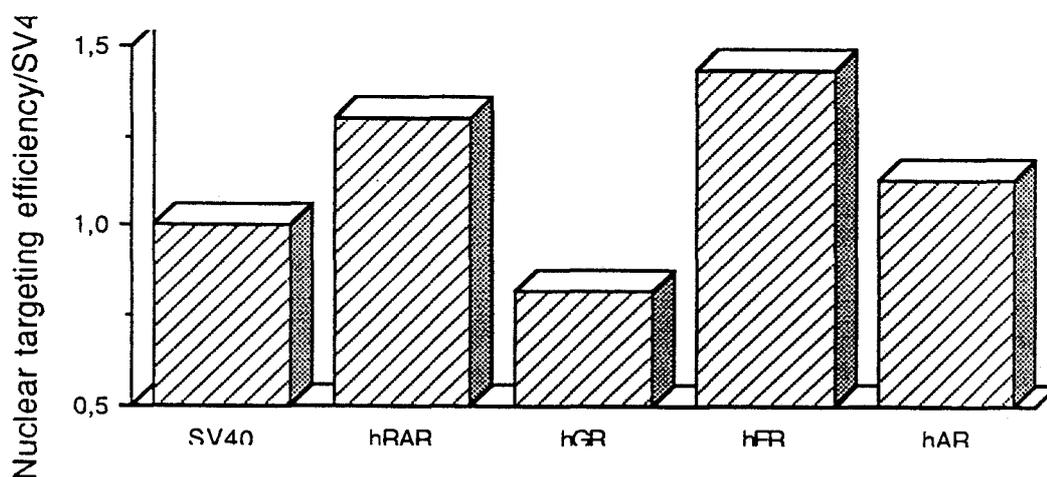


Figure 1: Comparison between nuclear targeting efficiencies of nuclear localization signal peptides. Values are calculated as indicated in materials and methods.

Competition with free peptide: A single competition experiment is described for hRAR-NLS-BSA complex versus free hRAR-NLS, this choice will be discussed further. For assay 0.5, 1, 2, 5 or 10 equivalents of free peptide were co-added with the complex in the medium. An equivalent is calculated taking into account the concentration of the complex multiplied by the coupling ratio. Figure 2 shows a graph in which, for each point, nuclear immunofluorescence/cytoplasmic immunofluorescence ratio was calculated as previously described except that the values were reduced to fix immunofluorescence of complex hRAR-NLS-BSA complex alone to 100%. Results indicate that hRAR-NLS is a good competitor in the conditions used in this work.

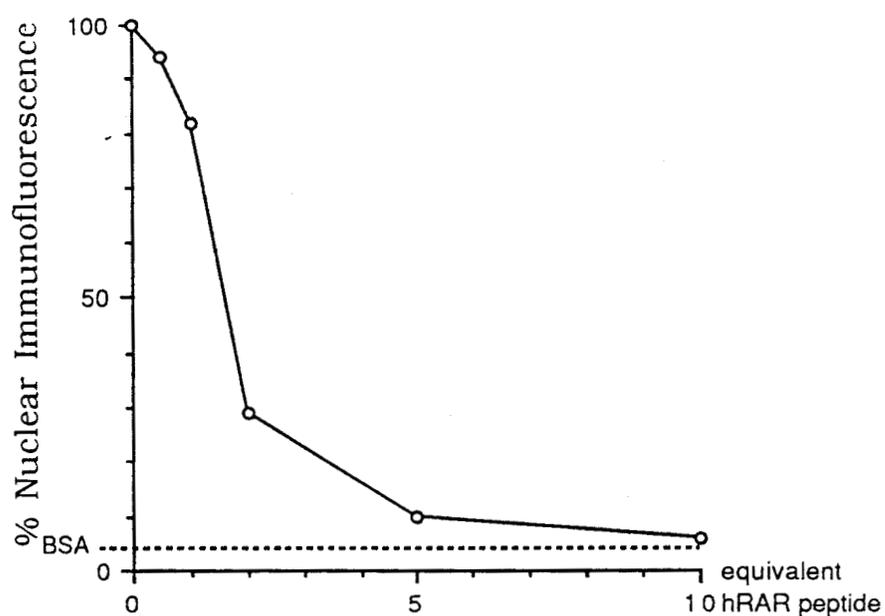


Figure 2: Competition for nuclear entry between hRAR-NLS-BSA complex and free hRAR-NLS. calculations for percentage of nuclear immunofluorescence and equivalent of competitor peptide are described in the text.

DISCUSSION

The work presented here clearly demonstrates that SV40 T-antigen related NLS-sequences of four receptors from the steroid/thyroid hormone receptors superfamily can target a normally nonnuclear protein to the nuclear compartment. The essential character for nuclear entry was already demonstrated for two of them by mutation studies (8,9) while for other members this role was inferential. However, this is the first time to our knowledge that the sufficient nature for nuclear targeting of such peptides, is reported out of the receptor context. Though there are sensible differences between evaluated nuclear targeting efficiencies, no relationship could be made with regulation of nuclear transport for the different receptors.

The competition study was organized for a single peptide hRAR, and this for at least two reasons: (i) we would like to verify if this kind of assay, already described using micro-injection (19), could be applied to the viral co-internalization method, and (ii) retinoic acid receptor is peculiar regarding all other members of the superfamily because in a recent report, Robbins *et al.* (20) put forward the presence of two basic aminoacids ten residues before presumed NLSs in number of nuclear proteins. These are conserved too throughout the superfamily except for retinoic acid receptors. These two basic aminoacids, which are thought to contribute to nuclear localization are pointing out the α -helix consecutive to the second zinc-finger (CII) of the DNA binding domain. However hRAR-NLS seems to be as efficient as other NLSs and the question of the role played by the short conserved basic stretch still remains.

Results from the competition experiment show that, in the conditions used here (45 minutes incubation time), hRAR-NLS alone is able to inhibit nuclear entry of the hRAR-NLS-BSA complex at ratios higher than 5 equivalents. At least, two explanations can be put forward: (i) the effect of position and number of NLSs seem to play an important role on transport (5) and (ii) it has been shown that competition essentially affects the rate of nuclear accumulation (19).

The peptides studied here were chosen to be heterogen, at least in sequence, and thus to be representative for the numerous members of the superfamily. Residues contained in the sequences are, for the majority, positively charged at physiological pH. These positive charges are not equally distributed in each peptide except for the presence of one basic aminoacid at a position that we have made to coincide with the first lysine of SV40 T-antigen. Actually, this position corresponds to the crucial Lys₁₂₈ in SV40 T-antigen which can be replaced (keeping nuclear targeting activity but less efficiency) by any positive charged group (21). Except for a basic residue at this position, we failed in defining a precise consensus because the studied peptides are all active for nuclear localization with positive charges unequally distributed for each. However, our results are in accordance with the widespread hypothesis that nuclear uptake is a receptor-mediated process involving semi-specific interactions. Works are in progress, taking into account the structural organization of such NLSs, in a way to establish a 3-dimensional map of charges repartition which should be important in the protein-protein interaction step that probably triggers the nuclear entry of macromolecules. Designing a simple consensual 3D-structure efficient for caryophylic function should be of utmost interest in targeting xenobiotics to the nucleus for instance.

REFERENCES

1. Dingwall, C., Sharnick, S.V. & Laskey, R.A. (1982) *Cell* **30**, 449-458.
2. Kalderon, D., Richardson, W.D., Markham, A.T. & Smith, A.E. (1984) *Nature* **311**, 33-38.
3. Kalderon, D., Roberts, B.L., Richardson, W.D. & Smith, A.E. (1984) *Cell* **39**, 499-509.
4. Lanford, R.E. & Butel, J.S. (1984) *Cell* **37**, 801-813.
5. Lanford, R.E., Kanda, P., & Kennedy, R.C. (1986) *Cell* **46**, 575-582.
6. Silver, P.A. (1991) *Cell* **64**, 489-497.
7. Wolff, B., Dickson, R.B. & Hanover, J.A. (1987) *TIPS* **8**, 119-121.
8. Guiochon-Mantel, A., Loosfelt, H., Lescop, P., Sokhavuth, S., Atger, M., Perrot-Applanat, M. & Milgrom, E. (1989) *Cell* **57**, 1147-1154.
9. Picard, D. & Yamamoto, K.R. (1987) *EMBO J.* **6**, 3333-3340.
10. Wikström, A.-C., Bakke, O., Okret, S., Brönnegård, M., & Gustafsson, J.-Å. (1987) *Endocrinology* **120**, 1232-1242.
11. King, W.J. & Greene, J.P. (1984) *Nature* **307**, 745-749.
12. Picard, D., Kumar, V., Chambon, P. & Yamamoto, K.R. (1990) *Cell regulation* **1**, 291-299.
13. Evans, R.M. (1988) *Science* **240**, 889-895.
14. Hamy, F., Verwaerde, P., Helbecque, N., Formstecher, P. & Hénichart, J.P. (1991) *Bioconj. Chem.*, in press.
15. Fields, G.B. & Noble, R.L. (1990) *Int. J. Peptide Protein Res.* **35**, 307-348.
16. Lerner, R.A., Green, N., Alexander, H., Liu, F.T., Sutcliffe, J.G. and Shinnick, T.M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3403-3407.
17. D'Halluin, J.C., Milleville, M., Boulanger, P.A. and Martin, G.R. (1978) *J. Virol.* **26**, 344-356.
18. Langranger, G., De Mey, J., Moeemans, M., Daneels, G., De Brabander, M. and Small, J.V. (1984) *J. Cell. Biol.* **99**, 1324-1334.
19. Goldfarb, D.S., Gariépy, J., Schoolnik, G. & Kornberg, R.D. (1986) *Nature* **322**, 641-644.
20. Robbins, J., Dilworth, S.M., Laskey, R.A. & Dingwall, C. (1991) *Cell* **64**, 615-623.
21. Lanford, R.E., White, R.G., Dunham, R.G. & Kanda, P. (1988) *Mol. Cell. Biol.* **8**, 2722-2729.

Un résultat principal qui ressort de cet article est qu'il n'y a pas de différence critique dans l'efficacité d'adressage nucléaire entre les divers signaux au sein de la superfamille ni même entre ces signaux et celui de l'antigène T de SV40. Comme on pouvait le prévoir, la localisation différentielle entre les différents récepteurs antérieurement à leur liaison à l'hormone n'est pas la conséquence de variations dans l'efficacité de leur signal de localisation nucléaire. Il existe d'autres d'hypothèses pour expliquer ces différences. En particulier, l'idée générale selon laquelle la liaison de l'hormone à son récepteur provoquerait chez ce dernier des changements allostériques qui, entre autres, exposeraient ou démasqueraient une partie du récepteur contenant le signal de localisation nucléaire. Pour le récepteur des glucocorticoïdes, cela coïnciderait avec sa dissociation de la protéine de choc thermique HSP90 (Sanchez et al., 1985) encore appelée "protéine chaperon". D'autres hypothèses font intervenir des processus biochimiques tels que la phosphorylation (Picard et Yamamoto, 1987).

L'expérience de compétition entre le complexe signal-sérum albumine et le signal apporte la suggestion que le transport nucléaire est un phénomène médié par un récepteur saturable. De fait, il est de mieux en mieux établi que l'adressage nucléaire fait appel à au moins une étape de reconnaissance moléculaire protéine-protéine.

La preuve expérimentale de l'existence d'un récepteur dans le transport nucléaire a été fournie en premier lieu par Goldfarb et ses collaborateurs (Goldfarb et al., 1986) dans une expérience de compétition avec le signal de l'antigène T de SV40 dont nous nous sommes inspirés pour notre étude. Ces expériences avaient démontré que le transport nucléaire était un processus saturable. L'hypothèse d'un transport faisant intervenir un récepteur a été confirmée pour d'autres découvertes indiquant que l'import des protéines est un processus à deux étapes : une liaison à une (des) protéine(s) dont la localisation n'est pas établie clairement (cytoplasme ou périphérie nucléaire) et une translocation dite "active" car elle est dépendante de la température et nécessite la consommation d'ATP (Newmeyer et Forbes, 1988; Richardson et al., 1988; Breeuwer et Goldfarb, 1990). La diversité des peptides capables de diriger une protéine au noyau fait admettre deux hypothèses pour l'étape d'interaction : soit il existe plusieurs protéines capables de reconnaître différentes variantes de signal, soit une protéine est capable de reconnaître tous les signaux.

Comment cette ou ces protéines reconnaissent-elles les signaux? Des mutations dans ces signaux démontrent qu'une haute teneur en résidus basiques n'est pas suffisante pour provoquer l'adressage nucléaire. Le remplacement de la lysine 128 de l'antigène T de SV40 par d'autres dérivés d'acides aminés à chaîne latérale portant une charge positive ne restaure pas une activité complète (Lanford et al., 1988). De même, un peptide correspondant au signal de l'antigène T de SV40 mais dont la séquence est inversée (rétropeptide) ne se lie pas au récepteur du signal de localisation nucléaire du foie de rat (Adam et al., 1989).

Cependant, au sens large, le signal de localisation nucléaire partage une caractéristique commune aux autres séquences signal : dans aucun des cas, il n'existe une séquence primaire consensus satisfaisante. Les signaux de sécrétion sont généralement hydrophobes, les signaux d'import mitochondrial sont une combinaison de résidus hydrophobes et basiques et les signaux de localisation nucléaire sont en général de caractère basique. De façon similaire, les groupements acides d'activation transcriptionnelle sont globalement de nature acide mais ils ont une faible homologie de

séquence primaire. Tous ces cas tombent sous la coupe du thème récurrent de la reconnaissance semi-spécifique : beaucoup de ligands sont admis, mais d'autres ligands potentiels sont exclus.

Ce concept constitue une motivation pour la continuité à donner à ces études : au vu des séquences primaires diverses des différents signaux (récepteurs hormonaux et autres récepteurs nucléaires) aptes à la caryophilie ainsi que des séquences apparentées mais qui ne fonctionnent pas, il serait pertinent de placer l'étude au niveau de la structure de ces séquences. En effet, il est de plus en plus clair que la reconnaissance moléculaire fait intervenir des séquences de macromolécules biologiques, mais aussi et surtout la structure tridimensionnelle agençant ces séquences. Dans le cas des signaux de localisation nucléaire, il ne serait pas surprenant que le positionnement des charges positives dans l'espace joue un rôle prépondérant. Il serait alors intéressant d'utiliser les techniques de modélisation moléculaire assistée par ordinateur afin de superposer les structures tridimensionnelles (pour celles qui sont connues) de ces signaux ainsi que leurs caractéristiques électroniques. Au moins dans le cas d'une superfamille comme celles des récepteurs hormonaux nucléaires, il est permis d'imaginer qu'en matière de signal d'adressage nucléaire il doit exister, non pas une séquence consensus, mais une structure tridimensionnelle consensus possédant une répartition de charges positives bien définie. Dès lors, on pourrait envisager la conception rationnelle d'une structure de signal, peptidique ou non, capable de conduire au noyau toute molécule à laquelle il aura été préalablement conjugué. Cela mettrait à la disposition des chercheurs un outil de vectorisation nucléaire utile dans l'élaboration de xénobiotiques ayant une action recherchée sur l'ADN par exemple.

Nous allons traiter dans une deuxième partie de la reconnaissance de l'ADN par les protéines de la superfamille des récepteurs des hormones stéroïdiennes/thyroïdiennes. Nous pourrions nous apercevoir que dans ce cas aussi, la séquence et la structure sont au moins aussi importantes dans cette interaction moléculaire spécifique.

ELEMENTS PEPTIDIQUES
DE
RECONNAISSANCE
D'UNE
SEQUENCE D'ADN

Des motifs de liaison à l'ADN.

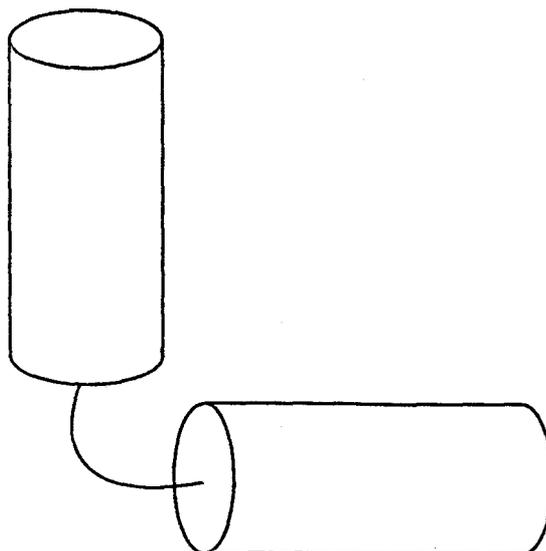
Les protéines reconnaissent l'ADN de la même façon qu'elles reconnaissent d'autres protéines c'est-à-dire qu'elles forment en premier lieu un contour ou forme tertiaire qui doit être compatible à la surface avec laquelle elles doivent interagir. Elles s'offrent ainsi l'opportunité d'un contact suffisamment proche et sur une surface suffisamment importante pour établir de nombreuses interactions. Comme dans le cas d'une interaction hormone-recepteur, enzyme-substrat, antigène-anticorps ou protéine-protéine, les contacts atomiques entre une protéine et l'ADN sont de natures variées comprenant des liaisons de type hydrogène, des interactions ioniques et des contacts de van der Waals. L'apport le plus important dans la compréhension de la spécificité de liaison entre une protéine et l'ADN est celui des études cristallographiques par diffraction de rayons X. Ce sujet ne pouvant être détaillé dans ce mémoire nous référons le lecteur à la revue de Pabo et Sauer (Pabo et Sauer, 1984).

Bien que la compréhension du phénomène de reconnaissance protéine-ADN passe nécessairement par des études structurales complexes, le problème s'est trouvé grandement simplifié par l'observation que la plupart des protéines reconnaissent l'ADN par des motifs structuraux communs. Nous allons maintenant faire la description de trois grandes classes de ces motifs : le motif hélice-coude-hélice, la crémaillère à leucine et les doigts de zinc.

Le motif hélice-coude-hélice.

Les structures de trois protéines liant l'ADN de manière spécifique ont été résolues en un laps de temps relativement court au début des années 1980. Chacune est une protéine de régulation génomique dans les cellules bactériennes; deux sont des petites protéines de liaison à l'ADN codées par le phage lambda (Cro et CI) (Anderson *et al.*, 1981; MacKey et Steitz, 1981), et la troisième est la protéine CAP (Pabo et Lewis, 1982) d'*Escherichia coli*. On sait que chacune de ces protéines se lie à l'ADN sous la forme d'un dimère et que le site oligonucléotidique reconnu est une séquence palindromique (de symétrie rotationnelle). A partir de la résolution de leur structure tridimensionnelle, il est apparu que leurs domaines de liaison à l'ADN ont en commun une succession distinctive de deux structures en hélice α séparées par un coude β prononcé ($\sim 90^\circ$) (107). Cette structure est appelée le motif hélice-coude-hélice (figure 6).

*Figure 6. Le motif hélice-coude-hélice
(Les cylindres représentent
les structures en hélice α).*



La symétrie et la position relative des deux éléments du dimère provoquent un positionnement correct pour l'interaction avec l'ADN (distance de 34 Å correspondant à un tour d'hélice d'un duplex d'ADN). Il a donc été proposé qu'une hélice de chaque motif viennent se lover dans deux grands sillons de la double hélice avec la symétrie d'ordre 2 du dimère coïncidant avec la symétrie d'ordre 2 de la séquence opérateur reconnue. Au sein d'un monomère, une des deux hélices (figure 7) se positionne dans le grand sillon établissant les interactions spécifiques avec les bases ; elle est nommée hélice de reconnaissance (Pabo et Sauer, 1984).

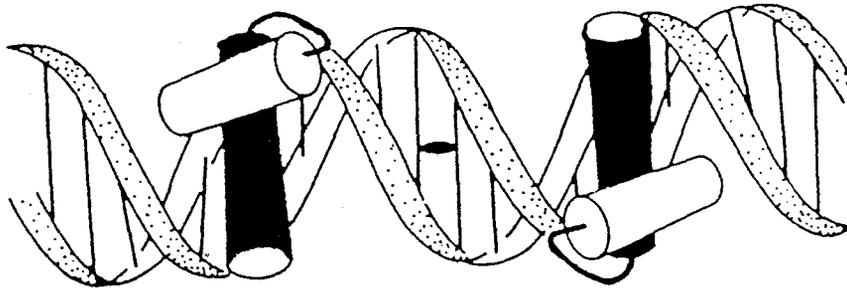


Figure 7. Interaction du motif hélice-coude-hélice de la protéine Cro du phage lambda avec l'ADN.

L'extrémité N-terminale de chacune des hélices pointe vers le squelette phosphate-sucré utilisant la demi-charge positive induite par le dipôle hélice pour le positionnement correct de l'hélice de reconnaissance. On présume alors que les chaînes latérales des résidus font saillie à la surface de l'hélice de reconnaissance et sont alors capables de contracter les interactions électrostatiques spécifiques avec les groupements de la succession de plateaux de paires de bases. Le motif hélice-coude-hélice a été identifié dans de très nombreuses protéines de régulation d'expression et de la différenciation tels que produits des gènes homéotiques chez les insectes.

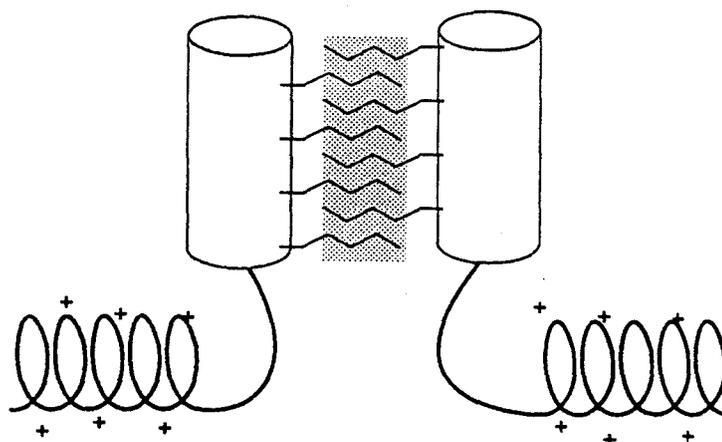
La crémaillère à leucine.

L'existence d'une autre catégorie de motif protéique de liaison à l'ADN fut en premier lieu mise à jour par une étude (Vogt *et al.*, 1987) faisant remarquer des similitudes dans la séquence primaire de plusieurs produits d'oncogènes (jun, myc, fos) et d'une protéine régulatrice d'expression chez la levure (GCN4), substances capables de reconnaître spécifiquement une séquence d'ADN. La structure sous-jacente commune à ces séquences a émergé grâce à la considération d'une autre protéine apparentée : C/EBP. C/EBP est une protéine thermostable que l'on trouve dans le noyau des cellules de foie de rat et qui fut initialement caractérisée en raison de ses capacités de liaison à des séquences d'ADN régulatrices associées à des gènes viraux (Graves *et al.*, 1986; Johnson *et al.*, 1987). Le clonage et le séquençage du gène codant pour C/EBP a conduit à la découverte que son domaine de liaison à l'ADN est similaire à celui des protéines transformantes Fos, Myc, et Jun

(Landschulz *et al.*, 1988). Les régions de similitude entre ces protéines sont relativement dépourvues de résidus incompatibles avec une hélice α (Pro, Gly, Tyr, Asn) (Vogt *et al.*, 1987; O'Neil *et al.*, 1990; Sakonju *et al.*, 1990). Sur la seule base de cette prétendue permmissivité d'hélice, la séquence en acides aminés du domaine de liaison à l'ADN de C/EBP fut disposée sur une hélice α idéalisée (Sakonju *et al.*, 1990). Cet exercice provoqua deux observations inattendues. Premièrement, cela montra qu'une région de 35 acides aminés représentant environ la moitié du domaine de liaison à l'ADN de C/EBP contient une répétition stricte de résidus leucine tous les 7 acides aminés (deux tours d'une hélice α). Deuxièmement, cela révéla une haute densité en acides aminés de charges opposées (basique et acide) juxtaposés de manière propice à un appariement intra-hélice. Cette haute fréquence de liaisons ioniques laissait donc prévoir pour ce domaine une structure hélicoïdale de stabilité peu commune. Chacune de ces propriétés est partagée par les protéines fos, jun, myc et GCN4 (Sakonju *et al.*, 1990).

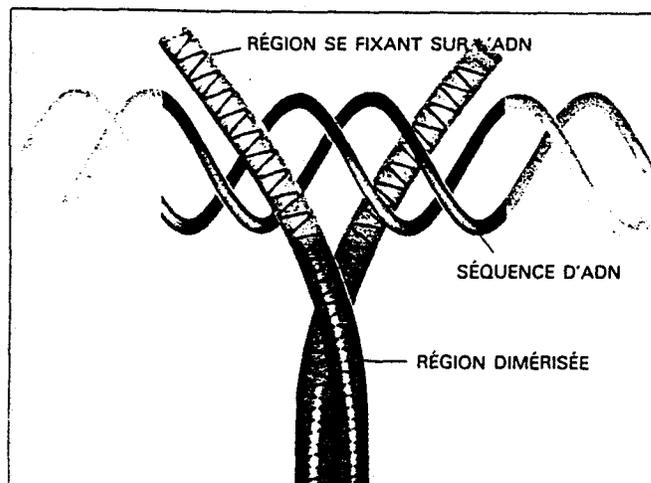
Le déploiement des résidus leucine tous les sept acides aminés est réminiscent d'une autre répétition d'ordre 7, que l'on retrouve dans des protéines de structure quaternaire arrangées en torsade d'hélices (ex: kératines, laminines et l'extrémité des chaînes lourdes de myosine). Ceci suggéra l'hypothèse que les polypeptides de cette classe pourraient se dimériser via l'hélice à répétition de leucines (Sakonju *et al.*, 1990). Cependant, à l'inverse des protéines standard à torsade d'hélice, l'interface putative de dimérisation de C/EBP et des protéines apparentées reposerait uniquement sur des interactions entre des leucines, à l'exception d'autres acides aminés. Une évaluation effectuée sur 7 protéines possédant ce motif montre une seule exception où une leucine est remplacée par une méthionine dans la protéine humaine l-myc. A part leur caractère hydrophobe général, les chaînes latérales de la leucine et de la méthionine ont en commun leur longueur et le fait qu'elles ne portent pas de groupement méthyle en ramification sur le carbone β (comme la valine ou l'isoleucine). Ces points communs conduisent à l'hypothèse que les chaînes latérales des leucines qui s'étendent sur la même surface de l'hélice forment des interdigitations intimes avec les leucines d'une hélice analogue d'un deuxième polypeptide formant une interface verrouillée à la manière d'une fermeture à glissière, d'où le nom anglo-saxon de "leucine-zipper" traduit en français par crémaillère à leucine (figure 8) (Sakonju *et al.*, 1990).

Figure 8. Le motif crémaillère à leucine. La zone grisée représente la zone d'interactions hydrophobes entre les deux hélices à leucines répétées.



La crémaillère à leucine fait partie intégrante du domaine de liaison à l'ADN de C/EBP. Lorsqu'on procède à des mutations ou des délétions dans la répétition de leucine de C/EBP, la protéine devient incapable d'interagir spécifiquement avec l'ADN. Cependant, la crémaillère à leucine n'est pas à proprement parler le motif de liaison à l'ADN de ces protéines. Un segment de 30 acides aminés situé directement en N-terminal de la répétition de leucines doit aussi demeurer intact pour que la protéine lie l'ADN. Ce segment est caractérisé dans chaque protéine par un haut degré de basicité. C'est ce segment basique qui entre en contact spécifique avec l'ADN par le grand sillon de la double hélice. La dimérisation facilitée et stabilisée par la crémaillère à leucine induit un positionnement bilatéral particulier qui est déterminant pour l'interaction du dimère avec la séquence reconnue. Des études de dichroïsme circulaire indiquent que les régions basiques ont une nette propension à être hélicoïdales en solution alors que la structure est totalement en hélice α lors de l'interaction avec l'ADN. Un modèle fort répandu de l'interaction est celui de la "prise en ciseaux" (figure 9) où l'ADN est embrassé par le dimère de part et d'autre de la double hélice, la "prise" étant verrouillée par la torsade d'hélice de la crémaillère à leucine.

ANCIEN MODÈLE DE FIXATION SUR L'ADN



MODÈLE ACTUEL

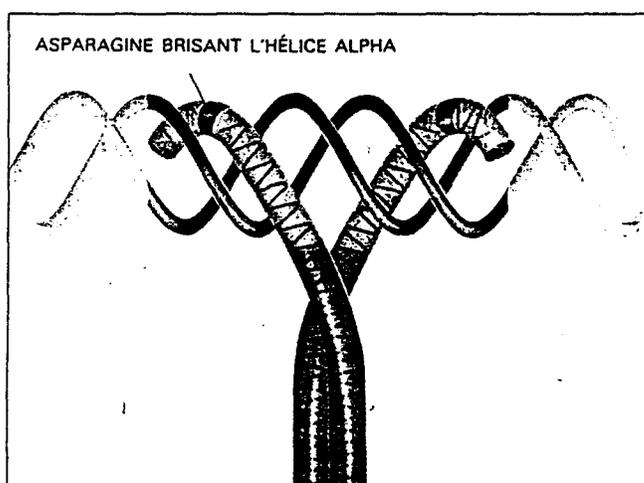


Figure 9. Interaction crémaillère à leucine-ADN

Ce modèle correspond bien à la symétrie rotationnelle observée dans certaines séquences oligonucléotidiques reconnues par ces motifs. L'hypothèse est de plus en plus admise que ce motif de crémaillère à leucine pourrait former des hétérodimères (en particulier jun/fos), ce qui aurait pour effet de multiplier les combinaisons de reconnaissance de larges séquences d'ADN.

Les motifs doigts de zinc.

La découverte du motif doigt de zinc est venue en conséquence d'études sur l'expression des gènes 5S de l'ADN ribosomal de *Xenopus laevis*. La transcription des gènes 5S est régulée par un promoteur interne (Bogehagen *et al.*, 1980; Miller *et al.*, 1985) auquel se lie une protéine appelée facteur de transcription IIIA (TFIIIA) (Engelke *et al.*, 1980; Sakonju et Brown, 1981). TFIIIA, en association avec deux autres protéines (TFIIIB et TFIIIC), forme un complexe stable avec le gène 5S, qui devient la cible de fixation de l'ARN polymérase III (Segall *et al.*, 1980; Shastry *et al.*, 1982).

Une grande avancée dans la compréhension de la nature de l'interface entre l'ADN et TFIIIA est venue de la découverte d'un motif répété au sein de sa séquence en acides aminés (Miller *et al.*, 1985; Brown *et al.*, 1985). On observait à neuf reprises un positionnement séquentiel et ordonné de résidus cystéine et histidine. (Cys-X₂-Cys-X₁₂₋₁₅-His-X₂₋₄-His). Ces observations, associées à des résultats antérieurs qui montraient que TFIIIA complexait le zinc (Hanas *et al.*, 1983), ont poussé Klug et ses collaborateurs à postuler que chaque répétition séquestre un seul ion Zn⁺⁺ via une coordination tétraédrique par quatre résidus spatialement conservés (2 cystéines et 2 histidines; figure 10).

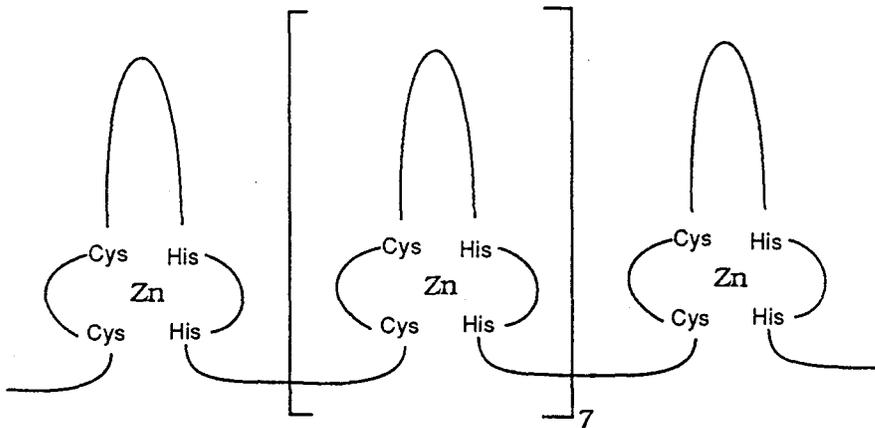


Figure 10. Représentation schématique du motif doigt de zinc C₂H₂ (type TFIIIA)

L'hypothèse fut émise par la suite que les 12 à 14 acides aminés intervenant entre les cystéines et histidines forment une boucle propre à faciliter une interaction spécifique avec le grand sillon de l'ADN (Fairall *et al.*, 1986; Rhodes et Klug, 1986). Bien qu'à ce moment précis il n'y eut pas de données cristallographiques disponibles pour mieux préciser le modèle du doigt de zinc, les résultats d'autres moyens d'investigation venaient confirmer cette interprétation. Par exemple la coordination tétraédrique du zinc par les doigts successifs de TFIIIA a été documentée par la diffraction de rayons X (Diakun *et al.*, 1986). L'interaction séquence-spécifique entre TFIIIA et son site de liaison sur l'ADN 5S s'est avérée être dépendante de la présence de zinc (Hanas *et al.*, 1983). Des dérivés obtenus par mutagenèse dirigée perdent séquentiellement leur capacité de liaison à l'ADN en corrélation avec une perte séquentielle dans la répétition des doigts de zinc (Vrana *et al.*, 1988).

Enfin, plusieurs expériences de protection de l'ADN par TFIIIA ont permis d'établir un modèle d'interaction entre TFIIIA et le grand sillon de l'ADN de part et d'autre de la double-hélice (figure 11) (Rhodes et Klug, 1986).

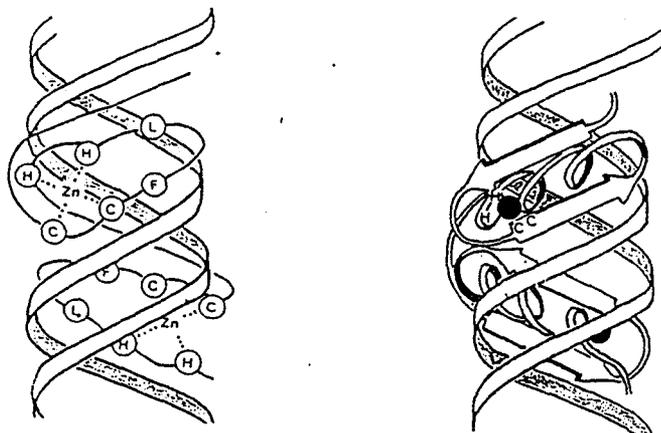


Figure 11. Schéma de l'interaction entre TFIIIA et l'ADN 5S

La distribution spatiale des résidus de cystéine et d'histidine rencontrée dans la séquence de TFIIIA était suffisamment distinctive pour permettre l'identification rapide d'autres protéines à doigt de zinc (ou protéines dactyles). Des constellations de cystéines et d'histidines organisées de manière hautement similaire à celles rencontrées dans TFIIIA (Cys₂/His₂) ont été repérées dans de nombreuses autres protéines de régulation chez la levure, l'insecte et l'homme. En outre, la constance en positionnement d'acides aminés du motif doigt de zinc a généré la synthèse d'oligonucléotides qui, à leur tour, ont permis de cloner des protéines dactyles dont, pour certaines, la fonction est toujours inconnue.

Un motif de séquence en acides aminés fortement apparenté (mais distinct) à l'arrangement Cys₂/His₂ des doigts de zinc du type TFIIIA est observé dans une large variété de protéines de liaison à l'ADN, allant des protéines de régulation chez la levure (Weinberger *et al.*, 1985; Green *et al.*, 1986; Greene *et al.*, 1986; Arrizza *et al.*, 1987; Weinberger *et al.*, 1986; Miesfeld *et al.*, 1986; Jeltsch *et al.*, 1986) jusqu'aux récepteurs hormonaux chez les mammifères (Laughon et Gesteland 1984, Kammerer *et al.*, 1984; Messenguy *et al.*, 1986; Baum *et al.*, 1987; Wray *et al.*, 1987). Ce motif apparenté se caractérise par des groupes de 4 cystéines arrangées en deux paires où les cystéines sont espacées de 2 à 5 acides aminés. Comme dans le cas du paradigme TFIIIA, ces motifs Cys₂-Cys₂ interviennent dans des groupes de 35 à 40 résidus qui ont tendance à être répétés. Dans ces protéines, la délétion d'un groupe Cys₂/Cys₂ ou la substitution d'une cystéine par un autre acide aminé élimine la capacité de liaison à l'ADN.

Le motif de liaison à l'ADN des récepteurs hormonaux nucléaires.

L'aspect le plus frappant dans la séquence des domaines de liaison à l'ADN des récepteurs hormonaux nucléaires est sans doute la conservation des résidus cystéine. Une comparaison des séquences en acides aminés du domaine de liaison à l'ADN (région C) des différents récepteurs (figure 12) révèle une identité certaine et une similitude entre ces molécules divergentes d'un point de vue évolutif. Sur 65 résidus, 20 sont invariants, 7 autres sont conservés dans 7 molécules sur 8 et plus de la moitié des résidus ont leur nature conservée (hydrophobe, basique...) dans les 5/8^e des récepteurs. Neuf des résidus invariants sont des cystéines (Figure 12).

a)

Récepteur	Séquence	
	← CI →	← CII →
hTR2 102	C V V C G D K A T G Y H Y R C I T C E G C	C K Y E G K C V I D K V T R N Q C Q E C R F K K C
hGR 421	C L V C S D E A S G C H Y G V L T C G S C	C A G R N D C I I D K I R R K N C P A C R Y R K C
hMR 603	C L V C G D E A S G C H Y G V V T C G S C	C A G R N D C I I D R I R R K N C P A C R L Q K C
hER 185	C A V C N D Y A S G Y H Y G V W S C E G C	C P A T N Q C T I D K N R R K S C Q A C R L R K C
hPR 567	C L I C G D E A S G C H Y G V L T C G S C	C A G R N D C I V D K I R R K N C P L C R L R K C
hAR 324	C L I C G D E A S G C H Y G A L T C G S C	C A S R N D C T I D K F R R K N C P S C R L R K C
cVDR 37	C G V C G D R A T G F H F N A M T C E G C	C P F N G D C K I T K D N R R H C Q A C R L K R C
hRAR ₅₈	C F V C Q D K S S G Y H Y G V S A C E G C	C H R D K N C I I N K V T R N R C Q Y C R L Q K C

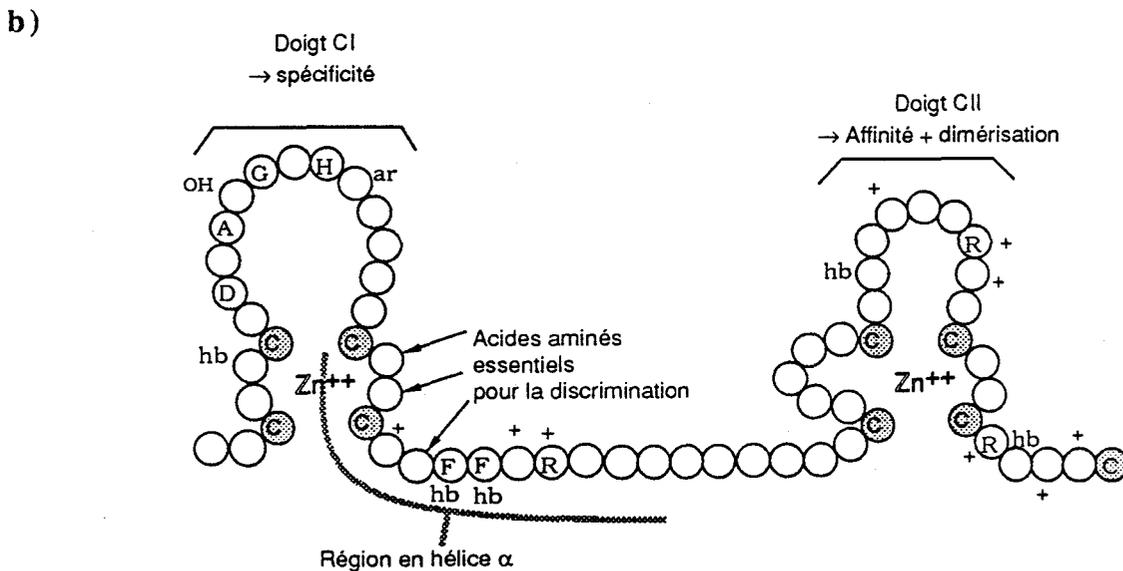


Figure 12. (a) Comparaison des séquences en acides aminés du domaine C de liaison à l'ADN de plusieurs récepteurs nucléaires. (b) Organisation putative en deux motifs liant le zinc (CI et CII) Les acides aminés conservés sont représentés par le code à une lettre. Les natures conservées sont indiquées : +=nature basique; hb=hydrophobe; OH=hydroxylé; ar=aromatique

Le positionnement de ces résidus est très similaire à celui du motif observé dans le facteur de transcription TFIIIA dans lequel des unités répétées contenant des cystéines et des histidines s'organisent en une structure en forme de doigt coordonnée par un atome de zinc.

En fait le positionnement des introns dans les gènes de la superfamille confirme que cette région C peut être séparée en deux parties. Chaque partie de la région C contient un motif similaire au doigt de zinc de TFIIIA (voir plus-haut) excepté le fait que l'ion métallique, dans le cas des récepteurs hormonaux est coordonné tétraédriquement par deux paires de cystéines par motif plutôt qu'une paire de cystéines et une paire d'histidines. L'utilisation de la spectroscopie EXAFS (Extended X-ray Absorption Fine Structure) a confirmé que cette région du GR de rat complexait effectivement deux atomes de zinc (Diakun *et al.*, 1986).

Contrairement au TFIIIA où les séquences de chaque doigt peuvent être alignées pour générer une séquence consensus, les deux motifs liant le zinc diffèrent sensiblement dans les récepteurs hormonaux. De fait, si les doigts de cette partie sont apparus par duplication, ils ont considérablement divergé depuis.

La première partie de la région C (CI) contient quatre cystéines conservées et une forte proportion d'acides aminés hydrophobes, tandis que la deuxième partie de la région contient cinq cystéines conservées et une forte proportion d'acides aminés basiques.

La région C est responsable de la spécificité d'activation des gènes cibles. Ceci fut démontré par l'utilisation de récepteurs chimères dans lesquels la région C d'un récepteur était remplacée par la région C d'un autre récepteur. Par exemple, un ER contenant la région C d'un GR active spécifiquement des gènes de réponse aux glucocorticoïdes mais cette fois sous la dépendance de la présence d'œstradiol.

Bien que chacune des deux parties de la région C des récepteurs soit nécessaire pour la liaison des récepteurs à l'ADN avec une affinité correcte, d'autres constructions de récepteurs chimères dans lesquelles des échanges individuels de doigts avaient été opérés, indiquent que la région contenant le premier doigt (CI) détermine la spécificité du gène cible du récepteur. Le deuxième doigt interviendrait alors dans l'affinité du récepteur pour l'ADN par des interactions non-spécifiques avec la double hélice (résidus basiques conservés) et par interaction avec une autre molécule de récepteur pour une fixation en dimère hautement coopérative (figure 12).

Les domaines de liaison de l'ADN des ER, PR et GR ont été exprimés chez *E. coli* et purifiés. Leur liaison spécifique à leur HRE respectif indique que ni d'autres protéines, ni des modifications post-traductionnelles ne sont requises pour la reconnaissance spécifique de l'ADN.

Au moment où notre étude a débuté, on ne connaissait relativement que peu de choses concernant les aspects structuraux et fonctionnels du domaine de liaison à l'ADN des récepteurs nucléaires : la spectroscopie d'absorption des rayons X (Diakun *et al.*, 1986) a montré que le domaine de liaison à l'ADN du récepteur des glucocorticoïdes lie le zinc via une coordination tétraédrique par des atomes de soufre. La technique d'échange de doigts a démontré que la région en doigt de zinc putatif N-terminale du domaine (CI, figure 12) détermine à elle seule la spécificité d'un récepteur envers son élément de réponse hormonal.

Nous avons choisi le doigt de zinc CI du récepteur humain des glucocorticoïdes comme modèle de fragment peptidique capable de reconnaître spécifiquement une séquence donnée d'ADN. Il est possible d'énumérer quelques-uns des critères qui ont motivé ce choix :

Le motif doigt de zinc : La reconnaissance moléculaire passe par la reconnaissance d'une séquence de la protéine mais aussi et surtout par sa structure tridimensionnelle. La complexation tétraédrique du zinc (liaisons stables) doit nécessairement apporter une contrainte structurale du motif et donc une structure rigide que l'on devrait retrouver dans un peptide synthétique extrait du contexte de la protéine.

Le domaine de liaison à l'ADN des récepteurs hormonaux : Le mécanisme d'action des récepteurs hormonaux fait l'objet de nombreuses études en particulier concernant l'interface protéine-ADN.

Le doigt de zinc putatif CI : La région N-terminale des récepteurs hormonaux putativement repliée en doigt de zinc est un motif unitaire relativement court (abordable par chimie de synthèse) qui semble directement impliqué dans une interaction spécifique avec une séquence donnée d'ADN. De plus, la permutation de reconnaissance entre deux HREs dépend de peu d'éléments (3 acides aminés).

Le récepteur humain des glucocorticoïdes : il est, à ce moment de l'étude, le mieux étudié des récepteurs nucléaires. La protéine partiellement purifiée est disponible par collaboration avec le groupe de Pierre Formstecher. L'action ubiquitaire des glucocorticoïdes sur l'organisme pourrait faire intervenir en toile de fond un éventuel intérêt thérapeutique.

Cette partie de l'étude consacrée à la reconnaissance d'une séquence d'ADN par un motif peptidique débute donc par la synthèse d'un peptide correspondant au doigt de zinc putatif N-terminal (CI, figure 12) du domaine de liaison à l'ADN du récepteur humain des glucocorticoïdes. Utilisant diverses techniques, nous avons pu obtenir des indices concernant la complexation du zinc, le repliement consécutif du motif et enfin la démonstration de son interaction spécifique avec une séquence oligonucléotidique correspondant à l'élément de réponse aux glucocorticoïdes. Ceci a fait l'objet de la rédaction d'un article "*A synthetic N-terminal zinc-finger of the glucocorticoid receptor binds DNA with moderate affinity, but in a sequence-specific manner.*" dont le manuscrit est en cours d'évaluation à "Journal of Molecular Recognition".

A SYNTHETIC N-TERMINAL "ZINC-FINGER"
OF THE GLUCOCORTICOID RECEPTOR BINDS DNA WITH
MODERATE AFFINITY, BUT IN A SEQUENCE-SPECIFIC MANNER.

F. HAMY¹, P.M. DANZE², N. HELBECQUE¹, J.P. HENICHART¹
and P. FORMSTECHE^{2,3}

¹INSERM U16, Place de Verdun, 59045 LILLE CEDEX
and ²LABORATOIRE DE BIOCHIMIE STRUCTURALE, Place de Verdun
59045 LILLE CEDEX (France)

³To whom correspondance should be addressed.

Running Title: Sequence-specific DNA-binding by zinc-fingers

Key-words: DNA binding, Glucocorticoid receptor,
Hormone response element, Peptide synthesis,
Zinc finger.

SUMMARY

Steroid hormone receptors can activate specific gene transcription after binding as hormone-receptor complexes to short DNA enhancer-like DNA sequences called hormone response elements. The receptor DNA binding domain comprises two putative "zinc finger peptides". We have studied the binding affinity and specificity to DNA of a synthetic 29 residues peptide corresponding to the N-terminal zinc finger of the glucocorticoid receptor (GR). Circular dichroism and $^1\text{H-NMR}$ showed a change of the peptide conformation in the presence of Zn^{2+} . Binding to a palindromic glucocorticoid response element (GRE) oligonucleotide was analyzed by nitrocellulose filter assay and demonstrated to be zinc-dependent and both peptide and DNA sequence specific. The N-terminal zinc finger peptide from the GR was able to bind GRE with a 5-fold higher affinity than the corresponding peptide from the retinoic acid receptor (RAR). K_d values were 0.4 and 2.0×10^{-6} M respectively. No GRE band-shift could be observed with the synthetic peptides in gel retardation experiments. However, the GR peptide was able to specifically inhibit, in a competitive manner, the band-shift observed with an activated wild type GR complex. With this method K_d values in the 10^{-10} and 10^{-5} M ranges were found for the native GR and the zinc finger peptides respectively. Finally, hydroxyl radicals foot-printing experiments revealed a twofold symmetric binding of either the activated GR or the peptide on half sites of GRE. All these data indicate that, despite a far lower affinity for GRE than the intact GR, the binding to DNA of a single synthetic zinc finger remains sequence-specific.

INTRODUCTION

The glucocorticoid receptor (GR) belongs to the steroid and thyroid receptors superfamily (Evans, 1988). Like the other members of this family, the glucocorticoid receptor possesses a DNA-binding domain which interacts with specific DNA target sequences called glucocorticoid responsive element (GRE) (Green and Chambon, 1988 ; Beato, 1989). Hormone receptor complex binding to GRE leads to transcriptional activation of specific gene expression by glucocorticoids. The DNA-binding domain, highly conserved in the steroid receptor superfamily, is rich in cysteines and contains two DNA-binding zinc finger motifs (Klug and Rhodes, 1987 ; Evans and Hollenberg, 1988). The presence of these two zinc fingers, encoded by separate exons of the receptor gene (Green and Chambon, 1988) appears required for correct DNA-binding and transcriptional activation (Hollenberg et al. 1987; Freedman et al. 1988). However, the functional importance of each finger is not the same. Indeed, finger swapping experiments demonstrated that the amino-terminal zinc finger determines target gene specificity (Green et al., 1988), i.e. discrimination between consensus sequences of the GRE and the ERE (estrogen responsive element) group (Beato, 1985). Three aminoacids located at the C-terminal side of this finger have been reported to play a key role in this specificity (Mader et al., 1989 ; Danielsen et al., 1989 ; Umesono and Evans, 1989). It has been suggested that the receptor binds to DNA as a dimer, with the first zinc finger lying in the major groove and recognizing one half of the palindromic hormone responsive element (HRE) in a base-specific way, whereas the C-terminal finger contacts the sugar-phosphate backbone of the flanking sequences (Green et al., 1988; Chalepakis et al., 1988). Direct demonstration of the role of zinc in correct folding and DNA-binding of a recombinant 19 kD protein encompassing the DNA-binding domain of the glucocorticoid receptor has recently been afforded (Freedman et al., 1988). By using EXAFS (extended X-ray absorption fine structure) and visible light spectroscopy, each Zn atom was found to be coordinated in a tetrahedral arrangement by four cysteines. The cysteine residues implicated in metal binding have been established by mutational analysis (Hollenberg and Evans, 1988 ; Severne et al., 1988). Very recently, a solution structure of a 71-residue GR fragment containing the two "zinc finger" domains has been determined by nuclear magnetic resonance spectroscopy (NMR) and distance geometry (Härd et al., 1990b). A model of the complex between this receptor fragment and the GRE has been proposed in which specific DNA binding involved the knuckle of the N-terminal zinc finger and a short α -helical region extending from the C-terminal end of the knuckle. Since all these data suggested a predominant role of the N-terminal zinc finger in GRE binding, we decided to synthesize this peptide and to study its conformation and GRE binding activity in the presence of zinc. Our aim was to see if a single finger peptide, could be able to bind to DNA in a sequence-specific manner. In order to test the specificity of the binding we used not

only unrelated peptides and oligonucleotides, but also a synthetic N-terminal zinc finger peptide from the retinoic acid receptor (RAR), another member of the nuclear receptor family (Evans, 1988).

MATERIALS AND METHODS.

Peptides synthesis.

All peptides were synthesized using the stepwise solid-phase method on a polystyrene-divinylbenzene matrix (Barany et al., 1987) and the Fmoc procedure (Sabatier et al., 1987). Purification was achieved using gel filtration on Bio-Gel P6 in 35 % aqueous acetic acid followed by standard reversed-phase high performance liquid chromatography on a C18 column. Peptide composition was confirmed by amino acid composition. Tert-butylthio group was used for protection of cysteines 88, 91, 105 and 108, and cysteines 421, 424, 438 and 441 in the hRAR α , and hGR "zinc finger peptides" respectively. In the latter peptide, cysteine 431 was protected using an acetamidomethyl group and used in the protected form in all the experiments reported here.

Spectrometry.

The circular dichroism spectra were recorded with a Jobin-Yvon dichrograph R.J. Mark III in a 0.01 cm pathlength cell. The ellipticity was expressed in degrees.cm².dmole⁻¹. Peptide concentrations were 0.2 mM. The peptide was studied in a 20 mM Tris buffer (pH 7.2) containing either 1 mM EDTA or 6 mM ZnCl₂. The NMR spectra were recorded on a Bruker AM 400 WB spectrometer using the same buffers as above. Peptide concentrations were 4 mM.

Crude cytosolic receptor preparation.

Cytosol containing glucocorticoid receptor was prepared from human lymphocytes. Cells were purified using Ficoll gradient, harvested, washed in PBS and lysed in buffer A (10 mM Tris, 1 mM EDTA, 60 mM NaCl, 1 mM PMSF, 10 % glycerol, 0.35 Triton X-100, pH 7.4).

The lysate was centrifuged at 45,000 rpm for 30 min at 4°C. The supernatant was collected and incubated with 1 mM dexamethasone at 4°C overnight. Just before use for gel retardation assay, cytosolic extracts were further incubated at 25°C for 30 min.

Purification of transformed hormone receptor complexes.

Adrenalectomized male Wistar rats were killed by cervical dislocation. Livers were removed and perfused with ice cold buffer A (10 mM Tes, 1 mM EDTA, 60 mM NaCl, 1 mM dithiothreitol, 1 mM PMSF, 10 % glycerol, pH 7.4). They were homogenized with one volume of buffer A (wt/vol) using teflon-glass Potter homogenizer, and centrifuged at 55,000 rpm for 60 minutes at 4°C. The supernatant was recovered by aspiration and the pH adjusted to 7.4.

Before incubation, cytosols were layered onto columns (1 x 3cm) filled with DNA cellulose and equilibrated with buffer A. The excluded fractions were pooled and incubated with 30 nM steroid (triamcinolone acetonide or RU 486), 16 hours at 4°C. Control cytosol extracts were adsorbed on a glucocorticoid receptor-specific affinity matrix as previously described (Idziorek et Al., 1985), in order to obtain glucocorticoid receptor-free cytosolic samples. All purification steps were performed at 4°C. After 16 hours, one part of each aliquot was heated at 25°C during 30 min and complemented with 10 mM molybdate. The other part was immediately incubated with 10 mM molybdate.

The aliquots were then chromatographed on DEAE Trisacryl columns equilibrated with buffer A in the presence of 10 mM molybdate (Danzé et al., 1990). The excluded fraction was collected. Protein concentration was determined by the method of Bensadoun and Weinstein, 1976, before use for band shift assay.

Nitrocellulose binding assay.

Unlabeled synthetic oligonucleotides were purchased from Institut Pasteur (Paris). The synthetic oligonucleotide 5'GATCTCTGGTACAATCTTGTTCTA 3' (GRE) was kinased according to Maniatis et al., 1982. Finger peptides were dissolved in buffer B (20 mM Tris-HCl, 1 mM ZnCl₂, 1 mM MgCl₂, 1 mM β-mercaptoethanol, 20 % glycerol, pH 7.4). Different concentrations of peptides were incubated with the ³²P end-labeled double stranded probe (GRE) in the presence or the absence of varying concentrations of nonlabeled synthetic probe (GRE, TRE or unrelated probe). Final composition of incubation buffer C was 20 mM Tris-HCl, 60 mM NaCl, 1 mM ZnCl₂, 0.2 mM MgCl₂, 0.2 mM β-mercaptoethanol, 4 % glycerol, pH 7.4. Synthetic peptides were incubated at room temperature during 30 minutes. Binding reaction samples (10 μl per aliquot) was then filtered through nitrocellulose membrane equilibrated in buffer D (20 mM Tris-HCl, 60 mM NaCl, pH 7.4). Each membrane was washed with 400 μl of the same buffer and the radioactivity retained on the membrane was counted. All points were measured in duplicate.

Band shift assay.

For the gel retardation assay (Schneider et al., 1986) 1 μ l (2-5 fmol) of the 32 P end-labeled double-stranded GRE oligonucleotide (5,000 - 10,000 cpm), 3 μ l of cytosolic extract and/or of synthetic peptide solution (5-30 mg/ml in buffer A containing or not 2 mM $ZnCl_2$), 1 μ l of (10x) buffer A and 1 μ g poly (dI-dC) were mixed and incubated at 4°C in the presence or the absence of $ZnCl_2$ (final volume 10 μ l). Protein-DNA complexes were resolved for 30 min in a 10 % polyacrylamide gel equilibrated in 25 mM Tris-HCl, 188 mM glycine, pH 8.6, at 15mA in a Biorad minicuve equilibrated in an ice-bath. Gels were then autoradiographed without intensifying screens and quantified by densitometric scanning of the autoradiograms.

Cloning of GRE in pBS vector.

In a way to insert the GRE element in a large DNA fragment, two oligonucleotides GRE-H1, 5'-AGCTTGGTTACAAACTGTTCTA-3' and GRE-H2, 5'-AGCTTAGAACAGTTTGTAACA-3' were prepared on a Applied Biosystem 381A DNA synthesizer and purified according to the manufacturer's instructions. Upon annealing, the element contain a GRE consensus sequence with HindIII cohesive ends.

Plasmid pBS was digested to completion by HindIII restriction enzyme. Phosphorylated synthetic double-stranded oligonucleotide GRE-H1/2 was then added and incubated with T4 DNA ligase. Positive clones were selected using hybridation with 32 P labelled single stranded GRE oligonucleotide. Following plasmid extraction, the PvuII-EcoRI fragment was 3'-end labelled with [α - 32 P] dATP using Kleenow fragment of DNA polymerase under standard conditions. After sequencing the PvuII-EcoRI fragment of several positive clones according to Maxam and Gilbert 1972, one of them was chosen containing a two repeated direct copy of GRE.

Hydroxyl radical footprinting:

Hydroxyl radical cutting reaction was performed essentially as described (21). Samples of the radioactive DNAs GRE-H2 and GRE-H3 (3pmol in base pairs) were incubated with 1ml of GR or peptide solution (0.1-50mM) at 37°C for 30 min. A stock solution of iron(II)-EDTA was prepared immediately before use by mixing 10ml freshly prepared 0.2mM $Fe(NH_4)_2(SO_4)_2$ with 20ml of 0.1mM Na_2 -EDTA (pH7.4) and 20ml freshly prepared 10 mM ascorbic acid. Hydroxyl radicals were generated by allowing that mixture to react with 20ml of 0.3% H_2O_2 . DNA backbone cleavage was initiated by immediately adding 7.5ml of the final mixture to the receptor- or peptide-DNA complex. Aliquots were removed after 1 and 5min digestion and the reaction quenched by adding 5ml of 0.1M thiourea. Samples were then

twice ethanol-precipitated and electrophoresed through a 8% polyacrylamide gel (0.3mm thick) containing 8.34M urea and 1X Tris-Borate-EDTA buffer(pH8.3). Electrophoresis was carried out at 1800V for about 2.5h. Gels were subsequently dried and autoradiographed.

RESULTS.

Finger peptides synthesis and zinc dependent folding.

The N-terminal zinc finger of hGR and hRAR were synthesized (Figure 1). The GR sequence was common to the human (sequence 420-442, Hollenberg et al., 1985) and rat (sequence 439-461, Miesfeld et al., 1986) receptors. This sequence was elongated by one glycine residue at the N-terminal end. In preliminary experiments, a simplified form of this peptide, in which the five C-terminal residues were replaced by glycine was also used. The RAR sequence was common to the human α and β receptors (sequences 86-114 and 79-107 respectively, Giguère et al., 1987 ; De Thé et al., 1987) and also to the corresponding murine receptors (Zelent et al., 1989). Physico-chemical tests were undertaken in order to ascertain that the synthesized peptides effectively complex metals.

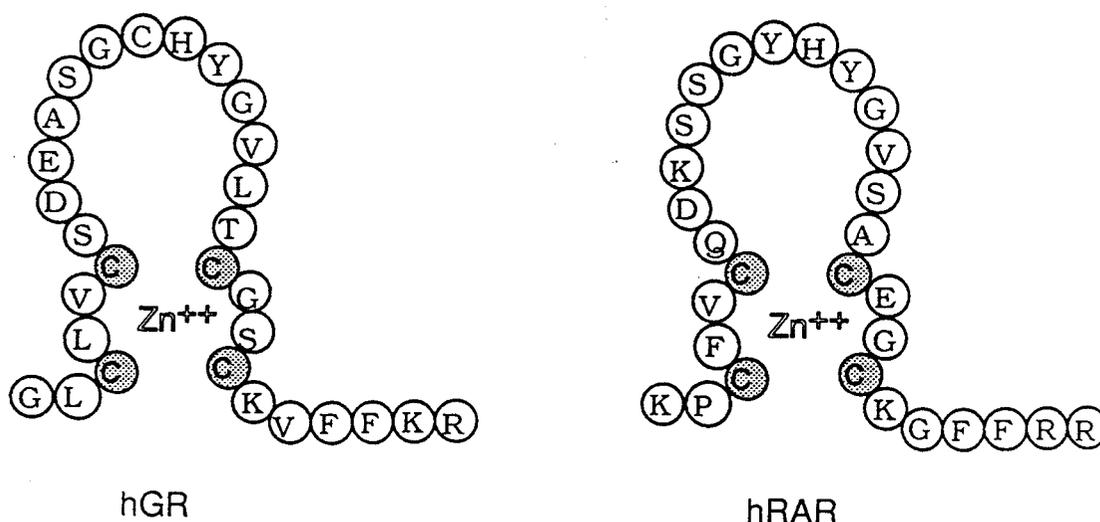


Figure 1. Amino acid sequence of the synthetic peptides.

Both peptides were highly hydrophobic and therefore concentrated solutions in polar solvents were difficult to obtain, specially in the case of the RAR peptide. One dimensional ^1H NMR spectra were recorded for the GR "finger" in the presence and absence of Zn^{2+} . The difference spectrum (Figure 2) exhibits changes around 3.0 ppm, corresponding to Cys β CH_2 resonances (Wuthrich, 1976), this being in favour of Zn^{2+} complexation by the cysteine thiol groups.

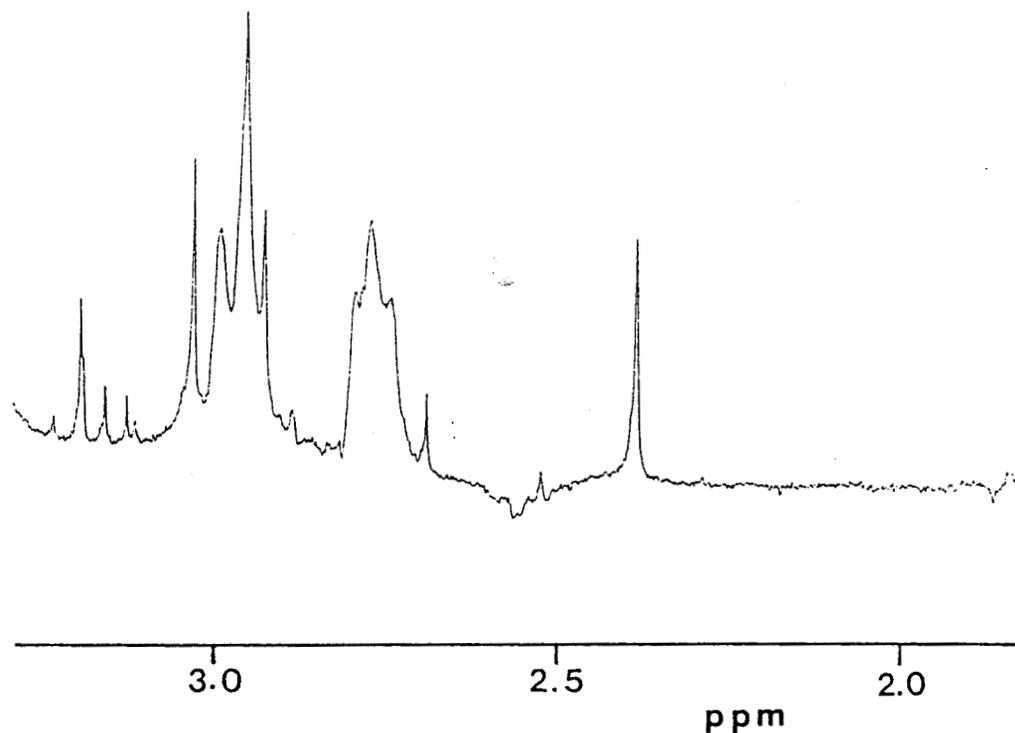


Figure 2. 400-MHz ^1H NMR difference spectrum of human glucocorticoid receptor (sequence 420-447) in the presence and in the absence of zinc.

A preliminary experiment using absorption spectroscopy showed that the RAR and GR "finger" are able to coordinate tetrahedrally Co^{2+} in the same way. On the other hand, circular dichroism spectroscopy was used to monitor secondary structural changes in the presumed "fingers". The CD spectrum of the reconstituted GR peptide in the presence of EDTA differs from that of the same peptide refolded in ZnCl_2 solutions (Figure 3), indicating an onset of secondary structure which could not be accurately precised.

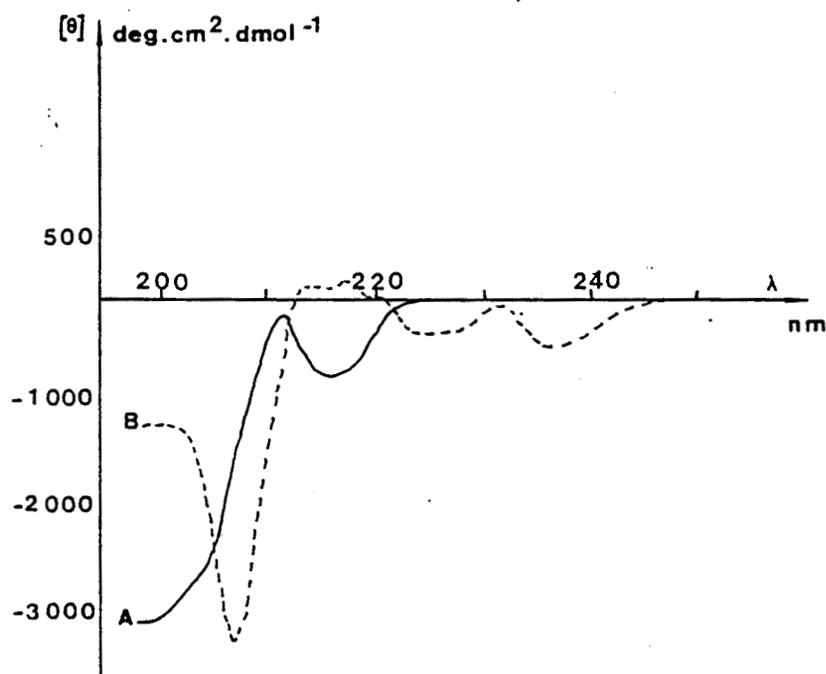


Figure 3. Circular dichroism spectra of human glucocorticoid receptor (sequence 420-447) in the presence (B) and absence (A) of zinc.

In the case of RAR no definite change could be seen. These results are in accordance with conformational energy minimization calculations (J.Poupaert, personal communication) showing that GR 439-447 residues are able to initiate an α helix whereas such a structure could not be predicted for the RAR sequence.

Specific GRE binding by the synthetic GR zinc finger can be demonstrated using nitrocellulose filter assay.

When increasing amounts of GR zinc finger peptide were incubated with a constant amount of ^{32}P -labeled oligonucleotide containing a typical GRE consensus sequence (Beato, 1989), a dose-dependent binding curve was obtained (Figure 4).

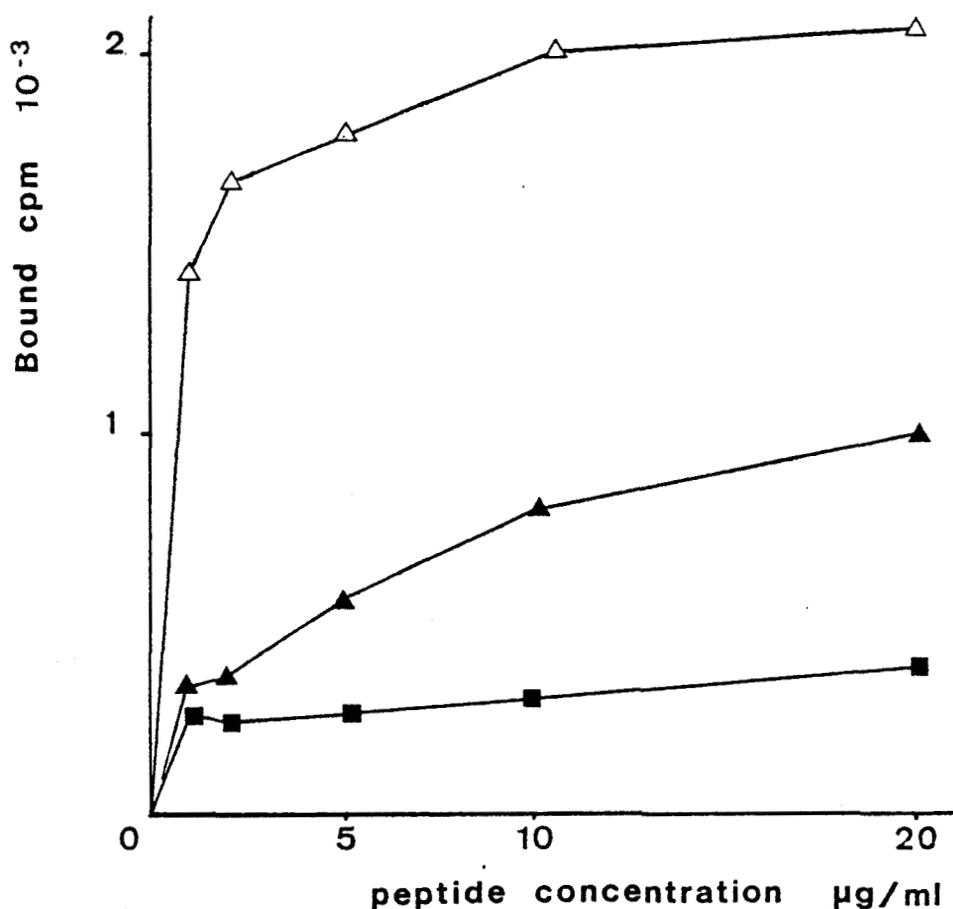


Figure 4. DNA-binding specificity of synthetic peptides to GRE. A ^{32}P end-labeled GRE oligonucleotide (10^{-3} M) was incubated with various concentrations of GR finger peptide (Δ), RAR finger peptide (\blacktriangle) and unrelated peptide (\blacksquare) in a final buffer containing 20 mM Tris-HCl, 60 mM NaCl, 1 mM ZnCl₂, 0.2 mM MgCl₂, 0.2 mM β -mercaptoethanol, 4% glycerol, pH 7.4. After 30 min. at room temperature, binding was determined by nitrocellulose filter assay. Each point was evaluated in duplicate.

The binding appeared rather specific since, when tested at the same concentration, an unrelated peptide yielded very low binding, whereas the related RAR zinc finger peptide yielded partial binding when compared to the GR peptide. The binding of the unrelated peptide could be considered as typical nonspecific binding since the curve was a straight line (the fact that this line did not include the origin indicated some background minimization problem in the nitrocellulose filter assay, which was empirically standardized and needs probably further optimization). However despite this limitation, 0.4×10^{-6} M and 2×10^{-6} M K_d values could be estimated from the data for GR and RAR zinc finger peptide-GRE interaction respectively. Moreover the specificity of the binding of the GRE sequence by the GR peptide was demonstrated by competition experiments (Figure 5).

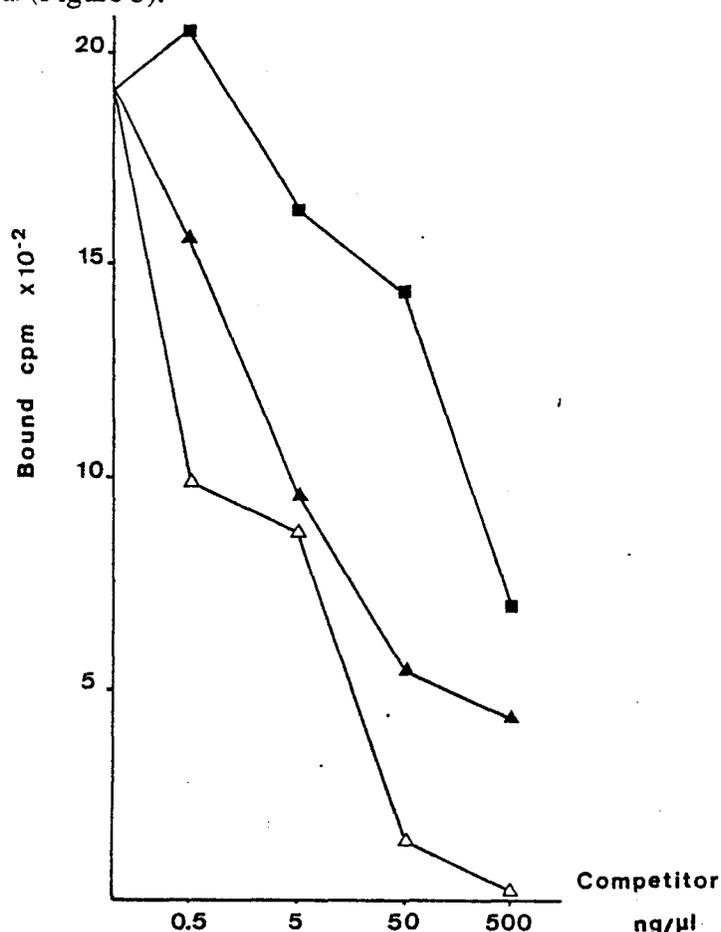


Figure 5. Specificity of GRE oligonucleotide for the GR zinc-finger peptide. GR or RAR zinc-finger peptide (20 mg/ml) was incubated with 32 P-end labeled GRE for 30 min at room temperature in the presence of various concentrations of unlabeled GRE (Δ), TRE (\blacktriangle) or of a 20 bp unrelated oligonucleotide (\blacksquare). Binding assay was as in figure 4.

The results showed that unlabeled GRE was a 10-fold more potent competitor than a palindromic thyroid responsive element (TRE), another hormone responsive element displaying some sequence homology with GRE (Umesono and Evans, 1989), and 1000-fold more potent than an unrelated oligonucleotide. Finally zinc dependence of finger peptide-GRE interaction was also demonstrated: very low binding was obtained in the absence of zinc (data not shown) and the binding observed in the presence of 1 mM Zn was completely abolished when the chelating agent EDTA was added in excess (Figure 6). This effect was dose-dependent and observed with both finger peptides.

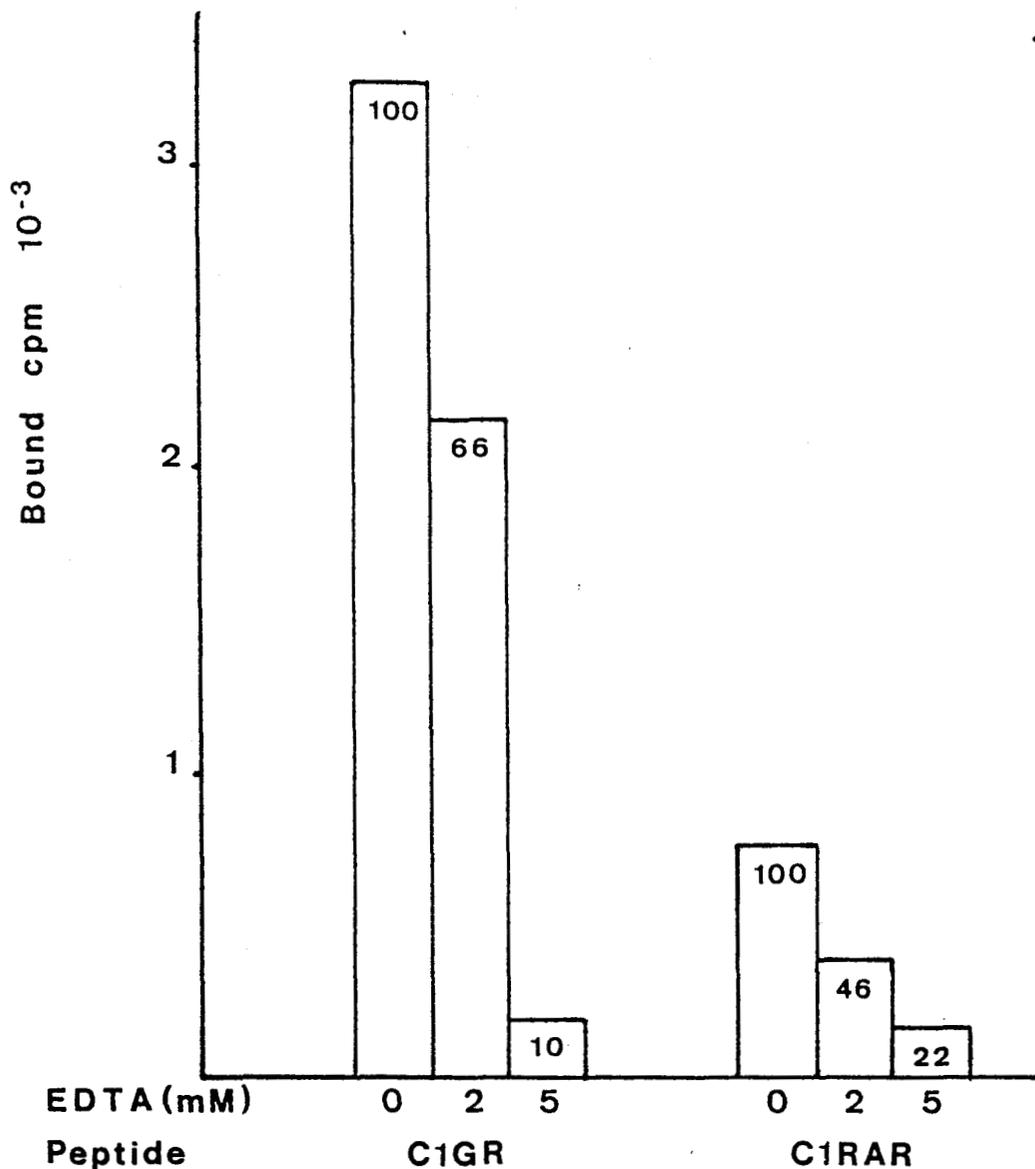


Figure 6. Zn²⁺ dependence of finger-GRE complex formation. GR or RAR zinc finger peptides were incubated (10 μ g/ml) with ³²P-end labeled GRE in buffer C in the presence of 0.2 or 5 mM EDTA. The nitrocellulose binding assay was then performed. After washing with buffer D, filters were counted for radioactivity. Each point was assayed in duplicate.

Competition with native glucocorticoid receptor for gel-retardation of labeled GRE can be used to demonstrate specific GRE binding by synthetic zinc fingers.

When the interaction of the zinc finger peptide with a ³²P-labeled GRE was studied by gel retardation assay, no band-shift could be observed, even in the presence of a high concentration of peptide (11 mg/ml, i.e. 300 nM or 50-100 fold excess versus GRE concentration) refolded in the presence of Zn²⁺. We checked that there was no oligonucleotide degradation by the peptide solution (data not shown). We concluded that the affinity for the GRE of a single finger was probably not strong enough to avoid complete dissociation of the peptide-GRE complex during the non-equilibrium conditions of the electrophoretic run. However, we postulated that if the zinc finger peptide has some affinity for the GRE, it must

be able to compete with the native glucocorticoid receptor for specific DNA-binding. Such a competition could be performed in equilibrium conditions and the effect of the peptide would be indirectly detected, by the appearance of a decrease in the band-shift obtained with a standard amount of native receptor alone.

The first step in the design of such a competition assay was the standardization of the receptor preparation and the band-shift assay conditions, in order to obtain reference retarded band, specifically due to the presence of the activated receptor. Best results (Figure 7) were obtained with receptor extracts partially purified and prepared in an homogeneously transformed form by a succession of DNA-cellulose chromatography (to remove interfering DNA binding proteins), incubation in the presence of triamcinolone acetonide (TA), heat activation and final purification on a DEAE trisacryl column (Danzé et al., 1990).

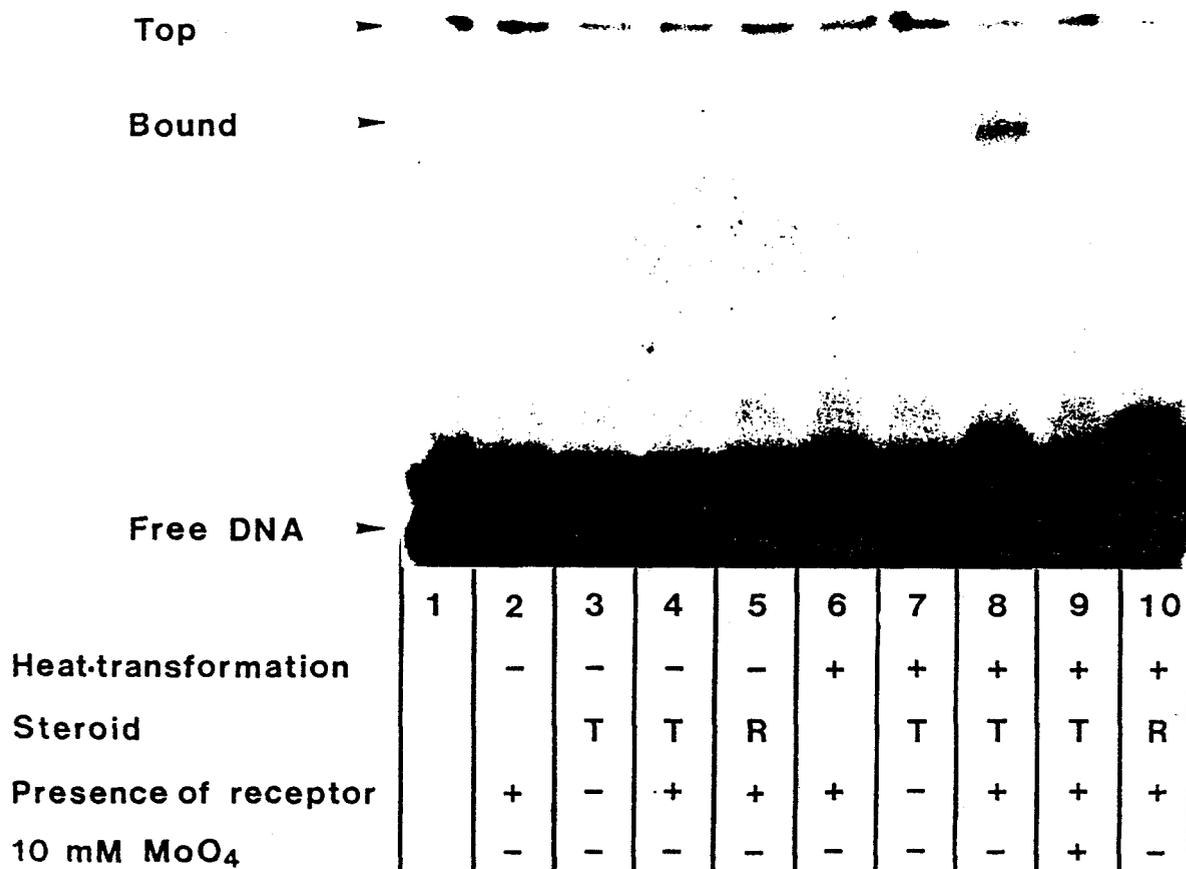


Figure 7. Gel retardation assay of homogeneously transformed receptor complexes.

Rat liver GR complexes were purified as described in material and methods. Aliquots were incubated with ³²P labeled GRE in the presence of poly (dIIdC). Before GR purification cytosolic extracts were incubated with either triamcinolone acetonide (T, lanes 3,4,7-9) or RU 486 (R, lanes 2 and 6) or were depleted in GR by affinity chromatography (lanes 3 and 7). Samples were then heated 30 min at 25 C (lanes 6-10) or maintained at 0°C for the same time (lanes 2-5). After receptor purification the protein concentration was assayed in all samples and adjusted to the same final value for the ³²P GRE binding step. A control sample containing only the ³²P GRE in the reaction buffer, and no GR was also run (lane 1).

A retarded band was only observed for a sample incubated with a glucocorticoid agonist and heat-activated (lane 8). No band-shift was observed in the absence of steroid (lanes 2 and 6) or of heating (lanes 2-5). Similar results were observed in the presence of sodium molybdate (lane 9), an inhibitor of steroid receptors transformation (Dahmer et al., 1984), or in the presence of RU 486 (lanes 5 and 10), an antiglucocorticoid which inhibits also GR transformation (Sablonnière et al., 1986). Moreover, no band-shift was observed with cytosolic samples specifically depleted in glucocorticoid receptor (lanes 3 and 7) by affinity chromatography on a dexamethasone bearing gel (Idziorek et al., 1985). Thus, the observed band-shift was clearly glucocorticoid receptor dependent. It was also specific for the GRE sequence as shown by competition experiments (Figure 8).



Figure 8 : DNA binding specificity of the GR band-shift. Gel retardation analysis of the binding of [32 P]GRE to human lymphocyte cytosolic extracts preincubated in the presence of dexamethasone and in the absence (lanes a,e,i) or in the presence of a various molar excess of unlabeled GRE (lanes b-d), TRE (lanes f-i) and unrelated oligonucleotide (lanes k-n). The sequences of the latter nucleotides were 5'-GATCTTCAGGTCATGACCTGAA-3'(TRE) and 5'-GTAGGCCACGTGACCGGG-3' (unrelated sequence).

Unlabeled GRE was a good competitor, whereas an unrelated oligonucleotide competed very weakly and TRE had an intermediate affinity. In fig. 8, the specific band-shift of [^{32}P]GRE appears as a double band. In this particular experiment crude human leukocyte cytosol was used and the double band can result either from receptor dimerization (Green et al., 1988) or from partial enzymatic degradation of the receptor, an event frequently observed in human leukocyte cytosol (Diestelhorst et al., 1987). The band-shift observed was dependent on Zn^{2+} concentration, with a maximum at 1.0 - 2.5 mM and a decrease at higher concentration (Figure 9).

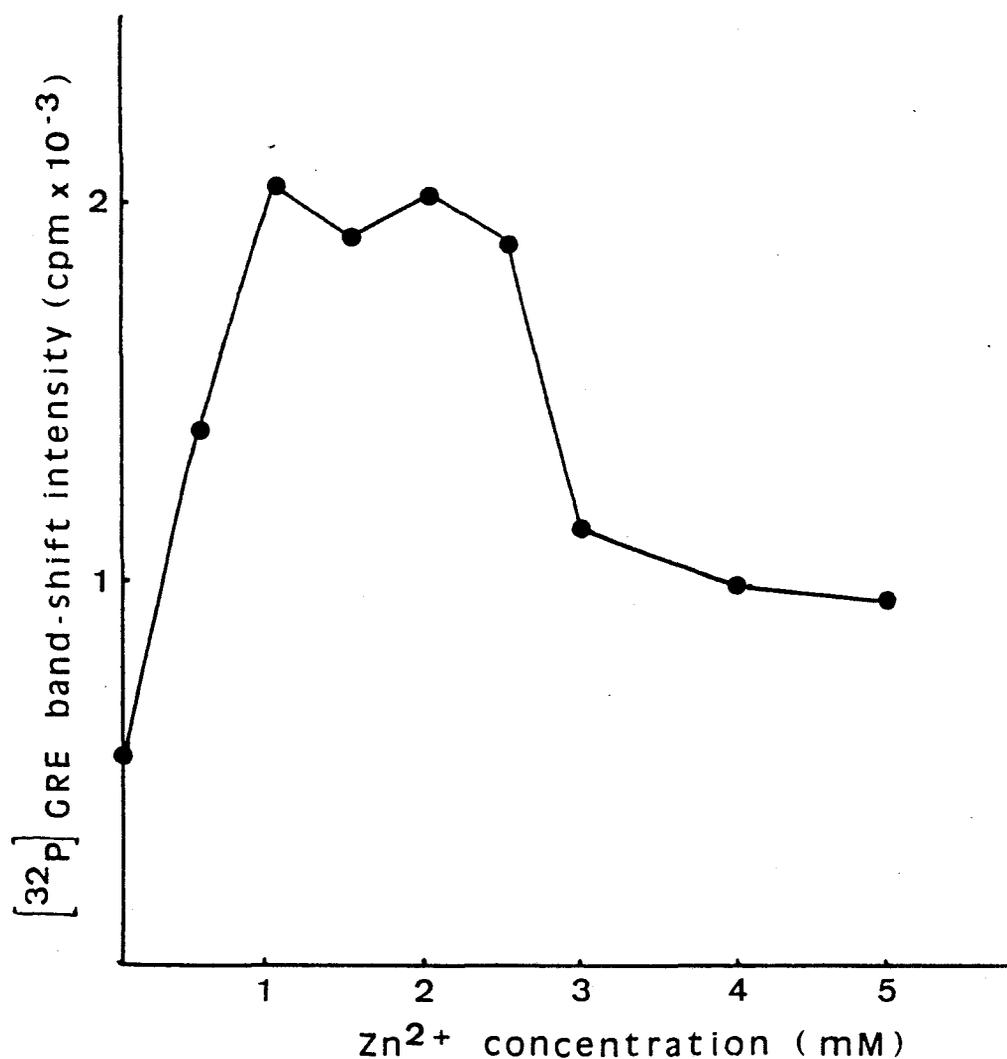


Figure 9 : Zn^{2+} concentration dependence of dexamethasone-induced band-shift.

Human leukocyte cytosolic extracts were incubated with 1 μM dexamethasone and tested for [^{32}P]GRE binding in a buffer containing various Zn^{2+} concentrations. Retarded bands were cut and counted for radioactivity after solubilization with H_2O_2

The second step was the realization of the competition assay with various peptides. In the incubation medium the GRE and receptor concentrations were roughly equivalent, in the nM range. The intensity of the shifted band was significantly decreased in the presence of increasing concentrations of zinc-finger peptides (Figure 10).

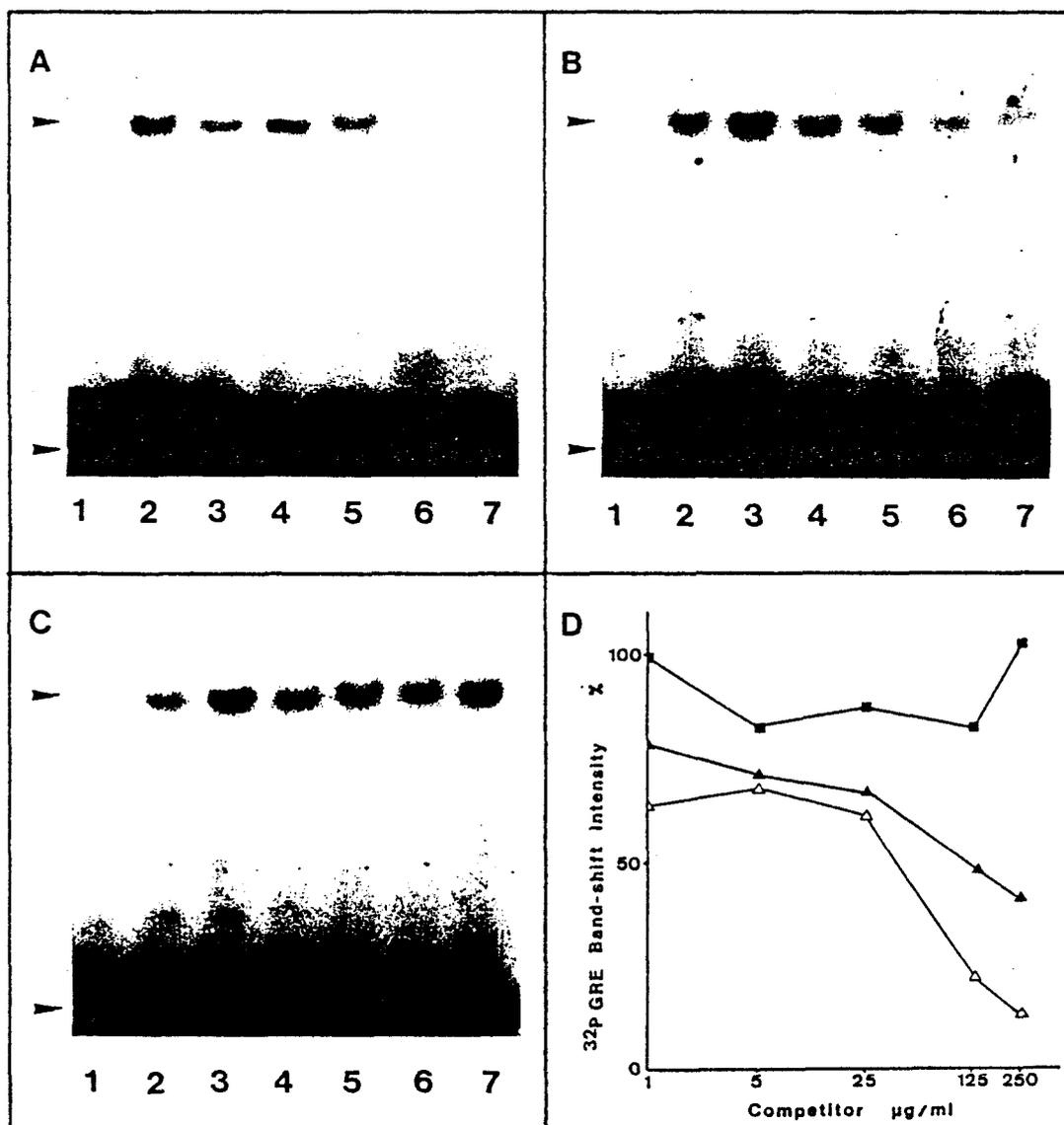


Figure 10 : Band-shift competition assay of the finger peptides binding to ^{32}P GRE

Homogeneously heat transformed and partially purified rat liver TA-GR complexes were incubated in the presence of a constant amount of ^{32}P GRE and increasing amounts of various competitor peptides : GR zinc finger (panel A), RAR zinc finger (panel B) or unrelated peptide (panel C). A band-shift assay was then performed and the gels were autoradiographed. Competitor peptide concentrations were 0, 1, 5, 25, 125 and 250 $\mu\text{g/ml}$ respectively in lanes 2-7. A control sample corresponding to non transformed GR complexes was run in parallel (lane 1). The retarded bands obtained in panels A, B and C were cut, solubilized with H_2O_2 and counted for radioactivity. Panel D depicted the competition curves obtained with the GR (Δ), RAR (\blacktriangle) and unrelated (\blacksquare) peptides. Results were expressed in per cent of the band-shift intensity obtained in the absence of competitor in each panel (lane 2).

As expected, the more potent competitor was the GR finger peptide (panels A and D). No competition was observed with an unrelated synthetic peptide of similar size (panels C and D), whereas a partial competition was observed with the RAR finger peptide (panels B and D). From the competition curves expressing the data in quantitative terms (panel D), peptide concentrations yielding 50% band-shift inhibition, i.e. 50% inhibition of GR-GRE binding, could be estimated and K_i values deduced. Calculated K_i were 1.1 and 3.3×10^{-5} M for GR and RAR finger peptides binding of GRE respectively. The equilibrium dissociation constant K_d of the native GR-GRE complex was determined in separate experiments (data not shown) using the procedure proposed by Baker et al. 1986. Calculated K_d , extrapolated at zero poly(dI-dC)-concentration, was 1.2×10^{-10} M. The K_i values reported here for GR and RAR peptides were obtained in the presence of a large excess of poly(dI-dC) and not corrected.

Footprinting experiments.

Hydroxyl radicals footprinting experiments have been simultaneously carried out with the purified GR and the GR finger peptide on the double GRE-inserted PvuII-EcoRI fragment from pBS (Assays with RAR finger peptide required too high peptide concentration for significant results). The effects of the whole GR and the finger peptide are reported on figure 11. The most obvious change in the presence of either GR (lane X-Y) or peptide (lanes Y-Z) is a decrease in the intensity of several bands with a two repeated shape. It is, clear that, for both receptor and GR finger peptide, GRE sequences are the major sites of protection. with an enhanced cleavage in the middle of the palindrome suggesting the expected interaction of one molecule of peptide per half-site of GRE.

Interestingly, one can observe that the pattern of protection in the presence of the peptide rigorously follows those in the presence of the whole GR indicating the same selectivity for DNA binding. However, densitometric analysis of the autoradiogram (not shown) show a lower intensity of protection with the peptide. This confirms the low-affinity of the peptide for DNA-binding already observed through nitrocellulose and band-shift experiments. Nuclease footprinting experiments have also been carried out (not shown); the results are similar to those of protection against hydroxyl radicals, although they are less distinct with the peptide, probably because nucleases are DNA-binding proteins which displaced the peptide during digestion.

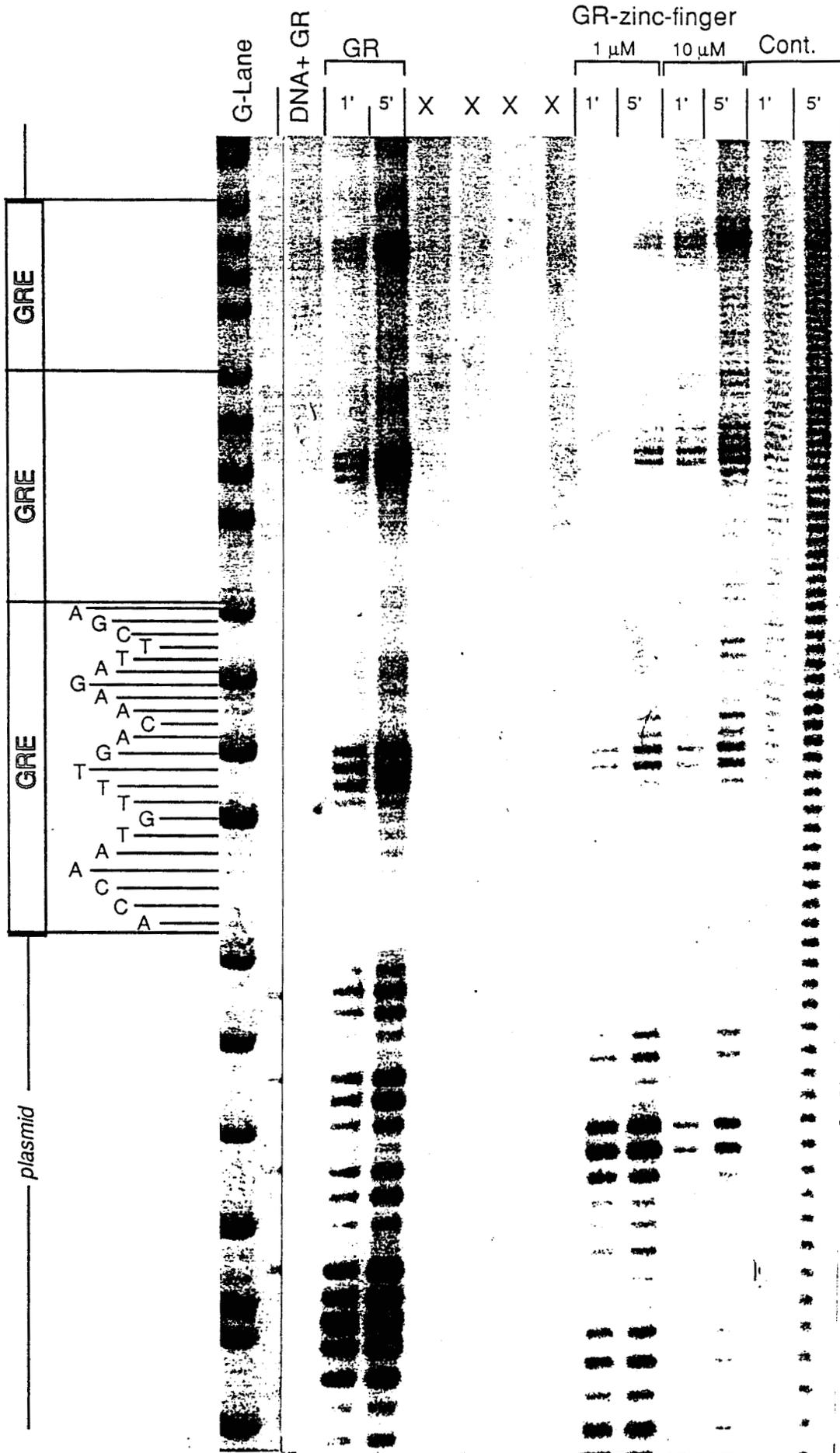


Figure 11. Hydroxyl radicals footprinting of glucocorticoid receptor (GR) and glucocorticoid receptor zinc-finger peptide. Lane (ADN + GR) correspond to DNA fragment GRE-3 + receptor without cleaving agent. The reader may disregard lanes noted (X) which correspond to the effects of RAR peptide which precipitated in the loading wells.

DISCUSSION.

Our synthetic GR N-terminal zinc finger appeared able to effectively complex metals and to fold in the presence of zinc. However, no well-defined secondary structure could be deduced from the CD spectrum. This feature is in agreement with the model very recently proposed for the DNA-binding domain of the GR (Hård et al., 1990). In this model, established by two-dimensional nuclear magnetic resonance study of a 71 residue protein fragment containing the two zinc finger subdomains, no α helical regions are present in the fingers themselves, which contain only very short stretches of antiparallel β sheet. The structure of the C₄ zinc fingers characteristic of the steroid receptor superfamily (Evans and Hollenberg, 1988 ; Berg, 1990) appears therefore quite different from the structure of "C₂H₂" zinc fingers found in other DNA binding proteins and characterized by an antiparallel β -sheet followed by an α -helix as proposed for the transcription factor IIIA (Berg, 1988) and demonstrated by 2D NMR for related peptides from other transcription factors (Lee et al., 1989 ; Parraga et al., 1988). The model was already unsuccessfully tested in the case of the GR (Chalepakis et al., 1988). In the model reported by Hård et al., 1990, two α -helical regions are found on the C-terminal side of both zinc fingers. The first α -helix was proposed as recognition helix involved in direct interaction with the GRE in the DNA major groove. The C-terminal half of this α -helical subdomain was absent in our synthetic zinc fingers. However our peptides bore the amino acids described in the finger knuckle and its C-terminal tail (Gly439, Ser440 and Val443 in the hGR sequence) as determinants of target gene specificity (Danielsen et al., 1989 ; Mader et al., 1989; Umesono and Evans, 1989). Even in our most simplified peptide, two of these three residues were still present. Our DNA binding experiments demonstrated that these simplified structures were sufficient for a rather specific recognition of a GRE sequence. These results contrast with the lack of detectable DNA binding of a single zinc finger previously reported for the glucocorticoid receptor (Freedman et al., 1988) and with the nonspecific binding observed for transcription factors IIIA (Frankel et al., 1987) and ADR1 (Parraga et al., 1988). This apparent discrepancy can be explained by technical considerations. Freedman et al., 1988, used a gel retardation assay for the direct detection of DNA binding and observed a specific band-shift only with a 150 aminoacid GR fragment encompassing both zinc fingers and the basic residues downstream of the fingers. A fragment obtained by trypsin digestion and comprising the intact N-terminal finger, the linker region and the N-terminal half of the second finger, did not bind DNA. Similarly, no band-shift could be observed with our synthetic zinc finger. However, we were able to demonstrate its DNA binding activity by an original band-shift competition assay. In this assay the synthetic zinc finger displayed a clear specificity for binding to GRE. This binding could not be explained by non-specific interactions involving the basic residues of the peptides, since the

GR peptide, despite a lower content in basic amino-acids, displayed a higher affinity for the GRE than the RAR peptide. This specificity was unequivocally supported by footprinting results. The peptide effectively binds DNA with the same site selectivity as the whole activated receptor

The lack of directly detectable band-shift when using a single zinc finger probably results from a very fast dissociation of the finger-GRE complexes. A half-life of 12-15 min has been reported for the GRE-native GR complexes (Schmid et al., 1989). In the case of the progesterone receptor a bacterially expressed DNA-binding domain yielded HRE-protein complexes dissociating far more rapidly (in less than 1 min) than the native receptor (Eul et al., 1989). The single finger-GRE complexes studied here were therefore logically expected to dissociate more rapidly than the intact GR or its entire DNA-binding domain. A too rapid dissociation precludes any complex detection by the nonequilibrium and time-lasting gel retardation assay. However finger-GRE complexes could be detected by the very rapid nitrocellulose filter assay and the "receptor-like" GRE specific binding by the equilibrium and rapid method of protection against hydroxyl radicals. The binding displayed a clear zinc dependence, also observed for the native receptor in the gel retardation assay. Thus zinc-dependent folding appears as a prerequisite for GRE binding of single finger peptides, result already reported for the entire GR DNA-binding domain (Freedman et al., 1988). The K_d found for the wild GR-GRE complexes was in the 10^{-9} - 10^{-10} M range already reported by others (Tully et al., 1990 ; Wrangé et al., 1989 ; Yang-Yen et al., 1990). The affinity observed here for the single GR N-terminal zinc finger was far lower, with a K_d in the 10^{-6} M range from the nitrocellulose filter assay. However, this four orders of magnitude lower affinity was not unexpected, since a two orders of magnitude fall in DNA-binding affinity (K_d in the 10^{-7} M range) was already reported for GR fragments encompassing the entire DNA-binding domain (Dahlman et al., 1989 ; Härd et al., 1990a). This decrease could be explained by the inability of our peptides to form the globular structure stabilized by the two Cys-Zn clusters of the entire DNA binding domain (Härd et al., 1990b) and by the loss of protein-protein contact involving the N-terminal domain of the native GR and affecting the DNA-binding properties of the receptor homodimer (Eriksson and Wrangé, 1990). In the case of our single finger peptide, the lack of the C-terminal part of the finger tail containing amino acids involved in specific GRE recognition, and the highly probable inability to dimerize, contrary to larger DNA binding domain fragments, explain the further loss in affinity observed. However, we wanted to stress in this report upon the fact that despite this rather low affinity, a single GR zinc finger displays specificity for binding to a GRE sequence or, at least, the same selective binding as the whole activated receptor.

11. Dahmer, M.K., Housley, P.R. and Pratt, W.B. (1984) Effects of molybdate and endogenous inhibitors on steroid-receptor inactivation, transformation and translocation. *Ann. Rev. Physiol.* **46**, 67-81.
12. Danielsen, M., Minck, L. and Ringold, G.M. (1989) Two aminoacids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. *Cell* **57**, 1131-1138.
13. Danze, P.M., Richard, C., Formstecher, P. and Dautrevaux, M. (1990) Steroid receptor exchange assay in the presence of aceto-nitrile. Applications to the study of glucocorticoid and antiglucocorticoid receptor complexes. *Steroids* **55**, 10-16.
14. De The, H., Marchio, A., Tiollais, P. and Dejean, A. (1987) A novel steroid thyroid hormone receptor related gene inappropriately expressed in human hepatocellular carcinoma. *Nature* **330**, 667-670.
15. Diestelhorst, C.W., Janiga, K.E., Howard, K.J., Stranjord, S.E. and Campbell, E.J. (1987) Neutrophil elastase produces 52 kD and 30 kD glucocorticoid receptor fragments in the cytosol of human leukaemia cells. *Blood* **70**, 860-868.
16. Eriksson, P. and Wrangé, O. (1990) Protein-protein contacts in the glucocorticoid receptor homodimer influence its DNA binding properties. *J. Biol. Chem.* **265**, 3535-3542.
17. Eul, J., Meyer, M.E., Tora, L., Bocquel, M.T., Quirin-Stricker, C., Chambon, P. and aronemeyer, H. (1989) Expression of active hormone and DNA-binding domains of the chicken progesterone receptor in *E. coli*. *EMBO J.* **8**, 83-90.
18. Evans, R.M. (1988) The steroid and thyroid hormone receptor superfamily. *Science* **240**, 889-895.
19. Evans, R.M. and Hollenberg, S.M. (1988) Zinc fingers: guilt by association. *Cell* **52**, 1-3.
20. Frankel, A.D., Berg, J.M. and Pabo, C.O. (1987) Metal-dependent folding of a single zinc finger from transcription factor III A. *Proc. Natl. Acad. Sci. USA* **84**, 4841-4845.
21. Freedman, L.P., Luisi, B., Korszun, Z.R., Basavappa, R., Sigler, P.B. and Yamamoto, K.R. (1988) The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature* **334**, 543-545.
22. Giguere, V., Ong, E.S., Segui, P. and Evans, R.M. (1987) Identification of a receptor for the morphogen retinoic acid. *Nature* **330**, 624-629.
23. Green, S. and Chambon, P. (1988) Nuclear receptors enhance our understanding of transcription regulation. *Trends Gen.* **4**, 309-314.
24. Green, S., Kumar, V., Theulaz, I., Wahli, W. and Chambon, P. (1988) The N-terminal DNA-binding "zinc finger" of the oestrogen and glucocorticoid receptors determines target gene specificity. *EMBO J.* **7**, 3037-3044
25. Hard, T., Dahlman, K., Carlstedt-Duke, J., Gustafsson, J.A. and Rigler, R. (1990) Cooperativity and specificity in the interactions between DNA and the glucocorticoid receptor DNA-binding domain. *Biochemistry* **29**, 5358-5364.

26. Hard, T., Kellenbach, E., Boelens, R., Maler, B.A., Dahlman, K., Freedman, L.P., Carlstedt-Duke, J., Yamamoto, K.R., Gustafsson, J.A. and Kaptein, R. (1990b) Solution structure of the glucocorticoid receptor DNA-binding domain. *Science* **249**, 157-160.
27. Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M. and Evans, R.M. (1985) Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* **318**, 635-641.
28. Hollenberg, S.M., Giguere, V., Segui, P. and Evans, R.M. (1987) Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor. *Cell* **49**, 39-46.
29. Idziorek, T., Formstecher, P., Danze, P.M., Sablonniere, B., Lustenberger, P., Richard, C., Dumur, V. and Dautrevaux, M. (1985) Characterization of the purified molybdate-stabilized glucocorticoid receptor from rat liver. *Eur. J. Biochem.* **153**, 65-74.
30. Klug, A. and Rhodes, D. (1987) Zinc fingers: a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12**, 464-469.
31. Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A. and Wright, P.E. (1989) Three-dimensional solution structure of a single finger DNA-binding domain. *Science* **245**, 635-637.
32. Mader, S., Kumar, V., De Verneuil, H. and Chambon, P. (1989) Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid responsive element. *Nature* **338**, 271-274.
33. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning*. Cold Spring Harbor, New-York pp. 125.
34. Miesfeld, R., Rusconi, S., Godowski, P.J., Maler, B.A., Okret, S., Wilkstrom, A.C., Gustafsson, J.A. and Yamamoto, R.R. (1986) Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA. *Cell* **46**, 389-399.
35. Nagai, K., Nakaseko, Y., Nasmyth, K. and Rhodes, D. (1988) Zinc-finger motifs expressed in E.coli and folded in vitro direct specific binding to DNA. *Nature* **332**, 284-286.
36. Parraga, G., Horvath, S.J., Eisen, A., Taylor, W.E., Hood, L., Young, E.T. and Klevit, E. (1988) Zinc-dependent structure of a single-finger domain of yeast ADRI. *Science* **241**, 1489-1492.
37. Sabatier, J.M., Tessier-Rochat, M., Granier, C., Van Rietschoten, J., Pedroso, E., Grandas, A., Albericio F. and Giralt, E. (1987) Convergent solid phase synthesis of a 19 VI Synthesis by the Fmoc procedure with a modified protocol of two protected segments. sequence 15-17 and 18-31 of the neurotoxin II of the scorpion *Androctonus australis Hector*. *Tetrahedron* **43**, 5973-5980.

38. Sablonniere, B., Danze, P.M., Formstecher, P., Lefebvre, P. and Dautrevaux, M. (1986) Physical characterization of the activated and non-activated forms of the glucocorticoid-receptor complex bound to the steroid antagonist [3H] RU 486. *J. Steroid Biochem.* **25**, 605-614.
39. Schmid, W., Strahle, V., Schutz, G., Schmitt, J. and Stunnenberg, H. (1989) Glucocorticoid receptor binds cooperatively to adjacent recognition sites. *EMBO J.* **8**, 2257-2263.
40. Schneider, R., Gander, I., Muller, U., Mertz, R. and Winnacker, E.L. (1986) A sensitive and rapid gel retention assay for nuclear factor I and other DNA-binding proteins in crude nuclear extracts. *Nucl. Acids Res.* **14**, 1303-1317.
41. Severne, Y., Wieland, S., Schaffner, W. and Rusconi, S. (1988) Metal binding finger structures in the glucocorticoid receptor defined by site-directed mutagenesis. *EMBO J.* **7**, 2503-2508.
42. Tully, D.B. and Cidlowski, J.A. (1990) Affinity of interactions between human glucocorticoid receptors and DNA: at physiologic ionic strength, stable binding occurs only with DNA, containing partially symmetric glucocorticoid response elements. *Biochemistry* **29**, 6662-6670.
43. Umesono, K. and Evans, R.M. (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57**, 1139-1146.
44. Wrangé, O., Eriksson, P. and Perlman, T. (1989) The purified activated glucocorticoid receptor is a homodimer. *J. Biol. Chem.* **264**, 5253-5259.
45. Wuthrich, K. (1976) in *"NMR in biological research: peptides and proteins"*. North Holland/American Elsevier, New-York pp.15-56.
46. Yang-Yeng, H.F., Chambard, J.C., Sun, Y.L., Smeal, T., Schmidt, T.J., Drouin, J. and Karin, M. (1990) Transcriptional interference between c-Jun and the glucocorticoid receptor: mutual inhibition of DNA binding due to direct protein-protein interaction. *Cell* **62**, 1205-1215.
47. Zelent, A., Krust, A., Petkovitch, M., Kastner, P. and Chambon, P. (1989) Cloning of murine α and β retinoic receptors and a novel receptor gamma predominantly expressed in skin. *Nature* **339**, 714-717.

- Par ce qui précède nous avons donc démontré qu'un peptide synthétique correspondant au doigt de zinc CI du récepteur des glucocorticoïdes est capable d'interagir spécifiquement avec la séquence reconnue par le récepteur entier mais avec une affinité moindre.

Dans la chronologie de ces études, la littérature a alors fait état de résultats intéressants obtenus par des techniques de biologie moléculaire (Mader *et al.*, 1989). Par mutation dirigée, il a pu être montré que seulement trois acides aminés situés à proximité du complexe zincique sont essentiels pour une discrimination de reconnaissance entre deux éléments de réponse hormonale (figure 13). Les acides aminés (Glu, Gly et Ala pour le récepteur des œstrogènes; Gly, Ser et Val pour le récepteur des glucocorticoïdes) occupant ces positions sont donc clairement impliqués dans la reconnaissance de l'ADN.

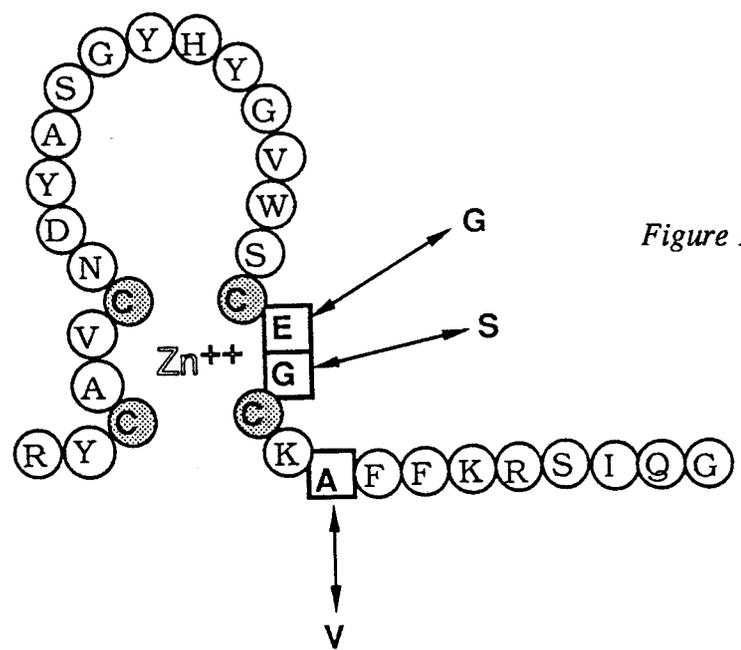


Figure 13. Résultats des études de mutagenèse sur le doigt CI du récepteur des œstrogènes.

Les acides aminés discriminants sont représentés dans des carrés. Les substitutions avec les acides aminés correspondants du récepteur des glucocorticoïdes sont indiquées par les flèches bicéphales.

Tenant compte de ces résultats, nous avons fait le pari que les acides aminés contenus dans la boucle entre la troisième et la quatrième cystéine étaient peu ou pas impliqués dans l'interaction spécifique avec l'ADN et nous avons, pour vérifier cette hypothèse, synthétisé deux versions simplifiées du doigt de zinc CI (figure 14).

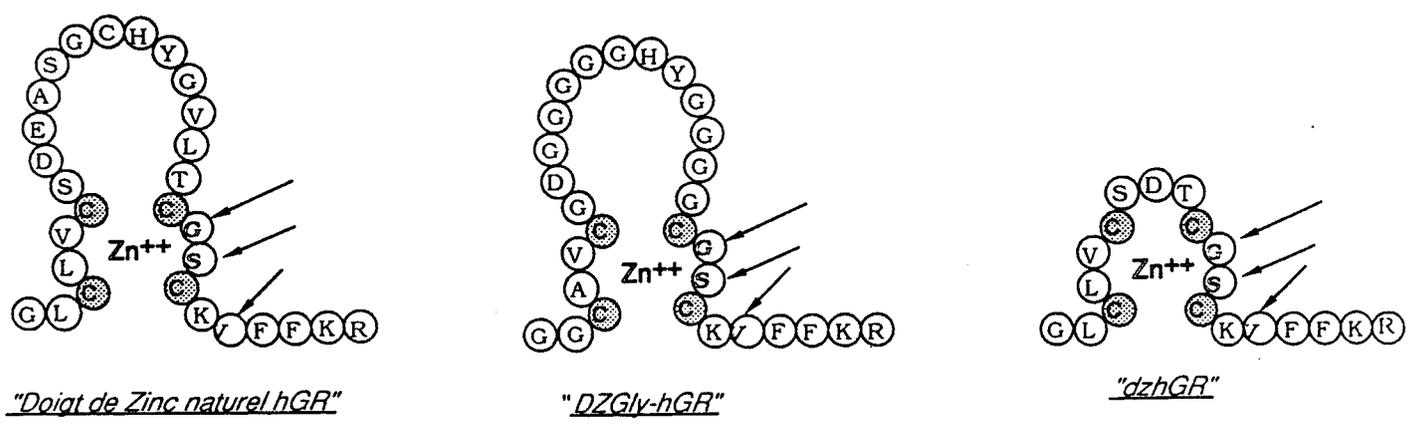


Figure 14. Versions simplifiées du peptide doigt de zinc CI de l'hGR.

Nous avons élaboré un premier peptide où les acides aminés de la boucle ont été remplacés par des résidus glycine (DZGly, figure 14), supprimant ainsi les possibilités d'interaction des chaînes latérales. Cependant, l'hydrophobie observée (et attendue) de ce peptide n'a pas permis de montrer de manière significative une interaction spécifique avec l'ADN. Dans une deuxième simplification du modèle doigt de zinc (dz, figure 14), nous avons opéré une délétion de la boucle conservant toutefois trois acides aminés afin de permettre au complexe de zinc de se former (vérification établie au préalable sur modèle moléculaire). Une fois encore, la liaison spécifique à l'ADN n'a pas pu être mise en évidence par les techniques classiques d'interaction.

Néanmoins, ce peptide simplifié a quand même pu être utilisé avec succès en tant qu' "outil de reconnaissance" pour diriger une coupure sur un site précis de l'ADN. Cet aspect sera discuté au cours du dernier chapitre.

La suite de l'étude fait appel à des résultats de la littérature et à notre propre réflexion concernant le niveau structural de ce motif qui vont maintenant être exposés.

Structure tridimensionnelle du domaine de liaison à l'ADN des⁶² récepteurs nucléaires.

Des études récentes de RMN-2D menées sur le récepteur des œstrogènes (Schwabe et al., 1990) et sur celui des glucocorticoïdes (Härd et al., 1990) ont permis une approche de la structure tridimensionnelle du domaine de liaison à l'ADN des récepteurs hormonaux. Pour ces études, des quantités importantes de peptide correspondant au domaine de liaison à l'ADN (env 70 aa, 2 doigts de zinc) de chaque récepteur ont été obtenues après clonage, expression chez *E. coli* et purification. La liaison spécifique de ces peptides biosynthétiques à leurs HRE correspondants est montrée par des techniques de retard sur gel et d'empreintes à la nucléase (Schwabe et al., 1990, Freedman et al., 1988).

Grâce à de puissants algorithmes, la RMN-2D permet d'aborder assez correctement une estimation de la structure tridimensionnelle du polypeptide en solution (pour revue, voir Wüthrich et al., 1990). La figure 15 montre une représentation schématique de cette structure du domaine de liaison à l'ADN du récepteur des glucocorticoïdes. La structure est constituée de deux motifs similaires (correspondant grossièrement aux deux doigts de zinc) repliés pour former une seule unité structurale globulaire. Dans chaque motif, on note l'existence d'une boucle qui s'étend entre les deux paires de cystéines liant le métal. La boucle du second motif, formée de l'enchaînement de neuf acides aminés, n'a pas une structure bien déterminée en solution. Dans chaque motif il existe une structure secondaire en hélice α qui débute à partir du résidu suivant la troisième cystéine complexant le zinc et qui s'étend sur 11-13 résidus. En absence de zinc, les récepteurs comme les domaines de liaison à l'ADN recombinants perdent la propriété de se lier au HRE. Ceci est vraisemblablement dû à l'absence d'organisation structurale induite par les complexes zinciques. En effet de nouvelles données de RMN-2D apportent maintenant la preuve que pour chaque motif le complexe de zinc induit le départ de l'hélice α , tandis qu'en l'absence de métal leur structure est désordonnée (D. Rhodes, communication personnelle).

Les deux hélices ont un haut potentiel amphipatique et sont arrangées entre elles de manière à ce qu'elles se croisent en angle droit approximativement en leur centre (figure 15). Les chaînes latérales hydrophobes de ces hélices forment une région hydrophobe importante impliquant au moins onze résidus parmi lesquels six sont invariants alors que les autres ont leur nature hydrophobe conservée dans les récepteurs de la superfamille. Ce noyau hydrophobe représente une armature importante de la structure globale.

A partir de la connaissance de la structure des domaines de liaison à l'ADN de ces deux récepteurs, il est clair que la plupart des acides aminés conservés dans la formule des récepteurs hormonaux jouent un rôle structural important. Ceci est bien illustré par les résidus hydrophobes formant une région hydrophobe stabilisante entre les deux hélices. La conservation de ces acides aminés importants du point de vue structural, ainsi que la similitude entre les deux structures suggèrent que ce motif représente une structure générale de l'interaction protéine-ADN dans les protéines de la superfamille des récepteurs aux hormones stéroïdiennes/thyroïdiennes.

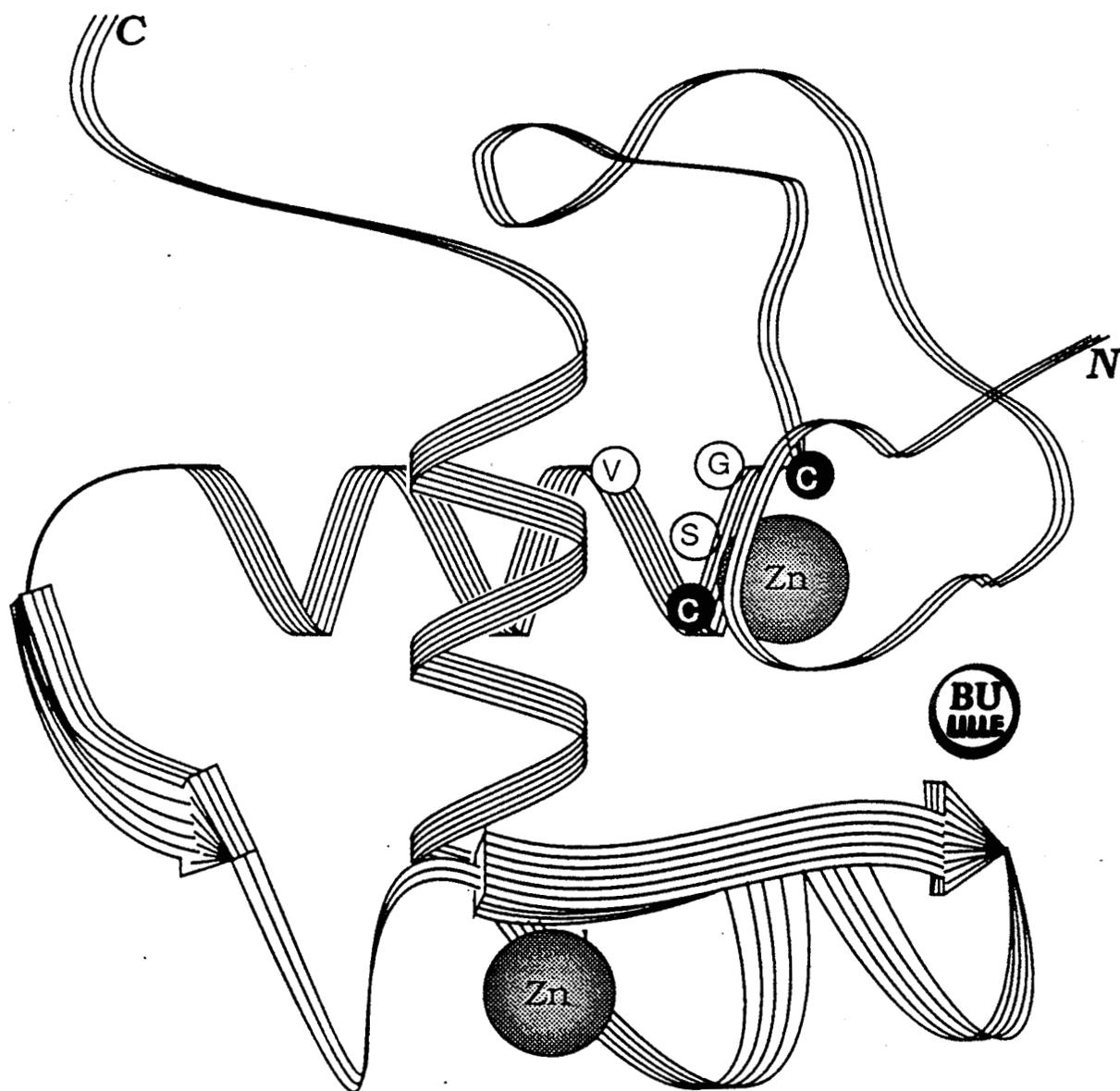


Figure 15. Représentation schématique de la structure tridimensionnelle du domaine de liaison à l'ADN du récepteur des glucocorticoïdes. (a) Les feuilletts β sont représentés par des flèches. Les ronds noirs représentent 2 atomes de zinc. (b) Les deux hélices α sont représentées par des cylindres et les atomes de zinc par des ronds blancs (d'après Härd et al., 1990).

En se référant à la structure de leur domaine de liaison à l'ADN, peut-on comprendre de quelle manière les récepteurs hormonaux se lient à l'ADN et la spécificité de séquence de cette liaison ? Quatre types d'observations ont contribué à la compréhension de l'interaction des récepteurs hormonaux avec l'ADN. En premier lieu, la comparaison des séquences des éléments de réponse hormonale reconnus par les différents récepteurs a montré l'importance des différences que les récepteurs doivent discriminer. Deuxièmement, les expériences de protection d'ADN et de retard sur gel ont révélé la nature de l'interaction entre la protéine et l'ADN. Troisièmement, des expériences utilisant des protéines comportant des substitutions spécifiques d'acides aminés ont identifié les résidus impliqués dans la discrimination des différents sites de liaison. Enfin, l'examen de la structure des domaines de liaison à l'ADN a mis à jour les caractères importants pour la reconnaissance de l'ADN.

Les récepteurs des hormones stéroïdiennes/thyroïdiennes se lient sous forme de dimères à des sites courts de séquence apparentée (cf introduction). Bien que les éléments de réponse hormonale déterminés *in-vivo* présentent des diversités dans leurs séquences, des éléments consensus ont pu être identifiés qui s'avèrent au moins aussi actifs que les éléments naturels. Ces éléments de réponse consensus sont reportés à la figure 16.

Chaque palindrome possède deux demi-sites de six paires de bases. L'ERE et le GRE ne diffèrent seulement que de deux paires de bases dans la séquence du demi-site, tandis que ERE et TRE ont des séquences de demi-site identiques mais qui sont espacées différemment. L'espacement dans le GRE et le TRE est tel que les deux demi-sites sont approximativement situés sur la même face de la double hélice d'ADN alors que les demi-sites du TRE sont situés de part et d'autre de la double hélice. En outre, certains récepteurs reconnaissent des éléments constitués de la répétition directe de deux demi-sites (contrairement au palindrome). Un homodimère de récepteur doit donc discriminer à la fois différentes séquences de demi-sites (ex GRE et TRE) mais aussi l'espacement et/ou l'orientation de ces demi-sites (ex ERE et TRE)

Les études de liaison à l'ADN utilisant la technique des empreintes et le retard sur gel (à la fois sur le récepteur entier ou sur le domaine recombinant de liaison à l'ADN pour le récepteur des glucocorticoïdes et celui des œstrogènes) montrent que le dimère de protéine entretient des contacts avec deux grands sillons adjacents sur une même face de la double hélice avec une symétrie parfaite centrée sur le milieu de la séquence palindromique de l'HRE. De plus, les domaines de liaison à l'ADN isolés se fixent à l'ADN de façon hautement coopérative sous la forme de deux monomères plutôt qu'un dimère préformé (Schwabe *et al.*, 1990).

GRE

MMTV -134 T G G T T T G G T A T C A A A T G T T C T G A T C T G -108
 MMTV -187 T T T A T G G T T A C A A A C T G T T C T T A A A A C -161
 hGH +87 C C T T T G G G C A C A A T G T G T C C T G A G G G G +113
 MSV -171 C A T C T G G G G A C C A T C T G T T C T T G G C C C -197
 MSV -245 T T C A G C T G T T C C A T C T G T T C T T G G C C C -271
 hMTIIa-268 G C A C C C G G T A C A C T G T G T C C T C C C G C T -242
 TO -439 C T C A T A T G C A C A G C G A G T T C T A G T G A G -413
 TO -1174 T G C T C C C T T T C A T G A T G T C C T G G C C C A -1200
 TAT -2420 T A C G C A G G A C T T G T T T G T T C T A G T C T T -2446
 TAT -2515 C T C T G C T G T A C A G G A T G T T C T A G C T A C -2489

<----->

consensus

T G N A C A N N N T G T T C T

ERE

xVit -338 A A A G T C A G G T C A C A G T G A C C T G A T C A -315
 cVit -625 T A T T C C T G G T C A G C G T G A C C G G A G C T -607
 cOvl -173 T T A T T C A G G T A A C A A T G T G T T T T C T G -199
 rPrl -1587 G C A T T T T T G T C A C T A T G T C C T A G A G T -1562

<----->

consensus

A G G T C A N N N T G A C C T

IRE

rGH -185 G T A A G A T C A G G G A C G T G A C C G C A G G -161

<----->

consensus

G A T C A N N N N N T G A C C

Figure 16. Séquence des éléments de réponse hormonale de différents récepteurs hormonaux nucléaires.

Les études de mutagenèse dirigée sur différents récepteurs ont révélé que les résidus de deux régions du domaine de liaison à l'ADN interviennent dans la discrimination des éléments de réponse (Mader *et al.*, 1989; Umesono et Evans, 1989). Dans la première région, trois résidus du premier complexe de zinc sont nécessaires et suffisants pour causer une reconnaissance différentielle de l'ERE et du GRE : Glu25, Gly26 et Ala 29 dans le récepteur des œstrogènes contre Gly, Ser et Val pour le récepteur des glucocorticoïdes (Mader *et al.*, 1989). Les résidus occupant ces positions sont donc clairement impliqués dans la reconnaissance de séquence d'un demi-site d'élément de réponse.

Dans la seconde région, cinq autres résidus, Pro44-Gln48, situés dans le second complexe de zinc sont impliqués dans la reconnaissance de l'espacement entre deux demi-sites (ex ERE et TRE) (Umesono et Evans, 1989). En corrélation avec la coopérativité de liaison à l'ADN pour les domaines

de liaison à l'ADN de l'ER et du GR, ceci suggère que des contacts protéine-protéine dans cette région influencent l'espacement et l'orientation relative de deux monomères de protéine. ⁶⁶

Comprendre comment les résidus identifiés par mutagenèse sont déterminants dans la discrimination entre deux éléments de réponse hormonale implique de connaître la localisation de ces résidus dans le contexte de la structure tridimensionnelle de la protéine. Ces trois résidus (Gly, Ser et Val) sont situés à l'extrémité N-terminale de la première hélice α (figure 15). Ils pointent sur la face externe de l'hélice, exposés au solvant. Sur la même face de l'hélice, on trouve trois résidus basiques conservés (2 lysines et 1 arginine). La nature même de cette hélice suggère donc qu'elle fasse fonction d'hélice de reconnaissance: les acides aminés discriminants entreraient en contact avec les bases qui diffèrent selon l'élément de réponse hormonale tandis que les acides aminés basiques conservés interagiraient plutôt avec les bases conservées dans les différents éléments de réponse hormonale ou bien avec le squelette phosphate sucre.

La prise en compte du processus de dimérisation implique que la reconnaissance moléculaire de l'ADN par un motif en doigt de zinc CI des récepteurs hormonaux nucléaires s'adresse à deux fois un demi-site d'élément de réponse plutôt qu'à l'élément de réponse hormonale entier. A ce point de l'étude, on se focalise donc sur la reconnaissance d'un demi-site du GRE (6 paires de bases) par le motif en doigt de zinc CI du récepteur des glucocorticoïdes et surtout tenant compte de sa composante apparemment essentielle : l'hélice α .

Des études de modélisation moléculaire ont prédit depuis longtemps qu'une hélice α peut être ajustée à la forme du grand sillon d'une double hélice d'ADN sous sa conformation B (Zubay *et al.*, 1959; Sung *et al.*, 1970; Adler *et al.*, 1972; Warrant et Kim, 1978). Il semble qu'à travers l'étude des différents motifs protéiques de liaison à l'ADN une hélice α joue très souvent un rôle prépondérant dans la reconnaissance. Une hélice α d'acides aminés, tenant compte des chaînes latérales, a un diamètre d'environ 12 Å alors que le grand sillon d'un ADN B a une largeur de 12Å et une profondeur de 6-8Å. La surface d'une hélice α peut donc parfaitement s'ajuster à l'intérieur du grand sillon (figure 7). La première condition à une reconnaissance moléculaire est donc satisfaite: il y a complémentarité de deux volumes.

Le squelette de l'hélice α peut être considéré comme une sorte d'échafaudage à partir duquel les chaînes latérales des acides aminés peuvent s'étendre et entrer en contact avec les plateaux de paires de bases par le grand sillon. Le nombre de paires de bases pouvant être contactées par une seule hélice α dépend de l'orientation de l'hélice par rapport au sillon et de la longueur des chaînes latérales. Un maximum de contacts est atteint lorsque l'hélice est parallèle à la direction locale du grand sillon. Avec cet arrangement, les chaînes latérales des acides aminés sont capables de contracter des interactions avec 4 à 6 paires de bases. Une extension de ce contact à un plus grand nombre de paires de bases est impossible par le simple fait que l'axe de l'hélice α est droit tandis que le sillon de l'ADN se courbe dans les deux directions. Cependant, des études de modélisation moléculaire indiquent que l'hélice α ne doit pas nécessairement être strictement parallèle à l'axe du grand sillon. Une variété d'orientations sont stériquement envisageables et de plus, certaines orientations permettraient un contact plus étendu avec l'ADN. Par exemple, une hélice α peut être positionnée de façon à ce que l'une de ses extrémités, plutôt que son centre, soit proche de l'axe de la double hélice et l'hélice α peut s'écarter d'un angle de 15 à 20° par rapport au grand sillon. Un tel arrangement est similaire à celui observé pour la protéine répresseur du phage lambda; il permet encore une reconnaissance de 4 à 5 paires de bases.

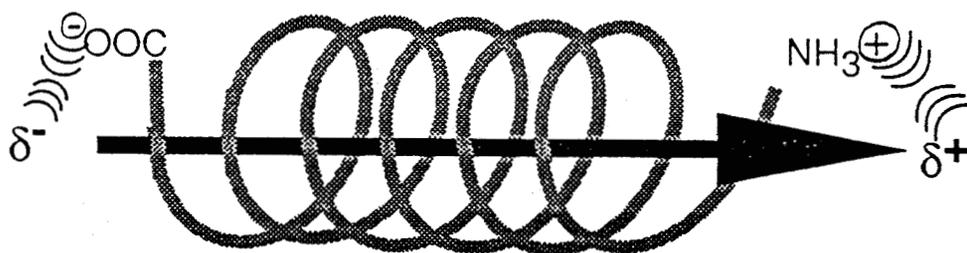


Figure 17. Dipôle induit par une structure en hélice α .

Il existe probablement une préférence pour un positionnement de l'extrémité N-terminale, plutôt que l'extrémité C-terminale, proche de l'axe de la double-hélice d'ADN. Cette préférence est d'ordre électrostatique; en effet, l'arrangement des dipôles créés par les liaisons intra-caténares induit la formation d'un dipôle global pour l'hélice α (figure 17). Il s'ensuit la création d'une charge partielle positive à l'extrémité N-terminale de la structure en hélice alors que le grand sillon de l'ADN est porteur d'une charge négative partielle (Pabo et Lewis, 1982, Ho' et al., 1978; Pullman et Pulmann, 1981). Ce phénomène aurait un effet coopératif sensible dans le bon positionnement des chaînes latérales des acides aminés proches de l'extrémité N-terminale pour interagir avec les paires de bases.

Les résultats relatifs à la structure tridimensionnelle du domaine de liaison à l'ADN de récepteurs hormonaux nucléaires ainsi que la réflexion sur la large utilisation de l'hélice α dans les motifs protéiques de liaison à l'ADN, nous ont amené à reconsidérer notre conception d'un motif peptidique de reconnaissance d'ADN sur le modèle du doigt de zinc CI du récepteur des glucocorticoïdes.

Par des études de prédiction de conformation, nous avons décelé depuis le début de ces études que la région située immédiatement en C-terminal du complexe de zinc de la partie CI était susceptible de former une hélice α . Les données de RMN-2D sur le domaine de liaison à l'ADN des récepteurs de l'œstrogène (Schwabe *et al.*, 1990) et celui des glucocorticoïdes (Hård *et al.*, 1990) sont venues confirmer ces présomptions et ont amené un autre élément important : la géométrie du complexe tétraédrique de zinc maintient les résidus situés entre la 3ème et la 4ème cystéine ligand du métal dans une structure très similaire à celle d'une hélice α et induit efficacement l'hélicité de la séquence juxtaposée (figure 15). De ce fait, les trois acides aminés essentiels pour la reconnaissance sont situés à l'extrémité d'une hélice α . La technique de modélisation moléculaire nous a permis de comparer les deux structures tridimensionnelles de cette région précise: les positionnements du squelette carboné et des chaînes latérales dans la structure tridimensionnelle induite par la géométrie du complexe de zinc sont superposables à ceux induits par la géométrie d'une hélice α (figure 18).

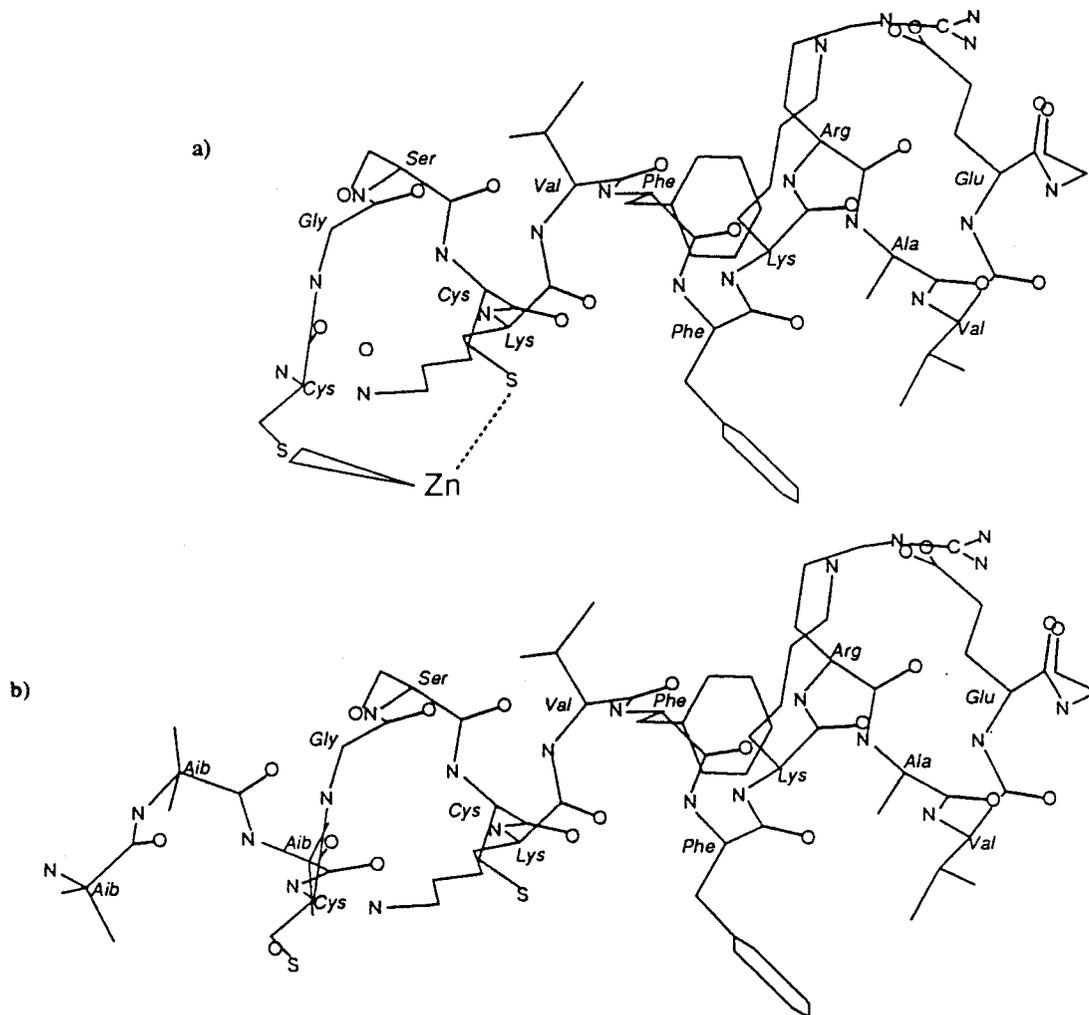


Figure 18. Structure tridimensionnelle de la région de reconnaissance du doigt de zinc CI du récepteur des glucocorticoïdes. a) Structure imposée par la géométrie du complexe de zinc. b) Structure en hélice α .

Dans la séquence du domaine de liaison à l'ADN des récepteurs hormonaux nucléaires (figure 12), les deux résidus situés entre la 3ème et la 4ème cystéine ligand du métal sont toujours de nature incompatible avec une structure en hélice α . En fait, ces acides aminés (Gly, Glu et Ser) sont même réputés pour être des "briseurs" d'hélice α (a fortiori lorsqu'ils sont adjacents). Nous avons donc émis l'hypothèse que toute la structure complexe du doigt de zinc CI avait pour finalité d'imposer une structure secondaire en hélice α (apte à la reconnaissance du grand sillon) à une région comportant des acides aminés incompatibles avec une telle structure. On peut expliquer alors la dépendance du zinc observée pour la liaison du récepteur à l'ADN.

Nous avons eu l'idée d'imposer à la séquence de reconnaissance une structure en hélice α , non plus par le biais du complexe de zinc, mais par l'incorporation d'un acide aminé particulier connu pour ses propriétés de fort inducteur d'hélice : l'acide α -aminoisobutyrique (Aib) ou α -méthylalanine (figure 19).

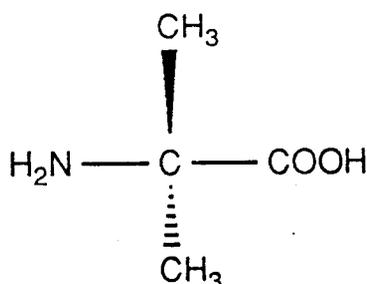


Figure 19. Structure de l'acide α -aminoisobutyrique (Aib)

L'acide α -aminoisobutyrique (figure 19) est un acide aminé achiral d'origine naturelle que l'on retrouve dans la structure d'antibiotiques de la famille des peptaibols par exemple (Venkataram Prasad et Balaram, 1981). Par synthèse peptidique en phase solide nous avons obtenu le peptide correspondant à la séquence Cys128-----Glu156 du récepteur des glucocorticoïdes (structure en hélice, figure 14). A chaque extrémité de la séquence nous avons ajouté trois résidus inducteurs d'hélice Aib. La synthèse d'un tel peptide tenant compte de la structure hélicoïdale comporte une contrainte: il faut normalement garder les extrémités COOH et NH₂ protégées. En effet, le zwitterion en solution possède des charges de même signe que la charge partielle induite par le dipôle hélice, ce qui a pour effet de déstabiliser la structure par simple répulsion de ces charges entre elles (figure 17). Toutefois, nous avons délibérément déprotégé les extrémités du peptide afin d'obtenir une synergie des charges positives dans le positionnement éventuel de l'hélice α dans le grand sillon de charge partielle négative.

Nous avons synthétisé les deux peptides HxP3 et HxP6 dont les séquences sont les suivantes:



Les deux peptides ont été obtenus par une méthode de synthèse peptidique en phase hétérogène sur un appareil semi-automatique LKB Biolynx 4175 utilisant la chimie des Fmoc (Fields et Noble, 1990) et le BOP comme agent de couplage des résidus Aib (Frérot *et al.*, 1991).

Les peptides complets sont clivés de la résine et déprotégés par un mélange TFA 95%/thioanisole 5% pendant trois heures à température ambiante. Ces conditions préservent les résidus cystéines sous leur forme tert-butyl mercaptan.

Chaque peptide est purifié par chromatographie d'exclusion stérique (Biogel P2, 200-400 mesh, 10% acide acétique) et leur pureté est vérifiée par chromatographie liquide à haute performance en phase inverse par passage isocratique d'acétonitrile à 20% sur une colonne de type C18.

L'identité des peptides synthétiques purifiés est vérifiée par spectrométrie de masse à temps de vol (HxP3: m=1822; HxP6: m=2084) ainsi que par analyse de la composition en acides aminés après hydrolyse acide (HCl 5-6N, 110 ° C, 24H, sous-vide).

La première chose à vérifier était que la structure induite aux peptides par les résidus Aib fût superposable à celle induite dans la région correspondante du récepteur par la géométrie du complexe de zinc. Nous avons effectué une étude de modélisation assistée par ordinateur produisant les structures d'énergie minimale dans le cas du complexe de zinc et d'une structure possédant les résidus Aib. Les calculs de modélisation ont été effectués au moyen de l'algorithme de champs de force d'Allinger (Allinger, 1976; Sprague *et al.*, 1987; Allinger *et al.*, 1988) qui est un outil du programme Chem 3D plus 2.0.1 (M. Rubenstein et S. Rubenstein, Cambridge Scientific Computing Inc., Cambridge, Massachusetts, U.S.A.).

Initialement un modèle est construit au moyen de sous structures préformées et d'énergie minimisée avec les acides aminés sous leur configuration β , disponibles dans la section MM2. Ce modèle subit alors une minimisation d'énergie. Les conformères de basse énergie sont recherchés par sondage de différents contacts potentiels de liaison hydrogène et les modèles résultants sont resoumis à un traitement de dynamique moléculaire utilisant une température cible de 310° K, une cinétique d'augmentation de température de 3,0 KCal/atom/psec à intervalles de 10 fsec et un temps d'équilibre de 3,0 psec.

Les résultats obtenus sont reportés à la figure 18. On note que chacune des structures est superposable à l'autre par son squelette peptidique et, chose encore plus importante, les groupements des acides aminés discriminants pour la reconnaissance (Gly, Ser et Val) pointent dans la même direction dans chacune des deux structures.

Donc, d'un point de vue théorique, nos peptides sont capables d'occuper le même volume spatial que dans le cas où ils sont complexés au zinc dans la région correspondante du récepteur et ils ont *a priori* une répartition tridimensionnelle analogue des groupements chimiques et des potentiels électrostatiques.

Il nous fallait maintenant vérifier que dans la pratique nos peptides fussent capables d'adopter une telle structure ordonnée en solution. Nous avons choisi pour cela d'utiliser la technique spectroscopique de dichroïsme circulaire.

Caractéristiques structurales des peptides HxP3 et HxP6.

Les spectres des peptides ont été enregistrés de 250 à 196 nm sur un spectromètre Jobin-Yvon R.J. mark III à température ambiante dans des cuves de 0,1 et 1 mm de trajet optique. Les concentrations en peptide étaient de 0,2mM dans l'eau ou le trifluoroéthanol.

Les spectres obtenus sont reportés à la figure 20.

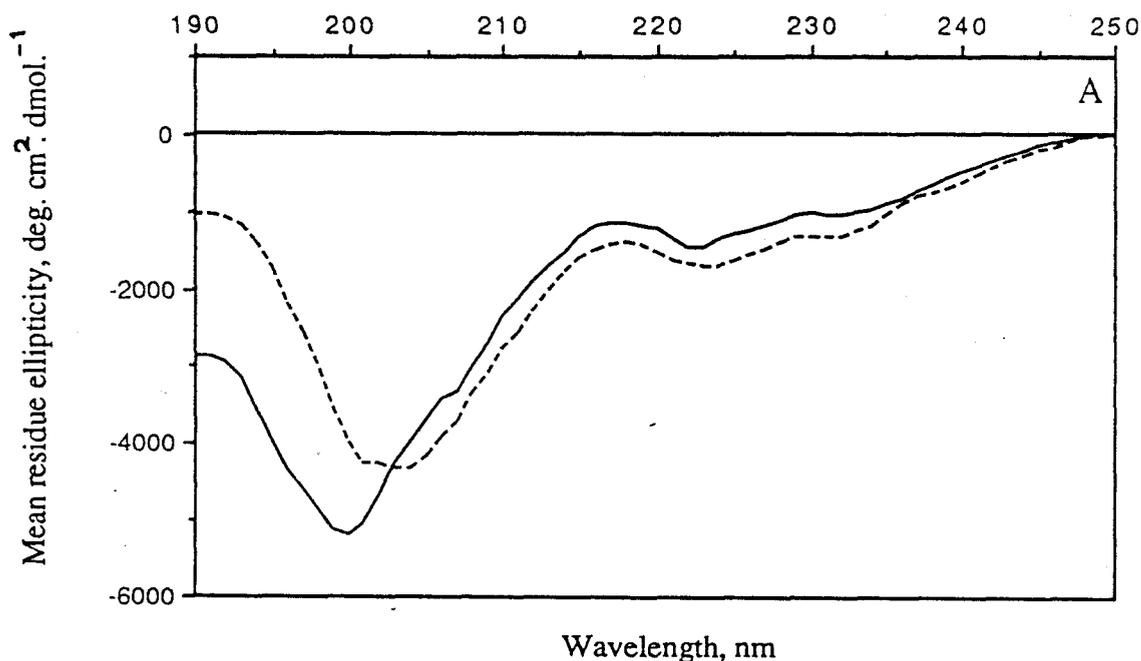


Figure 20. Spectres de dichroïsme circulaire des peptides HxP3 (—) et HxP6 (---) en solution 0,2mM dans l'eau.

On note premièrement que les deux peptides HxP3 et HxP6 donnent un spectre très peu différent dans l'eau. En solution aqueuse on note la présence de deux minima à 223 et 203 nm (HxP6), et 223 et 200 nm (HxP3) ce qui correspond à un mélange de conformations comprenant une pelote statistique et une hélice. Cette hélice peut être soit de type α soit de type 3_{10} , en effet il est très difficile de discerner les deux types de structures selon leurs seules propriétés optiques chirales (Venkataram Prasad et Balaram, 1984). Cependant, de façon à confirmer la tendance de ces peptides à adopter une structure ordonnée, nous les avons étudié en solution dans le trifluoroéthanol, un solvant bien connu pour entretenir de telles conformations. Dans ces conditions (figure 21), les spectres obtenus sont caractéristiques d'hélices α avec HxP6 étant plus hélicoïdal qu'HxP3.

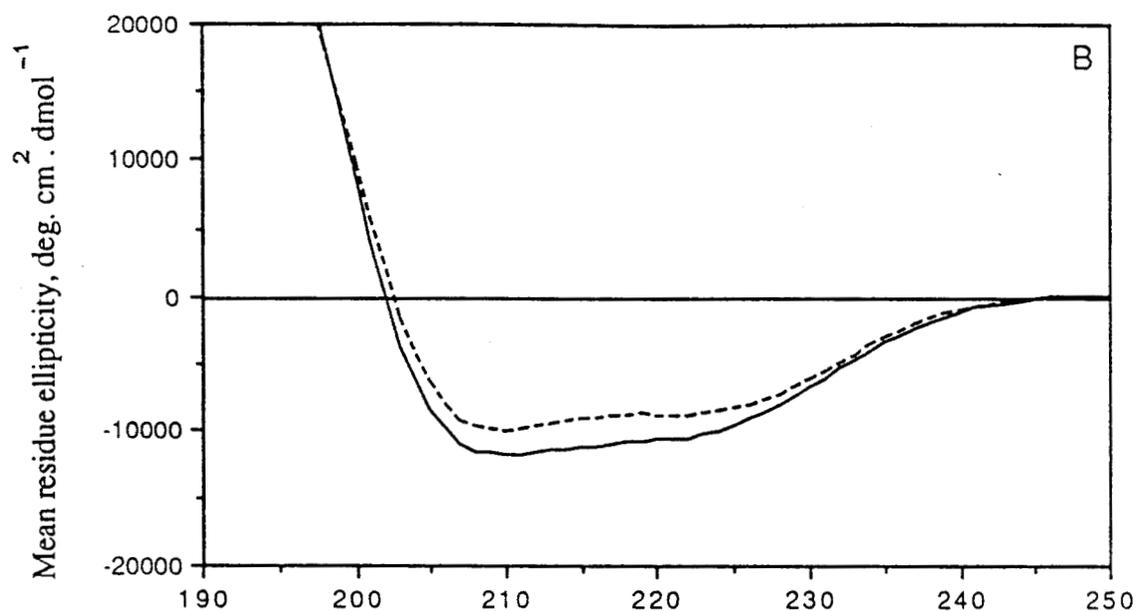


Figure 21. Spectres de dichroïsme circulaire des peptides HxP3 (—) et HxP6 (---) en solution 0,2mM dans le trifluoroéthanol.

Interaction des peptides HxP3 et HxP6 avec l'élément de réponse aux glucocorticoïdes.

Nous avons décidé dans un premier temps de tirer parti des propriétés spectroscopiques dichroïques des peptides pour étudier en présence de l'oligonucléotide correspondant au GRE et un autre oligonucléotide proche: l'élément de réponse à l'hormone thyroïdienne ou TRE.

5' -AGCTTAGAACAGTTTGTAACCA
ATCTTGTCAAACATTGGTTCGA -5'

GRE

5' -AGCTTTCAGGTCATGACCTGAA
AAGTCCAGTACTGGACTTTCGA -5'

TRE

Ces expériences d'interactions sont menées dans un tampon propice à l'interaction; Tris/HCl 10 mM (pH 7,5), dithiotréitol 0,1mM, EDTA 0,1 mM, glycérol 10% (vol:vol), sérumalbumine bovine 3 µg/µl et NaCl 100mM à température ambiante.

Les spectres dichroïques des peptides en interaction sont représentés à la figure 22. On observe dans chaque cas une variation de l'ellipticité dans la région 229-230 nm, Cette variation correspond à un changement de conformation du peptide lorsqu'il se lie à l'oligonucléotide. On note que ce changement intervient pour un rapport peptide/oligonucléotide plus élevé dans le cas du TRE que dans le cas de l'interaction avec le GRE. On peut donc déjà suggérer une préférence de nos peptides pour interagir avec le GRE plutôt qu'avec le TRE.

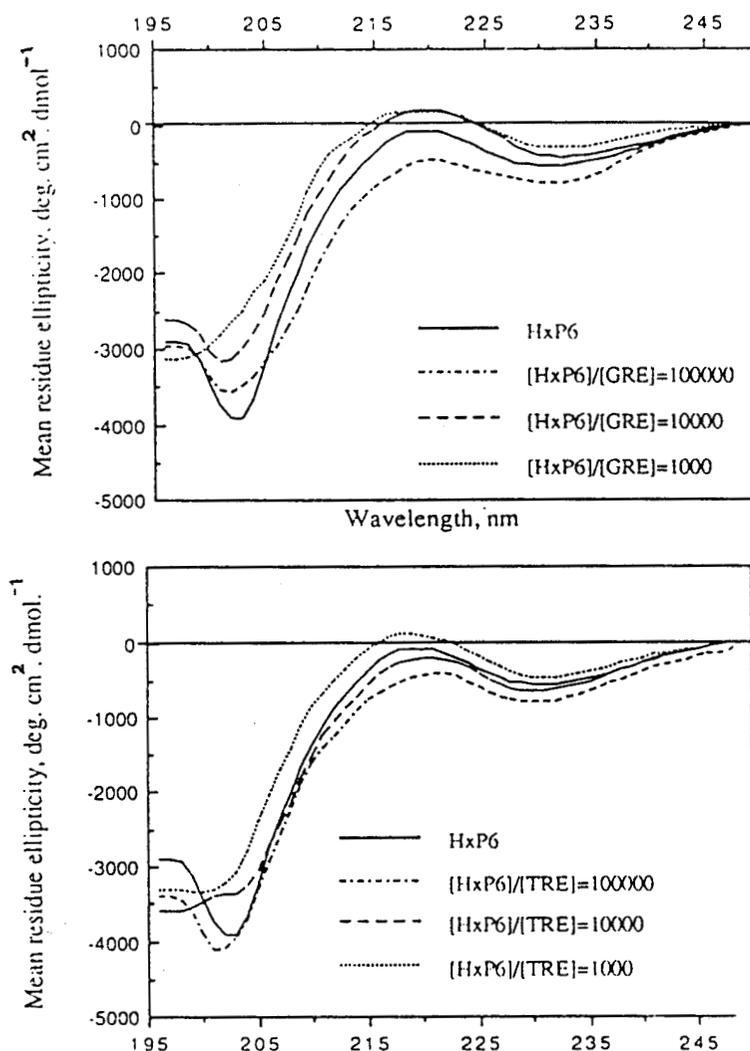


Figure 22. Spectres de dichroïsme circulaire de l'interaction de HxP6 avec l'oligonucléotide GRE (haut) et TRE (bas)
(Les oligonucléotides sont ajoutés à partir d'une solution mère 0,02mM.)

Nous avons ensuite réalisé une expérience préliminaire pour montrer l'interaction par la technique de retard sur gel. Il ne nous a pas été possible de démontrer directement l'interaction entre nos peptides et le GRE par apparition d'une bande retardée en électrophorèse en gel de polyacrylamide, c'est pourquoi nous avons utilisé un protocole indirect de compétition en retard sur gel. Cette expérience, décrite en détail à la page 48 de ce mémoire possède l'avantage de démontrer une interaction entre le peptide et l'oligonucléotide en compétition avec l'interaction du récepteur; c'est à dire que, dans le cas où l'on observe une interaction, on peut suggérer que celle-ci est de même nature que celle du récepteur entier.

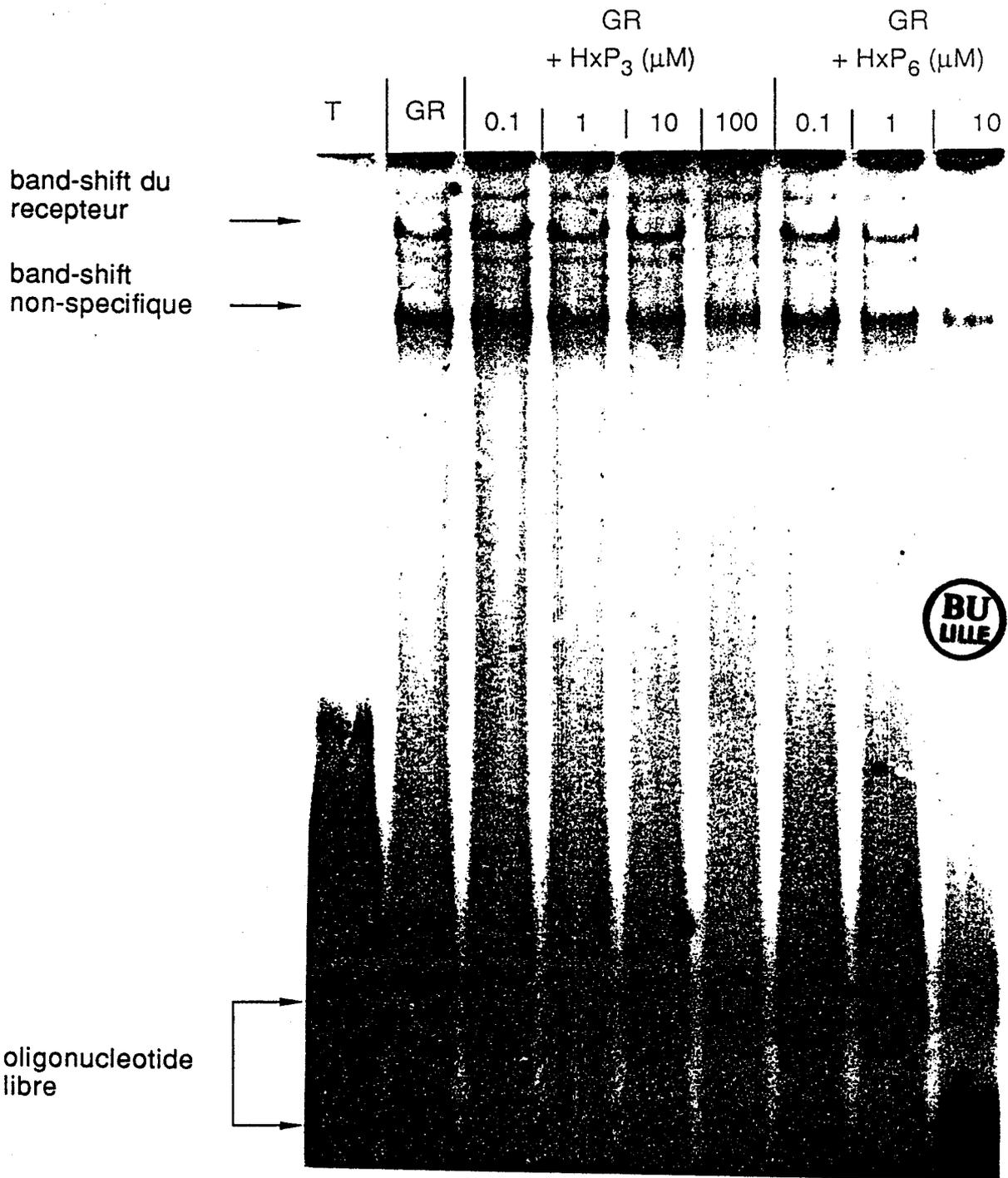


Figure 23. Compétition en retard sur gel entre les peptides HxP3 et HxP6 (aux concentrations indiquées et un extrait cytoplasmique enrichi en récepteur des glucocorticoïdes (GR))
La ligne T correspond à la migration de l'oligonucléotide GRE seul.

Cette expérience de compétition en retard sur gel a été réalisée comme suit: un oligonucléotide double-brin correspondant à la séquence de l'élément de réponse aux glucocorticoïdes est marqué aux extrémités 5' par incorporation de ^{32}P par la polynucléotide kinase de phage T4 (New England Biolabs) à partir de $[\gamma\text{-}^{32}\text{P}] \text{ATP}$ (3000Ci/mmol. Amersham). 2 à 5 fmoles (5.000-10.000 cpm) de GRE marqué sont incubés avec 1 μl de récepteur des glucocorticoïdes partiellement purifié, en présence ou non de quantités croissantes en peptide HxP3 ou HxP6, dans un tampon contenant Tris, HCl 10 mM, dithiothréitol 0,1 mM, EDTA 0,1 mM, glycérol 10% (vol:vol), sérumalbumine bovine 3 $\mu\text{g}/\mu\text{l}$ et NaCl 100 mM pendant 30 minutes à température ambiante. Chaque mélange est alors soumis à électrophorèse sur gel de polyacrylamide (rapport acrylamide/bisacrylamide 40:1) dans un tampon Tris,HCl 25 mM (pH 8,6), Glycine 188 mM à 15 mA pendant 50 min dans une minicuve Biorad équilibrée en bain de glace. Le gel est alors séché et autoradiographié.

La photo du gel est reportée à la figure 22. En présence d'extrait cytoplasmique enrichi en récepteur on note l'apparition de deux bandes retardées. La bande supérieure possède une migration identique à celle déterminée comme correspondant au complexe récepteur des glucocorticoïdes activé/GRE lors de la mise au point du protocole (cf p. 48 et suivantes).

Cette bande disparaît lorsque l'on augmente la quantité en peptide compétiteur, alors que la bande inférieure que nous attribuons à l'interaction non-spécifique avec une macromolécule présente dans l'extrait cytoplasmique, ne change pas d'intensité avec l'augmentation de la quantité en peptide.

Un fait important que l'on peut observer est qu'une concentration en HxP3 nécessaire pour éliminer la bande retardée est dix fois supérieure (100mM) à celle requise par le peptide HxP6 (10mM) pour obtenir le même effet.

Afin de confirmer et de bien visualiser l'interaction peptide/GRE nous avons entrepris d'utiliser la technique des empreintes. Le peptide qui reconnaît une certaine séquence d'ADN doit être capable de protéger spécifiquement celle-ci d'une dégradation par des agents chimiques, ici l'attaque du squelette phosphate-sucre par des espèces radicalaires. Pour cela il nous fallait construire un fragment d'ADN de séquence aléatoire contenant la séquence GRE insérée. Nous avons donc cloné la séquence GRE au sein d'un plasmide vecteur. Pour cela, 40 μg de plasmide pBS M13mp18(+) (Boehringer) sont linéarisés par 40 unités d'enzyme de restriction HindIII (Boehringer). On ajoute alors 10 μg d'oligonucléotide double-brin GRE possédant à chaque extrémité la séquence cohésive du site HindIII en présence de ligase de phage T4 (Boehringer).



Après transformation d'une souche JM109 d'*Escherichia coli*, les clones positifs sont sélectionnés par hybridation avec l'oligonucléotide simple brin marqué au ^{32}P à son extrémité 5' par la polynucléotide kinase de phage T4 (Boehringer). Après extraction des plasmides des colonies sélectionnées et multipliées, le fragment EcoRI-PvuII est isolé et marqué par incorporation d' $[\alpha\text{-}^{32}\text{P}]\text{ATP}$ à son extrémité 3' par le fragment klenow de l'ADN polymérase (Boehringer) dans les conditions standards. Après séquençage de ces fragments par la technique de Maxam et Gilbert

(Maxam et Gilbert, 1975), l'un d'entre eux est sélectionné contenant trois copies adjacentes de GRE insérées dans le site HindIII. C'est à ce fragment contenant 3 séquences GRE que l'on fait subir une attaque ménagée par les radicaux hydroxyles (OH°) en présence ou en l'absence de récepteur partiellement purifié ou de peptide HxP6 (Les empreintes ont été également effectuées en présence de peptide HxP3 mais les concentrations en peptide requises pour observer l'interaction étaient trop fortes pour considérer le résultat comme significatif).

La réaction de coupure ménagée par les radicaux hydroxyles est effectuée essentiellement de la manière décrite par Tullius et Dombroski (Tullius et Dombroski, 1984). 3 pmoles (en paires de bases) d'ADN (fragment EcoRI-PvuII+triple insert: 316 bp) sont incubés avec 1µl de récepteur ou de peptide (1-50µM) dans le même tampon que celui utilisé lors de l'interaction en retard sur gel sans EDTA pendant 30 minutes à 30°C. Une solution fraîche de Fe/EDTA est préparée immédiatement avant son utilisation en mélangeant 10µl de Fe(NH₄)₂(SO₄)₂ 0,2mM fraîchement préparé avec 20 µl d'EDTA-NA₂ 0,1M (pH 7,4) et 20 µl d'une solution d'ascorbate de sodium fraîchement préparée. Les radicaux hydroxyles sont générés par la réaction de ce mélange avec 20 µl d'H₂O₂ 0,3%. L'attaque de l'ADN est initiée par l'ajout immédiat de 7,5 µl de cette mixture au complexe récepteur/ADN ou peptide/ADN.

Des aliquots sont prélevés après 1 et 5 minutes et la réaction est stoppée par addition de 5 µl de thiourée 0,1M. Les échantillons subissent alors deux précipitations par l'éthanol à 95% et sont repris dans 2 µl de formamide 100% contenant du bleu de bromophénol. On procède alors à l'électrophorèse en gel de polyacrylamide 8%, urée 8M (0,3 mm d'épaisseur) contenant du tampon Tris-borate -EDTA 1X. La durée de l'électrophorèse est de 2,5H à 1800V. Le gel est alors séché et autoradiographié.

Les autoradiogrammes sont analysés à l'aide d'un microdensitomètre Joyce-Loebl produisant des profils de densitométrie dans lesquels l'intensité relative de chaque bande est mesurée. Les résultats sont reportés sous la forme de clivage fractionné $f=A_i/A_t$ où A_i est l'aire correspondant à la bande i et A_t est la somme des aires correspondant à toutes les bandes. Des pistes de contrôle (sans mixture Fe/EDTA) sont effectuées de chaque coté du gel et utilisées pour assurer une estimation précise des aires relatives des bandes ainsi que pour vérifier que les coupures sont bien dues à l'addition de Fe/EDTA et non pas au résultat d'une contamination. Pour la comparaison des différents profils de coupure, une attention toute particulière est portée pour s'assurer que l'étendue de digestion est très similaire et limitée à 20-40% du matériel de départ afin de minimiser l'incidence de coupures multiples dans un brin.

Sur la figure 24 on voit une photographie de l'autoradiogramme du gel d'empreintes pour le récepteur et le peptide HxP6. La caractéristique la plus remarquable dans ce résultat est que l'on peut observer à l'oeil nu (peu courant dans ce genre de technique) une décroissance de l'intensité de certaines bandes avec un profil identique en présence de récepteur ou de peptide à 5µM. Ce profil d'empreinte comporte dans chaque cas une décroissance répétée coïncidant avec les trois séquences GRE.

Ce profil des bandes n'est pas aussi régulier dans la partie hors-insert du fragment vecteur que dans le cas de l'échelle régulière obtenue dans le contrôle de digestion; ce qui suggère que des interactions se produisent aussi en dehors des séquences GRE insérées. Cependant, il est très clair que cette interaction est moins forte dans cette partie vecteur que dans les inserts et, en outre, elle est totalement différente dans le cas du récepteur que dans celui du peptide. On peut donc conclure que, au moins en ce qui concerne les régions GRE insérées, il y ait interaction sélective avec l'ADN et que le peptide HxP6 semble reconnaître l'ADN exactement de la même façon que le fait le récepteur entier.

Ces résultats sont confirmés par l'analyse des profils densitométriques (figure 25). Il en ressort très clairement que les régions correspondant aux inserts GRE sont des régions hautement protégées de la coupure par les radicaux alors que cette coupure est très prononcée au milieu de cette région. Ce résultat, parfaitement en accord avec celui de Chalepakis *et al.* (Chalepakis *et al.*, 1988) dans le même type d'expérience semble refléter l'interaction d'un dimère de récepteur avec les deux demi-sites du palindromes GRE alors que la région centrale de ce palindromes se trouve exposée à l'attaque chimique par les radicaux de petite taille (les empreintes à la DNase I ne permettent pas de visualiser cette région du GRE dont la protéine semble plus éloignée).

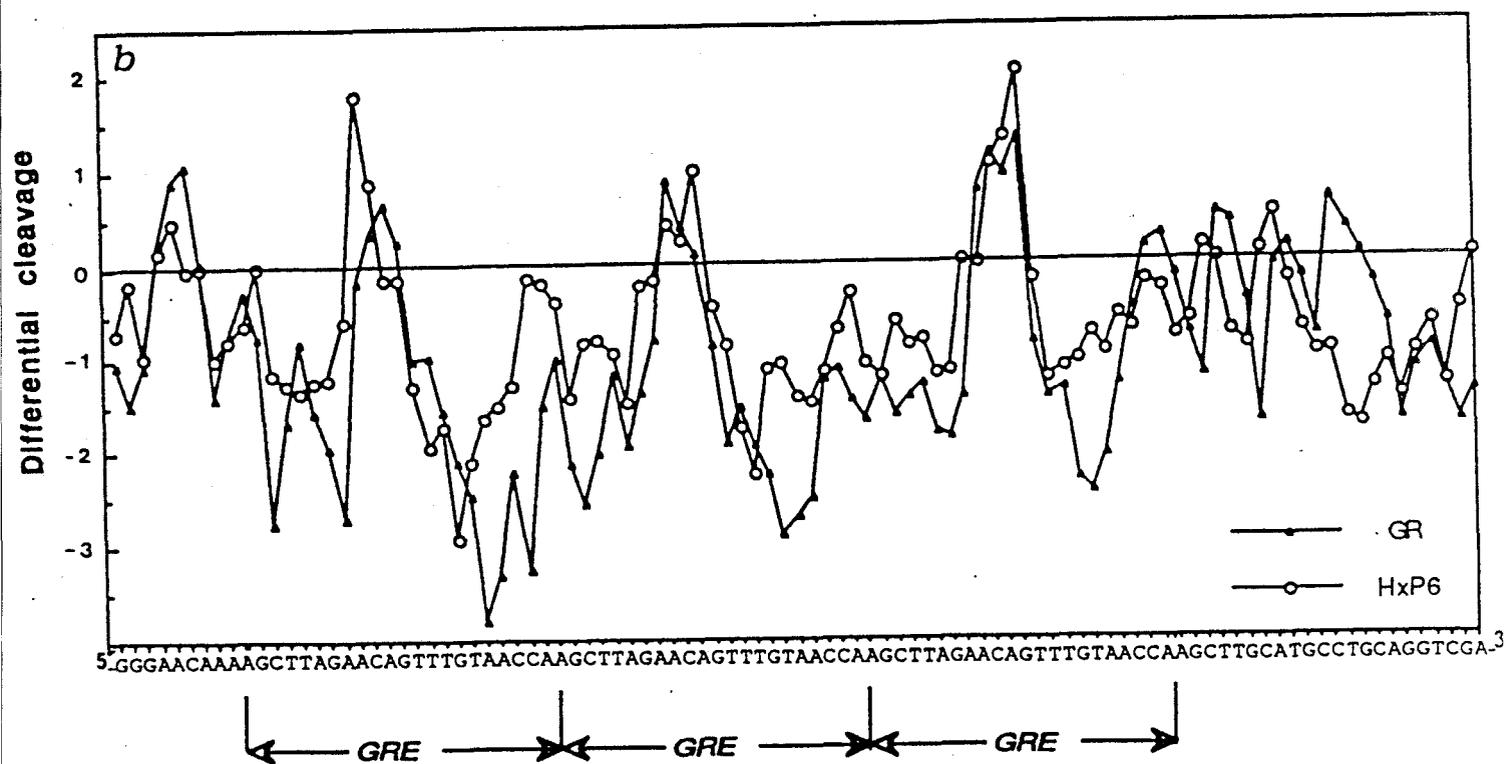


Figure 25. Profil densitométrique des empreintes de l'attaque par les radicaux hydroxyles en présence d'extrait cytoplasmique enrichi en récepteur des glucocorticoïdes (GR) et en présence de peptide HxP6.

Les valeurs positives indiquent une coupure plus prononcée et les négatives une protection

Enfin, la chose la plus importante à noter pour notre étude est que le profil densitométrique de protection avec le peptide HxP6 suit rigoureusement celui correspondant à la protection par le récepteur montrant ainsi le même processus de reconnaissance et suggérant donc la même façon d'interagir avec la double hélice d'ADN. Cependant on peut aussi observer que les intensités relatives de protection sont moindres avec le peptide HxP6 qu'avec le récepteur. Ceci est sans doute à rapprocher avec le résultat du retard sur gel pour lequel on pouvait déjà suggérer que le peptide reconnaît effectivement le GRE mais avec une faible affinité. On peut attribuer cette faible affinité observée au fait que le peptide ne bénéficie pas de la haute coopérativité de fixation qu'a le récepteur lorsqu'il interagit sous sa forme de dimère avec l'ADN. Un argument allant dans le même sens est le résultat d'études d'empreintes à la DNase I (non-montrés) qui donnent le même type de protection mais de façon beaucoup moins prononcée, probablement du au fait que la DNase I est elle-même une protéine affine pour l'ADN et qui doit déplacer le peptide lors de la digestion (ce genre de compétition n'intervient pas dans notre expérience de dégradation de l'ADN par des agents chimiques).

Le travail reporté ci-dessus est basé sur des études antérieures montrant que des caractéristiques importantes de l'interaction doigt de zinc/ADN peuvent être modélisées par de petits peptides obtenus par synthèse chimique.

Le rôle fondamental de trois résidus (Gly458, Ser459 et Val462) dans la spécificité d'interaction du récepteur envers l'ADN, montrés dans un premier temps par des expériences de mutagenèse (Mader et al., 1989), puis renforcées par des études de modélisation consécutives à une détermination de structure en solution par RMN, ajouté à l'observation que la structure hélice α joue un rôle majeur dans la reconnaissance de l'ADN par des protéines nous a amené à proposer un modèle très simplifié de la région du récepteur des glucocorticoïdes responsable de la reconnaissance du GRE.

Les résultats des études d'interaction que nous rapportons ici sont en parfait accord avec le modèle d'interaction proposé par Härd et al. (Härd et al., 1991) où un dimère du domaine de liaison à l'ADN recombinant du récepteur des glucocorticoïdes est artificiellement positionné face à un GRE en modélisation moléculaire. Ce modèle fait apparaître que les hélices α induites par le complexe de zinc de chaque région CI peuvent parfaitement s'ajuster dans le grand sillon de la double hélice sur la même face à un tour de spire d'intervalle tandis que la région de dimérisation du récepteur, située en face du centre du palindrome et plus éloignée de la double hélice.

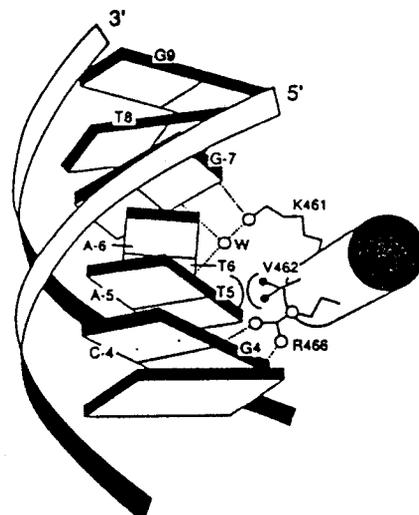
Ce travail préliminaire semble renforcer l'idée que les principaux motifs protéiques de reconnaissance d'ADN utilisent en fait un mode œcuménique d'interaction avec le grand sillon de l'ADN par l'intermédiaire de leur organisation régionale en hélice α . Dans le cas présent de la boucle à zinc CI du récepteur des glucocorticoïdes (ceci pourrait être étendu à d'autres membres de la superfamille), le complexe tétraédrique de zinc induit la structure en hélice α pour une région contenant des acides aminés (essentiels pour la reconnaissance) qui sont réputés pour être des briseurs de telles structures. En fait, tout ce passe comme si de tels résidus auraient été contraints d'être au sein

d'une telle structure figée autour d'un atome de zinc pour qu'ils puissent participer à la reconnaissance de l'ADN, permettant ainsi de nouvelles combinaisons d'interaction pour les protéines.

Des travaux sont en cours au laboratoire pour voir si de tels peptides synthétiques de structure induite en hélice α et correspondant aux séquences de la même région d'autres membres de la super famille des récepteurs hormonaux nucléaires sont capables de reconnaître sélectivement leur HRE correspondant. Si c'était le cas au moins pour certains d'entre eux, des études plus fines de ces interactions pourraient alors contribuer au décryptage du code mythique de reconnaissance protéine/ADN pour cette super famille de facteurs de transcription.

Très récemment est paru un article de Luisi et al. (Luisi et al., 1991) rapportant l'analyse cristallographique du complexe: domaine de liaison à l'ADN du récepteur des glucocorticoïdes /élément de réponse hormonale aux glucocorticoïdes. Les résultats relatifs à l'interaction spécifique de la région CI avec un demi-site de l'élément de réponse sont en parfaite harmonie avec nos concepts et viennent confirmer nos hypothèses. Entre autres, la reconnaissance d'un demi-site du GRE s'effectue via une hélice α qui pointe son extrémité N-terminale dans le grand sillon de la double hélice d'ADN (figure 26).

Figure 26 Représentation schématique de l'interaction entre l'hélice α du domaine CI de l'hGR et un demi-site du GRE

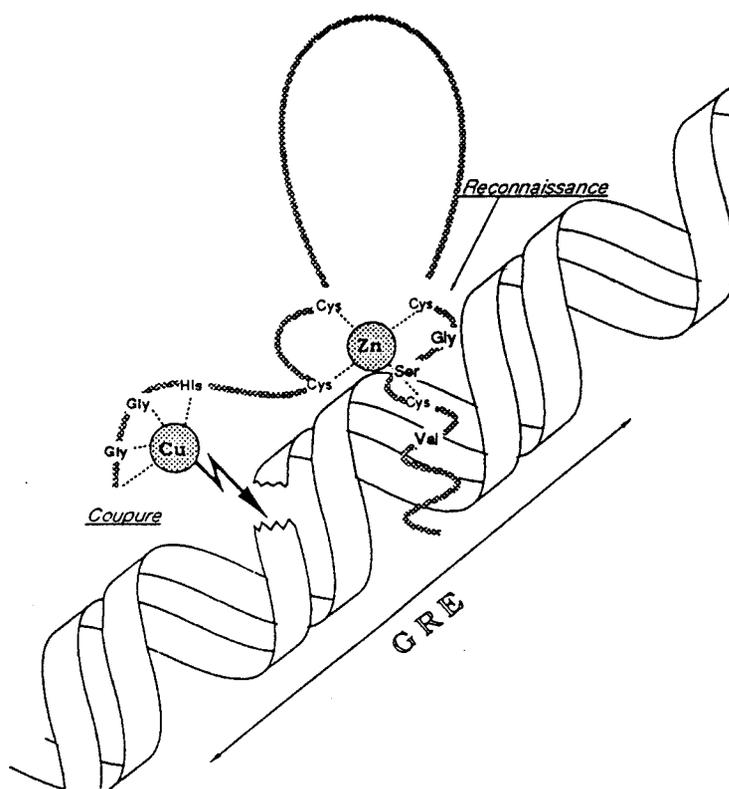


Cet arrangement positionne les chaînes latérales des acides aminés proches de l'extrémité N-terminale de manière à interagir avec la succession de plateaux de paires de bases du demi-site de GRE. Les résidus basiques conservés de cette hélice α parmi les différentes classes de récepteurs interagissent avec les bases conservées parmi les différents classes d'élément de réponse (figure 27). L'acteur essentiel de la spécificité de reconnaissance est une liaison de type van der Waals entre le méthyle γ de la valine 462 (numérotation hGR) et le groupement 5-méthyle de la thymine 5 du demi-site de GRE (figure 26). Il est alors permis d'émettre des hypothèses quant la discrimination entre GRE et ERE par seulement trois acides aminés. Les diverses classes d'éléments de réponse hormonale diffèrent essentiellement sur leurs positions 5 et 6 (positions relatives au centre de symétrie rotationnelle, figure 27).

La conclusion qui peut être tirée des études de ce chapitre vient corroborer celle énoncée à la fin du chapitre concernant les peptides signal de localisation nucléaire. Il est possible d'identifier au sein d'une protéine reconnaissant l'ADN un motif peptidique relativement court, responsable à lui seul de la spécificité de reconnaissance. Cette séquence, isolée de tout contexte protéique (peptide synthétique), garde la même propriété de reconnaissance si et seulement si il est tenu compte de sa structure tridimensionnelle.

Cependant, le but que nous nous étions fixé dans le cadre de la conception d'une molécule synthétique capable d'agir sur une séquence d'ADN donnée n'est pas encore atteint. Les peptides inspirés du motif CI du récepteur des glucocorticoïdes sont aptes à la reconnaissance d'une séquence de six paires de bases mais ils n'ont pas d'action directe sur la molécule d'ADN. C'est pourquoi, dans une troisième partie nous avons utilisé ce motif de reconnaissance comme "outil de reconnaissance" pour cibler une activité de coupure sur une séquence précise d'ADN. La reconnaissance d'une séquence de six paires de bases et la coupure dans cette séquence ou à proximité forment une double fonctionnalité que l'on retrouve chez un autre type de macromolécules biologiques : les enzymes de restriction des systèmes procaryotes. Autrement dit, le chapitre suivant qui traite de l'association d'un agent de coupure d'ADN avec un motif de reconnaissance nous engage dans la conception d'une endonucléase artificielle.

ELABORATION D'UNE ENDONUCLEASE ARTIFICIELLE



Production de radicaux libres et coupure d'ADN.

Un complexe, formé d'un agent chélateur et d'un métal approprié est capable de réduire l'oxygène moléculaire et de produire des espèces radicalaires comme le radical hydroxyle OH^\bullet ou l'anion superoxyde $\text{O}_2^{\bullet-}$ selon la réaction de Fenton. La réaction fait intervenir des métaux sous leur forme réduite et peut être catalysée par la régénération de cette forme en présence d'agents réducteurs (thiols, ascorbate) (figure 28).

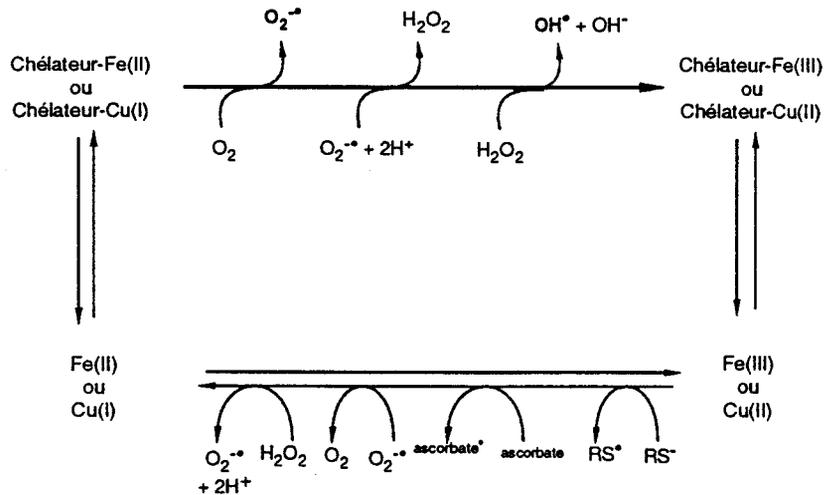


Figure 28: Mécanisme de production d'espèces radicalaires selon Fenton.

Si la production de telles espèces radicalaires a lieu au proche voisinage de l'ADN, on observe la dégradation de celui-ci par coupure aléatoire simple brin ou double brin. La chimie de la rupture de la chaîne ADN par des radicaux libres se décompose comme suit (figure 29) : on pense qu'il y a d'abord abstraction d'un hydrogène en position C'4, puis attaque subséquente par une molécule d'oxygène pour former un peroxyde. L'ouverture du cycle désoxyribose qui s'ensuit conduit à la libération de la base, d'un ester phosphorique de l'acide glycolique et d'une base propénal qui elle-même peut être dégradée pour donner le malonaldéhyde. Ce mécanisme d'action est celui utilisé *in-fine* par plusieurs agents antitumoraux (bléomycine, daunomycine).

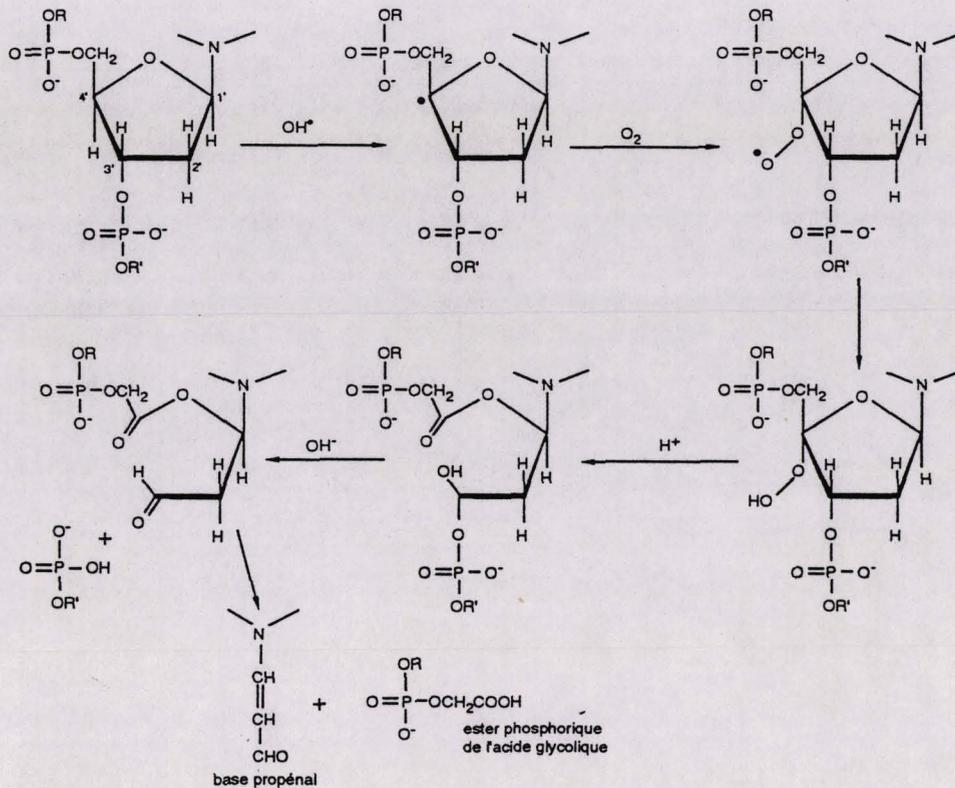


Figure 29. Mécanisme d'action du clivage d'ADN par les espèces radicalaires.

Réalisation d'hybrides : clivage spécifique d'un site d'ADN.

Des agents capables de cibler vers un site spécifique le clivage de l'ADN induit par des complexes chélateur-métal ont déjà été réalisés. Dans ces travaux les auteurs ont accompli l'incorporation d'un chélateur sur une drogue liant spécifiquement l'ADN (Schultz *et al.*, 1982), l'incorporation d'un chélateur en position N-terminale d'un peptide liant l'ADN (Sluka *et al.*, 1987; Mack *et al.*, 1988), l'incorporation d'un chélateur à de multiples positions d'une protéine se liant à l'ADN (Chen et Sigmann, 1987) ou l'incorporation d'un chélateur dans un oligonucléotide homopyrimidinique capable de former une triple hélice (Moser et Dervan, 1987).

Dans une récente publication (Ebright *et al.*, 1990), Ebright et ses collaborateurs décrivent la construction d'un agent de clivage de l'ADN site-spécifique par incorporation d'un complexe 1-10 phénanthroline-cuivre sur une position unique d'un motif hélice-tour-hélice. La fixation de la phénanthroline a été effectuée sur la cystéine 278 d'une sous-unité de la protéine CAP surexprimée chez *Escherichia coli*. Les auteurs ont en effet profité de l'opportunité que cette cystéine 278 était la seule accessible. Une coupure de l'ADN a été observée à proximité du site de fixation habituel de la protéine grâce au respect fortuit d'au moins trois critères : (i) l'introduction du chélateur n'a pas gêné outre mesure les contacts intra-moléculaires qui régissent la structure de la protéine CAP, (ii) le chélateur n'a pas empêché l'interaction CAP-ADN de se produire et (iii) le complexe métal-chélateur était placé suffisamment au proche contact du squelette phosphate-sucre pour que les radicaux libres OH^\bullet (très réactifs) puissent cliver la double hélice. Ces mêmes critères seront pris en compte lors de la construction de nos modèles hybrides.

Nous avons sélectionné comme agent de coupure le tripeptide Gly-Gly-His qui est capable de chélater un atome de cuivre (Sluka *et al.*, 1987) et ainsi de participer à la réaction de Fenton pour produire des espèces radicalaires oxygénées. Ces radicaux ont la propriété de couper l'ADN de la manière décrite au paragraphe 3.

En outre, la réalisation d'hybrides avec les peptides synthétiques DZGR et dzGR était facilement accessible par la simple addition de ces 3 acides aminés en cours de synthèse pour obtenir les peptides GGH-DZGR et GGH-dzGR. Les sites précis de l'interaction de la région CI avec l'ADN étant situés dans la portion C-terminale de la séquence, il nous a semblé judicieux d'ajouter la séquence Gly-Gly-His du côté N-terminal des peptides pour ne pas gêner les interactions probables de DZGR et de dzGR avec la séquence GRE (Glucocorticoid Responsive Element). Pour chaque hybride le but est d'aboutir à une molécule synthétique capable de reconnaître la séquence GRE et de produire des radicaux libres *in-situ* pour couper l'ADN à proximité.

Obtention des peptides GGHdzGR et GGHDZGR.

Les peptides GGHdzGR et GGHDZGR (figure 30) ont été obtenus par synthèse peptidique en phase hétérogène utilisant la chimie des Fmoc. Les réactions de couplage ont été effectuées dans le DMF, à l'exception des résidus 16 à 32 (numérotation de la figure 30) pour lesquels une meilleure solvataion ainsi que des rendements plus élevés ont été obtenus avec la N-méthyle pyrrolidone.

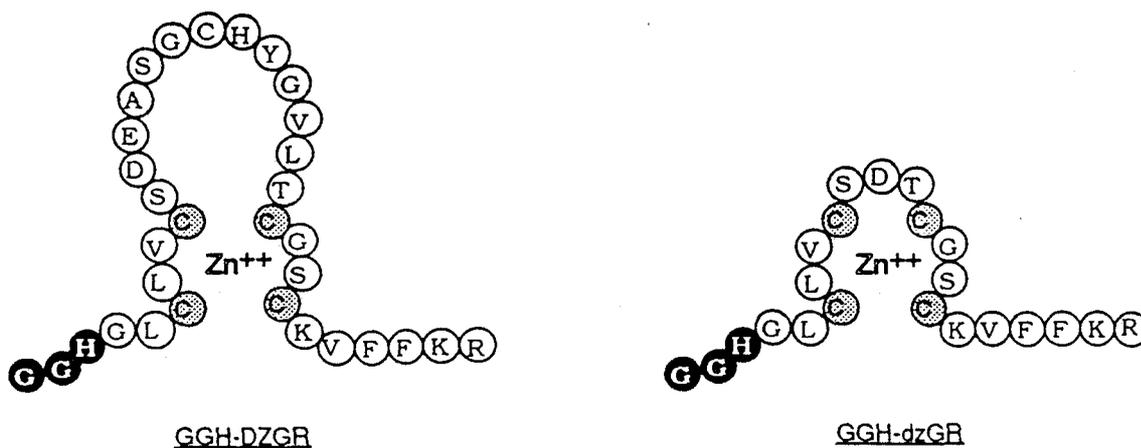


Figure 30. Peptides utilisés: GGHdzGR et GGHDZGR.

Les esters pentafluorophényliques des acides aminés ont été utilisés tout au long de la synthèse sauf pour les résidus Ser et Thr qui ont été utilisés sous la forme d'esters dihydro, 3-hydroxy, 4 oxo-,

1,2,3 benzotriaziniques. Le résidu cystéyl de la boucle de GGHDZGR était protégé par un groupement acétamidométhyle tandis que les quatre autres cystéines étaient protégées par un groupement S-tert butylthio. Cette protection étant enlevée sélectivement par une réduction à la tributylphosphine juste avant les expériences de complexation avec le zinc et/ou le cuivre. Le clivage des peptides de la résine a été réalisé en présence de thioanisole à 5% en tant que scavenger. Après une double étape de purification par gel filtration sur Biogel P4 dans l'acide acétique à 30% suivie d'une chromatographie liquide à haute performance en phase inverse sur colonne de type C18 (gradient d'acétonitrile de 10 à 60%), on obtient des peptides purs dont on vérifie la composition en acides aminés après hydrolyse acide (HCl 5-6N, 110°C, 24h sous vide). Les résultats d'analyse d'acides aminés sont les suivants (nombre théorique entre crochets):

GGHDZGR: Asx 0,98 [1]; Val 2,85 [3]; Thr 0,77 [1]; Ser 2,78 [3]; Glx 1,21 [1]; Gly 5,78 [6]; Cys2 2,40 [2,5]; Ala 0,95 [1]; Leu 3,15 [3]; Tyr 0,98 [1]; Phe 2,26 [2]; His 1,87 [2]; Lys 2,12 [2] et Arg 1,16 [1].

GGHdzGR: Asx 1,01 [1]; Thr 0,84 [1]; Ser 1,70 [2]; Gly 3,94 [4]; Cys2 1,96 [2]; Val 1,79 [2]; Leu 1,85 [2]; Phe 1,99 [2], His 0,98 [1]; Lys 2,16 [2] et Arg 1,07 [1].

Etudes de complexation

Il est connu comme une règle quasi-générale que la présence d'un résidu histidine en troisième position de tri- et de térapeptides provoque une affinité envers les ions Cu^{++} . Cet effet est dû à la présence de l'azote imidazole qui favorise la coordination *via* un système chélateur en anneau dans une géométrie de type plan-carré (figure 31).

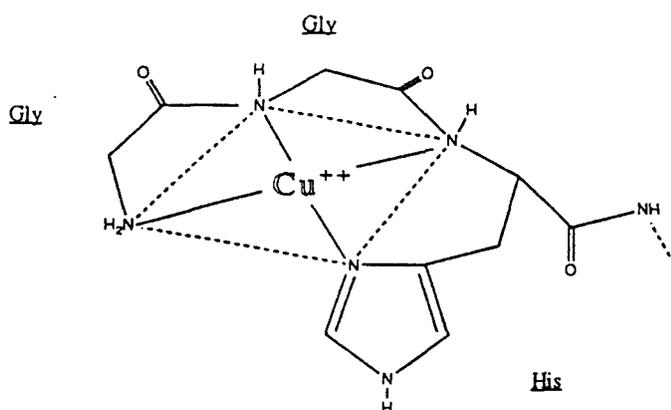


Figure 31. Représentation schématisée de la géométrie de type plan-carré adoptée par le complexe Gly-Gly-His/Cu(II) (d'après Mack et al., 1988).

Nous avons choisi de placer le motif tripeptidique Gly-Gly-His à l'extrémité N-terminale de nos peptides en doigt de zinc de telle façon que le peptide possède un résidu Histidine bien positionné.

En outre, ce motif tripeptidique a été très étudié de par sa présence à l'extrémité N-terminale d'albumines ("Copper-binding peptide").

Pour chaque peptide on enregistre un spectre de résonance paramagnétique électronique en présence de 1 éq. de Cu^{++} ou 1 éq. de Cu^{++} + 1 éq. de Zn^{++} , ces spectres sont reportés à la figure 32.

Si l'on se réfère aux études de Peisach et Blumberg (Peisach et Blumberg, 1983), les paramètres de RPE observés à pH physiologique sont en faveur de complexes impliquant 3 azotes et 1 oxygène ou bien 4 azotes.

Complex	$g_{//}$	g_{\perp}	$A_{//}$ (gauss)
GGH-DZGR.Cu(II)	2.230	2.050	190
GGH-DZGR.Cu(II) Zn(II)	2.246	2.051	180
GGH-dzGR.Cu(II)	2.180	2.044	200
GGH-dzGR.Cu(II) Zn(II)	2.231	2.047	185

Table II. Paramètres de résonance paramagnétique électronique pour les peptides GGHDzGR et GGHDZGR à pH physiologique

Les résultats de la figure 32 dont les paramètres sont dans la table II, reflètent la présence de deux types de complexes selon les conditions expérimentales. Le premier type, caractérisé par une valeur de $g_{//}$ autour de 2,23 et une valeur de $A_{//}$ de 180 gauss correspond au complexe plan-carré du cuivre impliquant quatre atomes d'azote dans la complexation (figure 31). L'autre complexe ($g_{//}=2,18$ et $A_{//}=200$ gauss), seulement trouvé dans le cas du peptide GGHdzGR/Cu(II) correspond à un autre type de complexe Gly-Gly-His/Cu(II) où les azotes amidiques ne prennent pas part à la complexation du cuivre. Si on augmente la concentration en cuivre, ce complexe disparaît et le complexe Gly-Gly-His/Cu(II) plan-carré bien caractérisé devient la seule espèce en solution. Ces résultats confirment aussi que pour chacun des peptides GGHDZGR et GGHdzGR, la complexation du cuivre est bien le fait du motif tripeptidique Gly-Gly-His et que les résidus cystéines ne sont en aucun cas impliqués dans la complexation de ce métal.

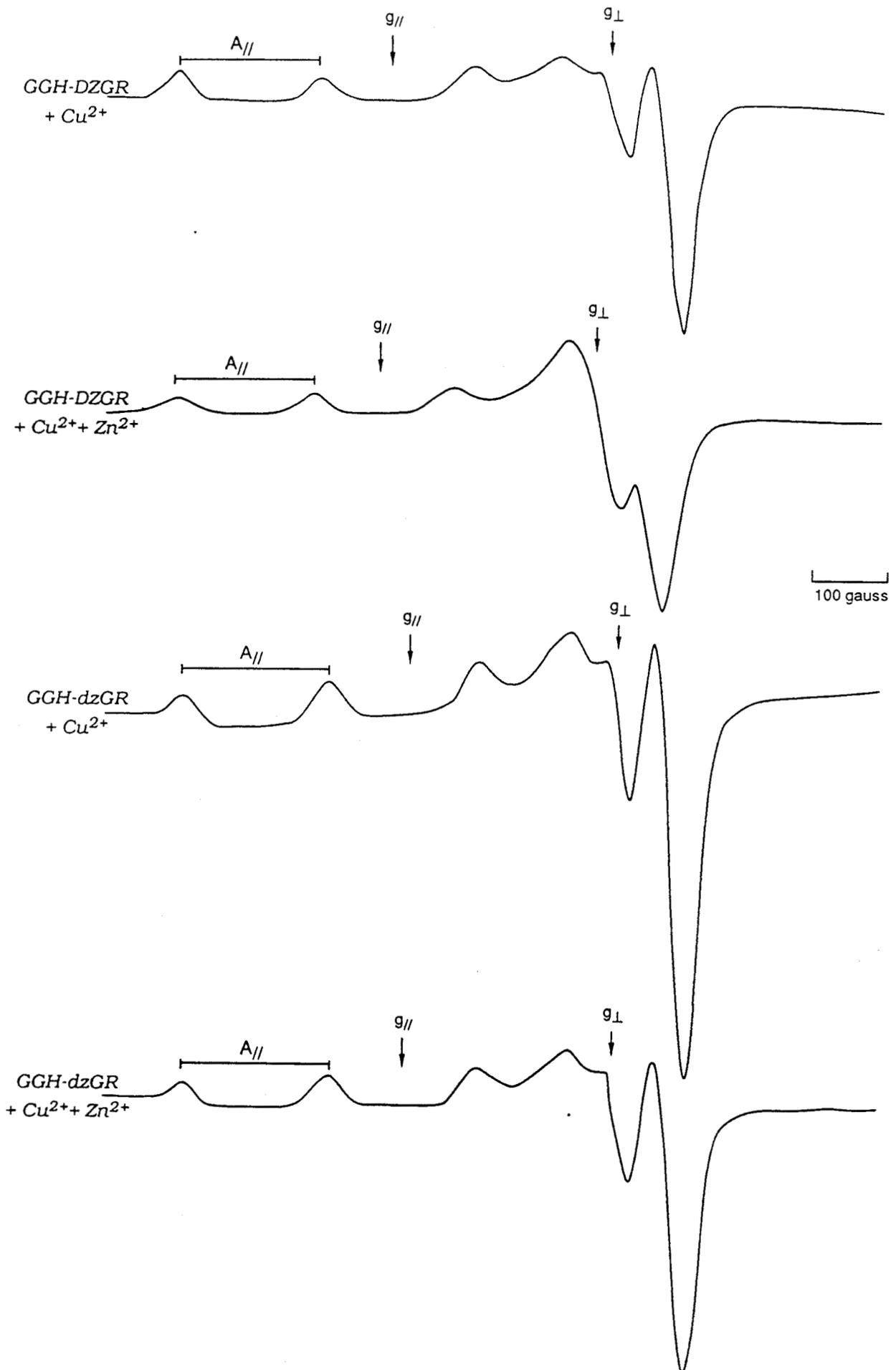


Figure 32. Spectres de résonance paramagnétique électronique des peptides GGHDzGR et GGHDZGR en présence de 1 éq. Cu^{2+} et 1 éq. Zn^{2+} + 1 éq. Cu^{2+}

Production de radicaux libres et coupure d'ADN.

Le tripeptide Gly-Gly-His complexé au cuivre est capable de produire des espèces radicalaires en présence d' H_2O_2 et d'un agent réducteur (Sluka *et al.*, 1987). Il était alors intéressant d'observer si nos peptides étaient dotés de la même propriété. Pour chaque peptide complexé au zinc et au cuivre nous avons pu montrer l'oxydation d' H_2O_2 et la production de radicaux OH° par la technique de spin-trapping au DMPO en résonance paramagnétique électronique. Des expériences de contrôle effectuées sur le DMPO, DMPO+ H_2O_2 et DMPO+ H_2O_2 + Cu^{++} , essentielles si on veut tirer des conclusions significatives basées sur le spin-trapping n'ont pas produit de tels signaux en RPE.

La production de radicaux libres étant démontrée, nous avons étudié la propriété de nos peptides à couper l'ADN. Nous avons fait une expérience tirant parti de la mobilité différentielle du plasmide pBr322 en électrophorèse en gel d'agarose selon qu'il a subi une coupure mono- ou double-brin ou selon qu'il est sous sa forme native super-enroulée. Pour l'induction de la coupure d'ADN par nos peptides GGHzGR et GGHDZGR, 10 μ g de plasmide pBr322 (Boehringer) sont incubés avec chacun des peptides (10^{-5} M) dans un tampon contenant Tris,HCl 50 mM (pH 7,5), $CuClO_4$ $0,8 \cdot 10^{-5}$ M, $ZnCl_2$ $0,8 \cdot 10^{-5}$ M, H_2O_2 0,3% et ascorbate de sodium 10^{-5} M pendant 30 minutes dans l'obscurité et à température ambiante. Les échantillons sont alors soumis à une électrophorèse en gel d'agarose 0,8% contenant du tampon Tris-borate-EDTA 1X et du bromure d'éthidium (5 μ g/ml). La figure 33 montre la photo du gel d'agarose sous éclairage ultra-violet.

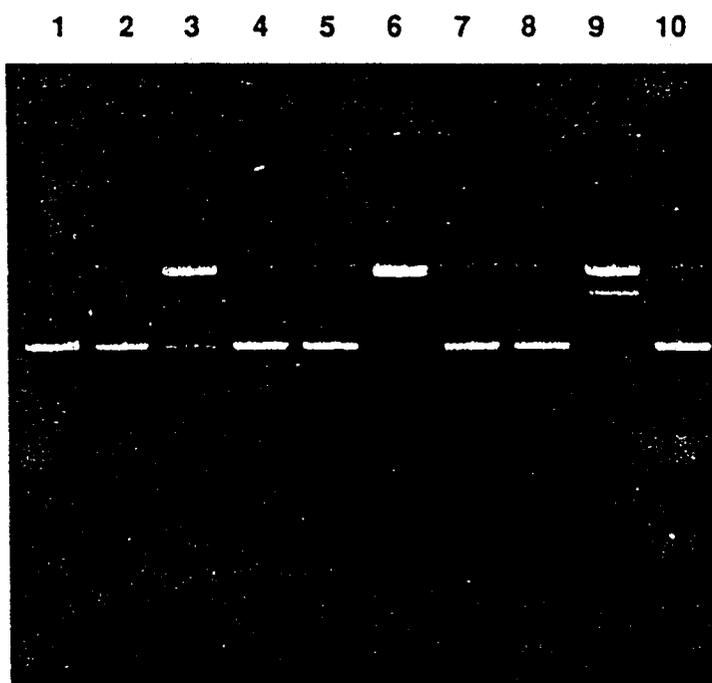


Figure 33. Migration électrophorétique en gel d'agarose 0.8% des produits de coupure du plasmide pBR 322.

Ligne 1, GGHzGR; Ligne 2, GGHzGR+ Cu^{++} ; Ligne 3, GGHzGR+ Zn^{++} + Cu^{++} ; Ligne 4, GGHDZGR;
Ligne 5, GGHDZGR+ Cu^{++} ; Ligne 6, GGHDZGR+ Zn^{++} + Cu^{++} ; Ligne 7, H_2O_2 + Cu^{++} +ascorbate;
Ligne 8, H_2O_2 +ascorbate+ Zn^{++} ; Ligne 9, coupure référence double brin par la bléomycine+ Fe^{++} ;
Ligne 10, plasmide seul.

Les peptides, dans les conditions de production de radicaux libres induisent au plasmide pBR322 une ou plusieurs coupures sur l'un de ses brins alors que le tripeptide Gly-Gly-His dans les

Les peptides, dans les conditions de production de radicaux libres induisent au plasmide pBR322 une ou plusieurs coupures sur l'un de ses brins alors que le tripeptide Gly-Gly-His dans les mêmes conditions et à même concentration n'induit pas la formation de cette forme relâchée (forme II, figure 33). De plus, le peptide GGHDZGR ne donne lieu à une coupure qu'en présence de cuivre et de zinc ce qui suggère (i) que la coupure est rendue possible grâce à une interaction de la partie en doigt de zinc avec l'ADN et (ii) cet apport d'affinité n'a lieu que, si et seulement si, le peptide est en présence de zinc supposant que des changements critiques interviennent dans sa structure en présence de cet ion et le rendent apte à se fixer à l'ADN. Des contrôles ont été effectués en présence de $\text{Cu(II)} + \text{H}_2\text{O}_2$, $\text{Cu(II)} + \text{H}_2\text{O}_2 + \text{ascorbate}$, $\text{Cu(II)} + \text{Zn(II)} + \text{H}_2\text{O}_2 + \text{ascorbate}$, de façon à s'assurer que la coupure à cette concentration est bien due au complexe peptide/métal. Nous avons qualifié ce type d'expérience de coupure de plasmide sous le terme de coupure aléatoire, cependant il existe au sein de la séquence du plasmide pBr322 au moins plusieurs fois la séquence TGTTCT correspondant à un demi-site du palindrome GRE. Nous avons cherché par la suite à démontrer une sélectivité de la coupure de l'ADN par nos peptides sur le site GRE.

Pour cela nous avons entrepris en parallèle à une expérience d'empreintes au radicaux hydroxyles destinée à visualiser le site d'interaction de nos peptides, une expérience de coupure spécifique pour déceler le site préférentiel de coupure par nos peptides hybrides. Ces deux types d'expériences ont été effectués sur le même fragment d'ADN (EcoRI-PvuII du plasmide pBS M13mp18(+)) contenant la séquence GRE insérée trois fois dans le site HindIII) que dans les études précédentes faisant appel à cette technique. La synthèse des peptides, le clonage du GRE, l'obtention du fragment radiomarké ainsi que la technique des empreintes suivent les mêmes protocoles que ceux décrits précédemment dans ce mémoire.

Pour la technique de coupure spécifique, 10 fmoles (300 μCi) de fragment d'ADN triple-insert sont incubés avec le peptide GGHDZGR ou GGHDzGR à 10^{-5} M pendant 80 minutes dans l'obscurité à température ambiante dans un tampon contenant Tris, HCl 50 mM, en présence lorsqu'elle est indiquée d' H_2O_2 0,3%, Cu(II) $0,8 \cdot 10^{-5}$ M, Zn(II) $0,8 \cdot 10^{-5}$ M et d'ascorbate de sodium 10^{-5} M. La réaction est stoppée par l'addition de 3 μl d'EDTA 5 mM, les échantillons sont précipités par addition de trois volumes d'éthanol puis rincés par de l'éthanol 70%. Les échantillons sont ensuite séchés et repris dans 3 μl de formamide 100% contenant du bleu de bromophénol et déposés sur un gel de polyacrylamide (acrylamide 38:bis acrylamide 2) contenant de l'urée 8M et du TBE 1X. L'électrophorèse est effectuée à 2000 V pendant 2h30. Le gel est alors séché et autoradiographié à -70 pendant une nuit.

Les résultats des expériences d'empreintes sont reportés à la figure 34 uniquement dans le cas du peptide GGHDZGR en comparaison avec les effets du récepteur des glucocorticoïdes partiellement purifié, en effet nous n'avons pas pu mettre en évidence une protection contre l'attaque par les radicaux par le peptide GGHDzGR à une concentration significative.

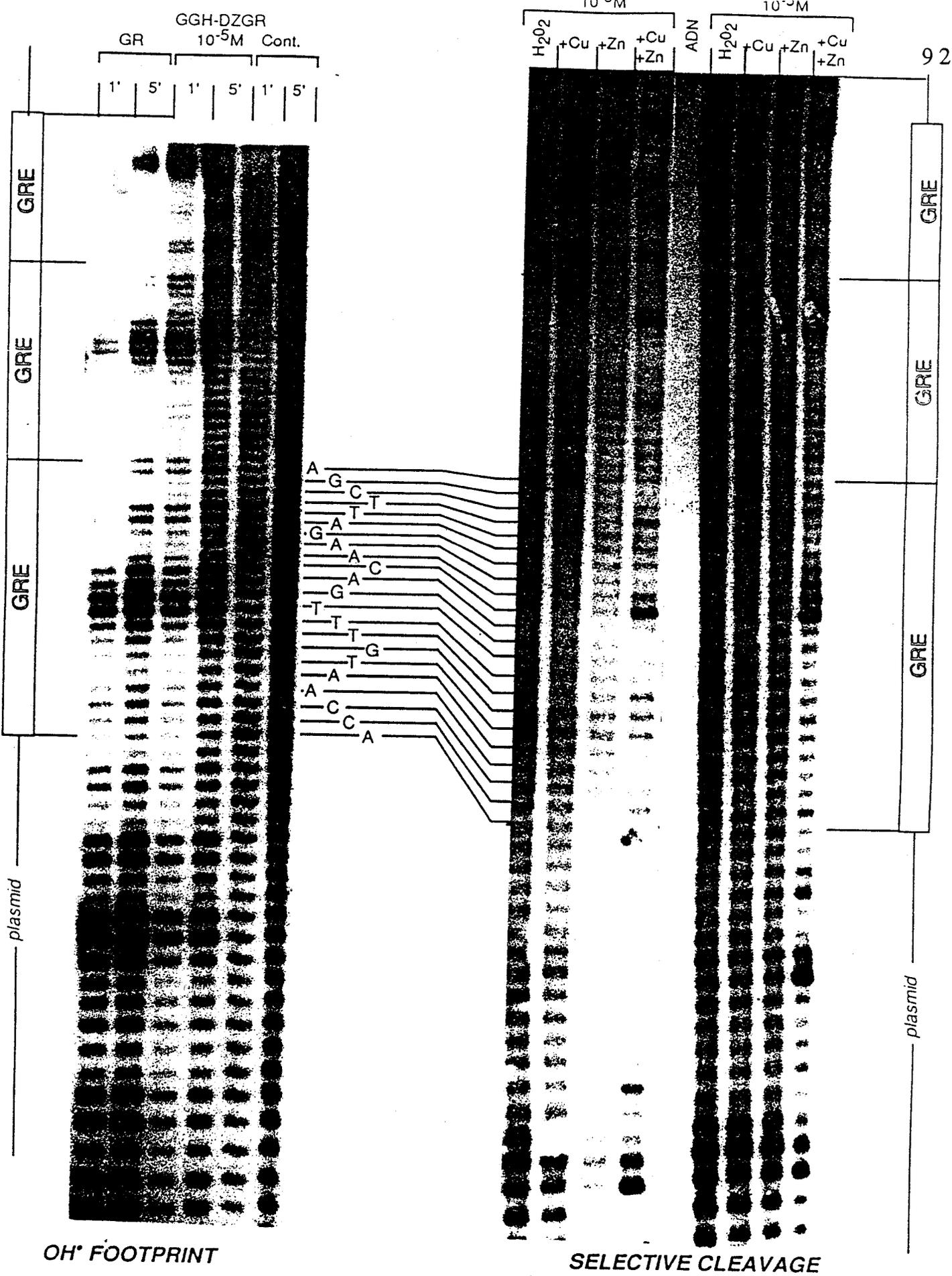


Figure 34. Empreintes des radicaux hydroxyles et coupure sélective par les peptides GGHDZGR et GGHDZGR.

du peptide comme en présence du récepteur sur les deux demi-sites symétriques du palindrome GRE avec une coupure plus prononcée au centre de ce palindrome. L'analyse densitométrique à la figure 35 fait apparaître que le profil de protection (parties positives de la courbe) contre l'attaque par les radicaux hydroxyles par le peptide GGHDZGR suit parfaitement celui en présence de récepteur partiellement purifié. Cependant l'intensité de protection est plus faible en présence du peptide. Une fois encore, on suggère l'interaction de deux molécules de peptides GGHDZGR avec les deux demi-sites de chaque GRE inséré, de la même manière que le fait le dimère de récepteur aux glucocorticoïdes.

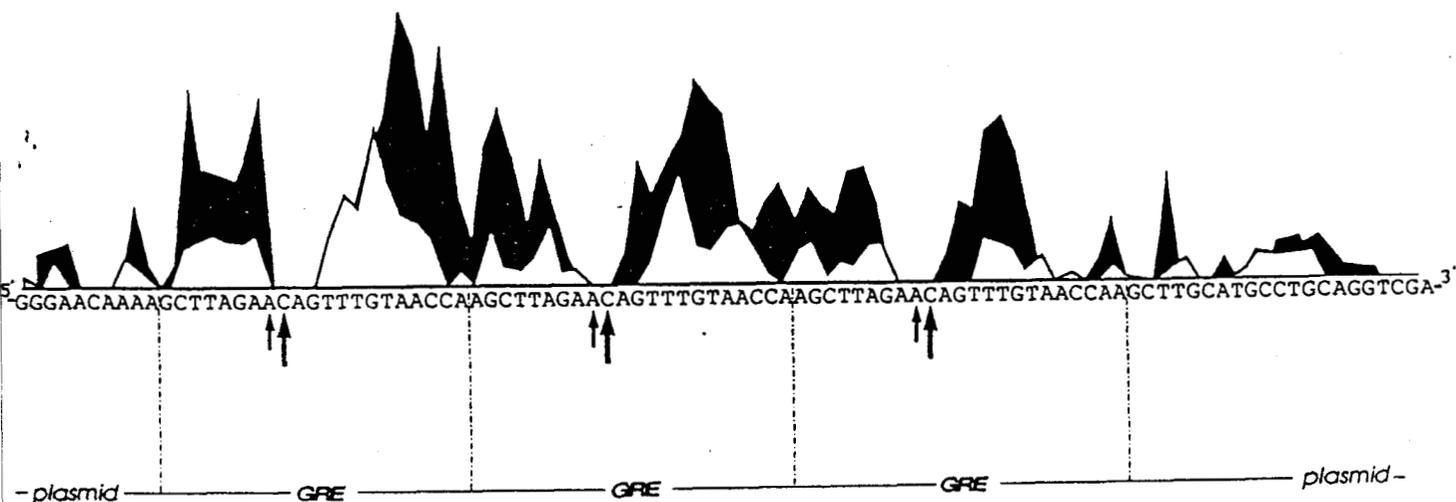


Figure 35. Analyse densitométrique des autoradiogrammes issus de la techniques des empreintes et de la coupure sélective par les peptides GGHDzGR et GGHDZGR
 Les aires blanches représentent l'extension de protection par les peptides et les aires noires la protection observée pour le récepteur.
 Les sites de coupure sélective sont représentés par des flèches noires dont la longueur est proportionnelle à l'intensité de coupure

En parallèle à ces études d'empreintes, on a fait apparaître sur la figure 35 les résultats de la coupure sélective du même fragment d'ADN en présence des peptides GGHDzGR et GGHDZGR. On observe que les peptides, en présence d' H_2O_2 et d'un agent réducteur (ascorbate) mais en l'absence de métal (lignes 1 et 6) ou en présence soit de Cu(II) (lignes 2 et 7), soit de Zn(II) (lignes 3 et 8) donnent lieu à une faible coupure répartie uniformément sur toute la longueur du fragment d'ADN. Tandis que chaque peptide, en présence de Cu(II) et de Zn(II) (lignes 4 et 9) montre des sites de coupure préférentielle bien localisés. Ces sites de coupure sélective communs aux peptides GGHDzGR et GGHDZGR sont représentés par des flèches pleines sur le profil d'analyse densitométrique des empreintes. Cette coupure est localisée en un endroit précis de chaque palindrome GRE inséré : sur le résidu C4 et à un niveau moindre sur le résidu A3.

empreintes. Cette coupure est localisée en un endroit précis de chaque palindrôme GRE inséré : sur le résidu C4 et à un niveau moindre sur le résidu A3.

Afin de comprendre cette coïncidence entre protection et coupure nous avons repris les résultats des études cristallographiques de Luisi *et al.* Selon ces données cristallographiques, il est possible de schématiser l'interaction entre le peptide GGHDZGR et un demi-site de l'élément palindromique GRE de la manière représentée à la figure 36.

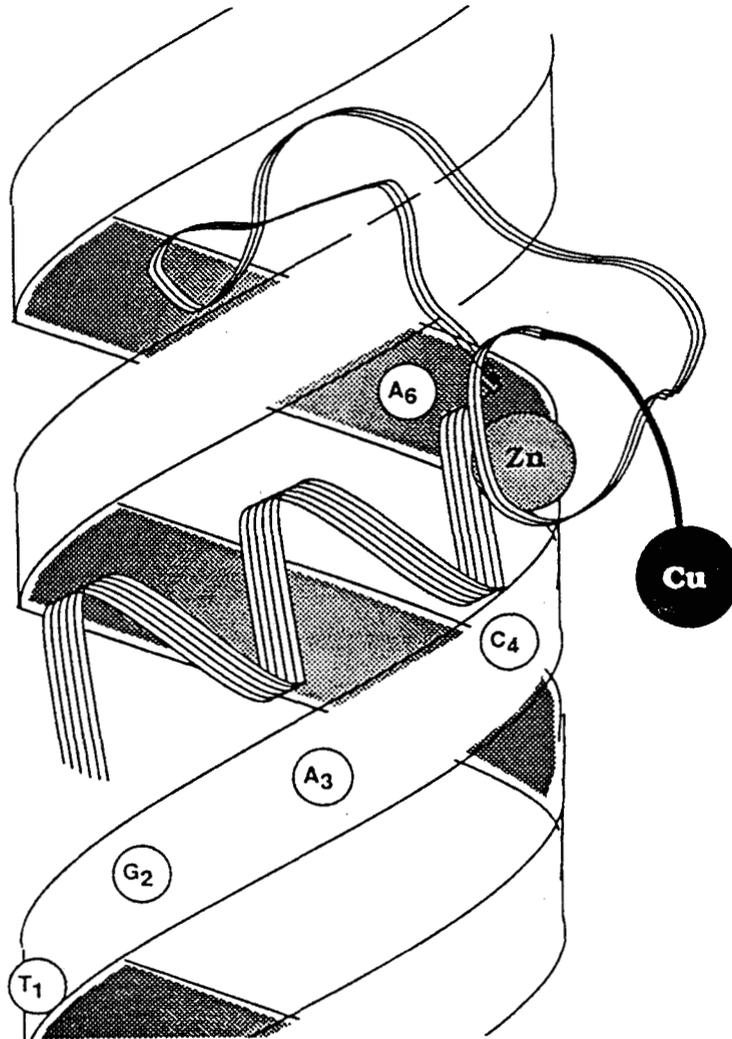


Figure 36. Représentation schématique du peptide GGHDZGR en interaction avec un demi-site de GRE (d'après Luisi *et al.*, 1991).

L'hélice α adjacente à la boucle du doigt de zinc CI et induite par le complexe tétraédrique du métal repose dans le grand sillon de la double hélice permettant aux groupements qu'elle supporte d'entrer en interaction avec certains plateaux de paires de bases (ex: interaction van der Waals entre Valine 465 et Thymine 5). Nos résultats d'empreintes sont en parfait accord avec ce modèle et nous proposons que la partie doigt de zinc de nos peptides ainsi que la version simplifiée vienne se loger dans le grand sillon et reconnaisse l'ADN de la même façon que le fait la région correspondante du récepteur. Selon ce modèle d'interaction, la présence du motif Gly-Gly-His en position N-terminale n'interfère pas avec l'interaction doigt de zinc/grand sillon et la délétion de la boucle n'a pas de

phosphate-sucre exactement situé au niveau du lien désoxyribose entre les bases C4 et A5 expliquant ainsi la coupure observée à cet endroit précis du demi-palindrome GRE par les deux peptides.

La plus faible intensité de coupure que l'on observe en présence du peptide GGHdzGR est à mettre en parallèle avec son incapacité à donner une empreinte très claire. En fait, nous avons pu observer que ce peptide ne possède qu'une faible affinité pour le zinc, i.e. en présence d'un équivalent de Zn(II), seule une faible proportion de peptide forme le complexe tétraédrique et ainsi seule une très faible proportion de ce peptide possède la bonne structure (hélice α) capable de reconnaître l'ADN. Au cours de la brève exposition aux radicaux hydroxyles (1 et 5 minutes) dans la technique des empreintes, trop peu de complexe GGHdzGR/GRE sont formés pour donner lieu à une protection détectable à cette concentration. Cependant, au cours de l'expérience relativement longue de coupure sélective, des complexes GGHdzGR/GRE ont le temps de se former et la coupure de se produire. Même si l'interaction est faible et ainsi la vitesse de dissociation du complexe est plus forte que la vitesse d'association, le phénomène irréversible de coupure oxydative donne lieu à une quantité suffisante de fragment correspondant pour être détecté par la sensibilité de la méthode autoradiographique.

Il reste encore deux questions soulevées par les résultats de la figure 35:

1. Les empreintes suggèrent que l'interaction se fait entre deux molécules de peptides pour un palindrome GRE, or on n'observe qu'une seule coupure par palindrome, sur le demi-site "inférieur". Ce résultat est à rapprocher de celui de la coupure aléatoire sur plasmide pBR322 où l'on avait conclu à une coupure monobrin de l'ADN par les peptides GGHdzGR et GGHDZGR. Si, de la même façon qu'on l'a fait à la figure 36, on modélise l'interaction d'une autre molécule de peptide sur l'autre demi-site du GRE on s'aperçoit que la coupure, qui est monobrin, s'effectuera en théorie sur l'autre brin et qu'il est donc logique qu'elle n'apparût pas dans ce résultat ne concernant qu'un seul des deux brins.

2. Dans le cas de la coupure par GGHDZGR (ligne 9) on observe un site de coupure prononcé qui se trouve en dehors des GRE insérés (bases 284 et 285). Cependant on observe que cette coupure a lieu sur des bases C et A de la même façon que la coupure avait lieu sur les bases C4 et A5 dans les GRE insérés. En outre on observe à distance adéquate de ce site de coupure, la présence de deux plateaux A-T qui sont des acteurs majeurs de la reconnaissance par interaction valine/thymine selon les données cristallographiques.

C'est pourquoi, ces résultats nous poussent à suggérer que nous sommes parvenus à élaborer des molécules peptidiques relativement simples possédant une activité de coupure d'ADN assez sélective à des endroits bien définis reconnus: Les éléments de réponse hormonale aux glucocorticoïdes.

Nos modèles peptidiques hybrides sont donc capables de reconnaître et de se fixer à une séquence donnée de 6 paires de bases de l'ADN (un demi-site de GRE) et de couper un brin de la double hélice sur ce site de manière très précise. Ces endonucléases artificielles peuvent être qualifiées de hautement spécifiques comparativement au modèle naturel des enzymes de restriction. Les représentants les plus performants de cette classe d'enzymes (les enzymes à coupure rare ou "rare cutters") exercent la double fonction reconnaissance / coupure sur des séquences de 8 paires de bases.

représentants les plus performants de cette classe d'enzymes (les enzymes à coupure rare ou "rare cutters") exercent la double fonction reconnaissance / coupure sur des séquences de 8 paires de bases. On peut considérer que nos modèles GGH-DZGR et GGH-dzGR coupent les deux brins d'un duplex d'ADN contenant deux demi-sites de GRE disposés en palindrome (*i.e.* symétrie rotationnelle). Deux peptides n'étant pas liés par processus de dimérisation, l'espacement entre les deux demi-sites n'est pas défini. Le résultat de la coupure d'une telle séquence donne alors 2 extrémités cohésives dont la longueur varie avec l'espace entre les deux demi-sites (figure 37).

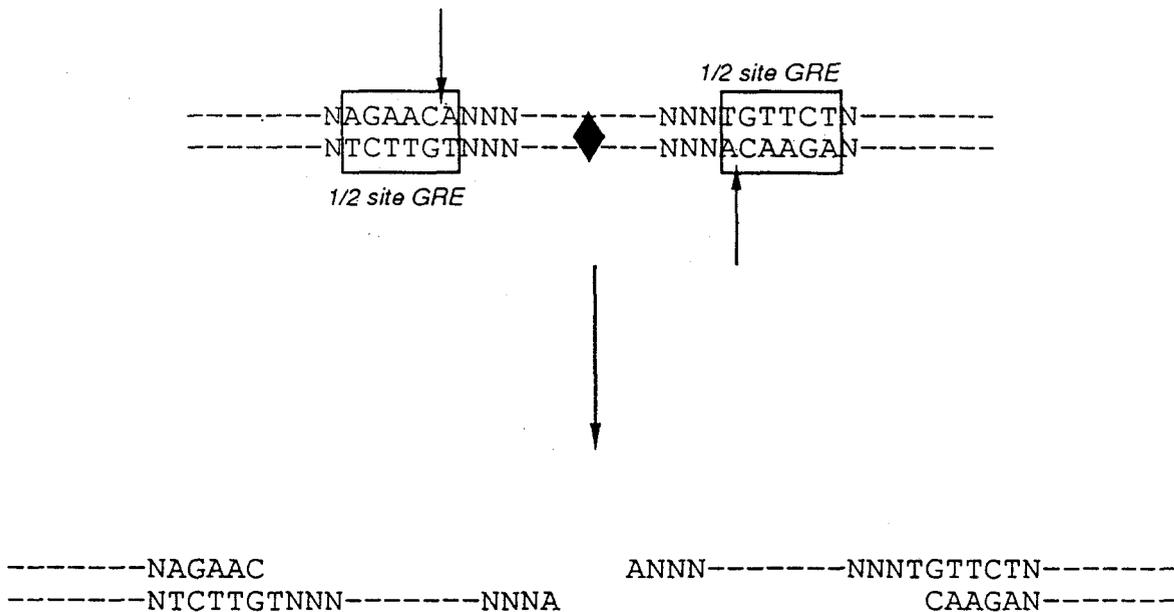


Figure 37 Représentation schématique de la coupure par GGH-DZGR ou GGH-dzGR d'un ADN contenant 2 demi-palindromes (encadrés) de GRE. (N) correspond à n'importe quelle base et (◆) indique le centre de symétrie rotationnelle.

On peut raisonnablement estimer qu'une telle coupure devrait se produire efficacement sur une séquence totale allant jusqu'à 25-30 paires de bases. Au-delà, la coupure serait effective mais l'appariement entre les bases des extrémités cohésives serait trop fort pour la dissociation des deux brins. Des vérifications très prudentes restent à mettre en œuvre mais on peut d'ores et déjà estimer que de telles endonucléases artificielles s'avéreraient d'une utilité certaine pour la biologie moléculaire, par exemple. De fait, si l'explosion de l'ingénierie génétique des quinze dernières années est essentiellement due à la découverte des enzymes de restriction, leur manque relatif de spécificité est vite apparu limitant. Par exemple, une endonucléase reconnaissant un site de 8 paires de bases coupe statistiquement 72 fois le génôme d'*Escherichia coli* alors que le génôme humain est disloqué en

50 000 pièces. Suivant ce raisonnement, une enzyme ayant une spécificité de reconnaissance de 20 paires de bases aurait statistiquement un site de coupure toutes les 10^{12} paires de bases soit un site tous les 1000 génômes humains.

S'inspirant de nos modèles, la conception d'autres endonucléases artificielles de performance similaire semble abordable. Les chercheurs pourraient alors disposer d'une panoplie d'enzymes de restriction synthétiques de très haute spécificité rendant ainsi plus aisés des projets de cartographie et de séquençage de grands génômes tels que ceux d'organismes supérieurs voire celui de l'être humain.

CONCLUSIONS

Les thèmes abordés et les résultats acquis au cours de ce mémoire amènent plusieurs réflexions. D'un point de vue très général, il semble que (était-il nécessaire de le démontrer ?) la reconnaissance moléculaire soit prédominante dans la totalité des processus biologiques. Nous avons pu nous apercevoir qu'il est absolument indispensable de tenir compte de la structure tridimensionnelle des macromolécules biologiques pour rendre compte de leur fonction. Les techniques de biochimie classiques et de biologie moléculaire, malgré leur puissance d'investigation aboutissent souvent à des informations relatives à la structure primaire des macromolécules : les modifications chimiques d'un acide aminé ou les techniques de délétion/mutation permettent de déterminer en termes de séquence le site actif d'une enzyme, un domaine de liaison à l'ADN ou une séquence opérateur d'un gène. Ces approches sont tout à fait nécessaires pour aborder la compréhension d'un mécanisme biologique; cependant elles ne sont pas suffisantes pour appréhender un éclaircissement total des rouages impliqués. Sous le thème récurrent de la "Science Une et Indivisible", il doit être fait appel à d'autres outils. L'approche de deux étapes du mécanisme d'action des récepteurs stéroïdiens effectuée dans les travaux de recherche décrits dans ce mémoire est celle d'une "mimétique" chimique. Bénéficiant de l'apport des techniques biochimiques et biologiques, des domaines fonctionnels sont identifiés responsables d'étapes du mécanisme d'action des récepteurs stéroïdiens. La chimie de synthèse permet alors de reproduire ces régions *ab-nihilo* et de démontrer leur fonctionnement intrinsèque hors du contexte de la protéine, de la cellule, de l'organisme. L'approche chimique nous a aussi permis de démontrer que le caractère nécessaire d'une séquence de ces domaines devient suffisant à condition que l'on respecte son aspect structural tridimensionnel. Par exemple, une séquence de doigt de zinc ne reconnaît l'ADN que si le complexe de zinc est formé induisant par sa géométrie la bonne structure ou bien si la région précise d'interaction possède une structure en hélice α .

De manière plus modeste on doit considérer que nous n'avons pas l'exclusivité de cette approche, car elle est largement utilisée dans d'autres travaux s'intéressant à la reconnaissance moléculaire des molécules biologiques : modélisation de sites actifs d'enzymes, élaboration d'épitopes synthétiques ou conception de ligands ou de substrats synthétiques.

Bien qu'il faille rester souvent très prudent lorsqu'on aborde le sujet des applications potentielles d'une recherche, il nous faut rappeler que ces travaux ont été effectués dans le cadre d'une unité de l'Institut National de la Santé et de la Recherche Médicale. C'est pourquoi les perspectives ambitieuses que l'on peut imaginer à ces études sont surtout du domaine de la thérapeutique.

Les motivations pour l'étude sur les séquences signal d'adressage nucléaire étaient avouées dès le chapitre des hypothèses de travail : trouver une molécule relativement simple capable de transporter très efficacement un xénobiotique vers le noyau de la cellule. Une application toute désignée serait de rendre caryophiles des substances antitumorales conçues pour agir sur l'ADN et par conséquent dans le noyau pour les cellules eucaryotes.

Les hypothèses d'applications que l'on peut formuler pour un motif simple (tel que le doigt de zinc CI des récepteurs stéroïdiens) capable de reconnaître l'ADN sont multiples. Prenons le cas isolé du récepteur des glucocorticoïdes dont il est maintenant possible d'imaginer qu'une molécule simple

puisse prendre la place par compétition sur l'élément de réponse hormonal. Nous avons envisagé le cas où cette fixation induirait à elle seule une régulation de la transcription (par courbure de l'ADN et rapprochement d'éléments actifs dans la transcription par exemple); des résultats préliminaires non décrits dans ce mémoire nous permettent d'affirmer que cette hypothèse est peu probable (au moins pour le motif de reconnaissance seul). Ce qui est plus plausible, c'est une compétition pour la fixation entre le récepteur et le motif de reconnaissance syn.hétique et ainsi une levée de la régulation transcriptionnelle induite par le récepteur. Autrement dit, une nouvelle voie de conception d'antagonistes pharmacologiques de l'effet de glucocorticoïdes. Si, comme nous sommes en train de le vérifier au laboratoire, le concept de motif simple de reconnaissance d'ADN peut être généralisé à toute la superfamille des récepteurs stéroïdiens, on aurait alors accès à une palette d'agents pharmacologiques hyper-sélectifs ayant des activités aussi diverses que celles des hormones stéroïdiennes (inflammation, hypertension artérielle, cancers hormonodépendants, contraception), des hormones thyroïdiennes (croissance, développement), de la vitamine D3 (rachitisme, croissance), de l'acide rétinoïque (différenciation, morphogénèse, hépatocarcinome) ou d'autres ligands de fonction non encore soupçonnée.

L'utilisation du motif de reconnaissance de l'ADN dans l'élaboration d'endonucléases artificielles à très haute spécificité est évidente en ce qui concerne l'outil de biologie moléculaire : couper le génôme humain en des sites très rares rendrait sa cartographie et son séquençage plus aisé et plus rapide, avec toutes les implications que cela comporte (localisation d'oncogènes, détermination des maladies génétiques, thérapie génique, etc.). De plus, être capable de reconnaître très spécifiquement une séquence d'ADN et de couper la double hélice à ce site précis pourrait donner naissance à une nouvelle voie de conception de molécules à activité d'excision hyper-sélective sur une région du génôme. Pour le moment de telles molécules ayant une action sur le patrimoine génétique d'un organisme n'ont qu'une utilisation restreinte aux médicaments antitumoraux. Dans le cas (tout à fait envisageable) où la spécificité de reconnaissance s'effectuerait sur une séquence mutée ou modifiée dont la présence est responsable de l'expression de produits d'oncogènes, l'excision ou une dégradation de cette région très précise pourrait aboutir à de nouvelles voies de chimiothérapie antitumorale hyper sélectives.

REFERENCES
BIBLIOGRAPHIQUES



- Adam, S.A., Lobl, T.J., Mitchell, M.A. & Gerace, L. (1989) "Identification of specific binding proteins for a nuclear location sequence." *Nature* **337**, 276-279.
- Adler, K., Beyreuther, K., Fanning, E., Geister, N., Gronnenborn, B., Klemm, A., Muller-Hill, B., Pfahl, M. & Schmitz, A. (1972) "How Lac-repressor binds to DNA." *Nature* **237**, 322-327.
- Anderson, W.F., Ohlendorf, D.H., Takeda, Y. & Matthews, B.W. (1981) "Structure of the Cro-repressor from bacteriophage λ and its interaction with DNA." *Nature* **290**, 754-758.
- Arriza, J., Weinberger, C., Cerelli, G., Glaser, T.M., Handelin, B.L., Housman, D.E. & Evans, R.M. (1987) "Cloning of human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor." *Science* **237**, 268-270.
- Ashbrenner, M. (1980) "Chromosomal action of ecdysone." *Nature* **225**, 435-437.
- Bailly, C. & Hénichart, J.P. (1991) "DNA recognition by intercalator-minor groove binder hybrid molecules." *Bioconj. Chem.* **2**, 379-393.
- Bailly, C., Pommery, N. & Hénichart, J.P. (1988) "DNA-synthesis and tumor growth inhibition by AGGA, a bleomycin-amsacrine hybrid derivative." *Cancer Lett.* **38**, 321-328.
- Bailly, C., Pommery, N., Houssin, R. & Hénichart J.P. (1989a) "Design, synthesis, DNA-binding and biological activity of a series of DNA minor groove binding intercalating drugs." *J. Pharm. Sci.* **78**, 910-917.
- Bailly, C., Catteau, J.P., Hénichart, J.P., Reszka, K., Shea, R.G., Krowicki, K. & Lown, J.W. (1989b) "Subcellular distribution of a nitroxide spin-labeled netropsin in living KB cells." *Biochem. Pharmacol.* **38**, 1625-1630.
- Banerji, J., Olson, L. & Schaffner, W.A. (1983) "A lymphocyte specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes." *Cell* **33**, 729-737.
- Baum, J.A., Geever, R. & Giles, N.H. (1987) "Expression of qa-1F activator protein: identification of upstream binding sites in the qa gene cluster and localization of the DNA-binding domain." *Mol. Cell. Biol.* **7**, 1256-1266.
- Baxter, J.D. & Tyrrell, J.B. in *Endocrinology and Metabolism*, P. Felig, D. Baxter, A.E. Broadus, L.A. Frohman, Eds. (McGraw-Hill, New-York), pp. 385-510.

- Benbrook, D. & Pfahl, M. (1987) "A novel thyroid hormone receptor encoded by a cDNA clone from a human testis library." *Science* **238**, 788-792.
- Benoist, C. & Chambon, P. (1981) "In-vivo sequence requirements of the SV40 early promoter region." *Nature* **290**, 304-312.
- Bentley, P.J. (1986) *Comparative vertebrate endocrinology* (Cambridge Univ. Press, Cambridge) 65-76.
- Bernier, J.L., Lohez, M., Helbecque, N. & Henichart, J.P. (1989) "Effect of the combination of m-AMSA and doxorubicin on their redox properties and on DNA cleavage." *Chem. Biol. Interact.* **70**, 103-115.
- Breeuwer, M. & Goldfarb, D.S. (1990) "Facilitated nuclear transport of histone H1 and other small nucleophilic proteins." *Cell* **60**, 999-1008.
- Bogenhagen, D.F., Sakonju, S. & Brown, D.D. (1980) "A control region in the center of the 5S RNA gene directs specific initiation of transcription. II. The 3' border region." *Cell* **19**, 27-35.
- Broackus, H. in *Endocrinology and Metabolism*, P. Felig, D. Baxter, A.E. Broadus, L.A. Frohman, Eds. (McGraw-Hill, New-York), pp. 963-1080.
- Brown, R.S., Sander, C. & Argos, P. (1985) "The primary structure of transcription factor TFIID has 12 consecutive repeats." *FEBS Lett.* **186**, 271-274.
- Carlstedt-Duke, J., Okret, S., Wrangé, O. & Gustafsson, J.-A. (1982) "Immunochemical analysis of the glucocorticoid receptor: identification of a third domain separate from the steroid- and DNA- binding domains." *Proc. Natl. Acad. Sci. USA* **79**, 4260-4267.
- Chalepakis, G., Postma, J.P.M. & Beato, M. (1988) "A model for hormone receptor binding to the mouse mammary tumor virus regulatory element based on hydroxyl radical foot-printing". *Nucleic Acids Res.* **16**, 10237-10247.
- Chambon, P., Dierich, A., Gaub, M.P., Jakowlev, S., Jongstra, J., Krust, A., LePennec, J.P., Oudet, P. & Reudelhuber, P. (1984) "Promoter elements of genes coding for proteins and modulation of transcription by estrogens and progesterone." *Rec. Prog. Hormone Res.* **40**, 1-78.

- Chandler, V.L., Maler, B.A. & Yamamoto, K.R. (1983) "DNA sequences bound specifically by glucocorticoid receptor *in-vitro* render a heterologous promoter hormone responsive *in-vivo*." *Cell* **33**, 489-497.
- Chen, C.H. & Sigman, D. (1987) "Chemical conversion of a DNA-binding protein into a site-specific nuclease." *Science* **237**, 1197-1201.
- Compton, J., Schrader, W.T. & O'Malley, B.W. (1983) "DNA sequence preferences of the progesterone receptor." *Proc. Natl. Acad. Sci. USA* **80**, 161-168.
- Debuire, B., Henry, C., Benaissa, M., Biserte, G., Claverie, J.M., Saule, S., Martin, P. & Stehelin, D. (1984) "Sequencing the erb-A gene of avian erythroblastosis virus reveals a new type of oncogene." *Science* **224**, 1456-1459.
- Dejean, A., Bougueleret, L., Grzesnick, K.H. & Tiollais, P. (1986) "Hepatitis B virus DNA integration in a sequence homologous to v-erbA and steroid receptors genes in an hepatocellular carcinoma." *Nature* **322**, 70-72.
- Dellwege, H.G., Hotz, A., Mugele, K. & Gehring, U. (1982) "Active domains in wild-type and mutant glucocorticoid receptors." *EMBO J.* **1**, 285-289.
- DeRobertis, E.M., Longthorne, R.F. & Gurden, J.B. (1978) "Intracellular migration of proteins in *Xenopus* oocytes." *Nature* **272**, 254-256.
- DeThé, H., Marchio, A., Tiollais, P. & Dejean, A. (1987) "A novel steroid-thyroid hormone receptor related gene inappropriately expressed in human hepatocellular carcinoma." *Nature* **330**, 667-670.
- DeThé, H., del Mar Vivanco-Ruiz, M., Tiollais, P., Stunnenberg, H. & Dejean, A. (1990) "Identification of a retinoic acid responsive element in the retinoic acid receptor β gene." *Nature* **343**, 177-180.
- Diakun, G.P., Fairall, L. & Klug, A. (1986) "EXAFS study of the zinc-binding sites in the protein transcription factor IIIA." *Nature* **324**, 698-699.
- Dingwall, C., Sharnick, L. & Laskey, R.A. (1982) "A polypeptide domain that specifies migration of nucleoplasmin into the nucleus." *Cell* **30**, 449-458.
- Druege, D., Klein-Hitpaß, L., Green, S., Stack, G., Chambon, P. & Ryffel, G.U. (1986) "Introduction of estrogen responsiveness in human cell lines." *Nucl. Acids Res.* **14**, 9329-9334.

- Ebright, R.H., Ebright, Y.W., Pendergrast, P.S. & Gunasekera, A. (1990) "Conversion of a helix-turn-helix motif sequence-specific DNA binding protein into a site-specific DNA cleaving agent." *Proc. Natl. Acad. Sci. USA* **87**, 2882-2886.
- Engelke, D.R., Ng, S.Y., Shastry, B.S. & Roeder, R.G. (1980) "Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes." *Cell* **19**, 717-728.
- Evans, R.M., Binsberg, S.C. & Rosenfeld, M.G. (1982) "Glucocorticoid and thyroid hormones transcriptionally regulate growth hormone gene expression." *Proc. Natl. Acad. Sci. USA* **79**, 7659-7666.
- Evans, R.M. (1988) "The steroid and thyroid hormone receptor superfamily." *Science* **240**, 889-895.
- Fairall, L., Rhodes, D. & Klug, A. (1986) "Mapping of the sites of protection on a 5S RNA gene by *Xenopus* transcription factor IIIA. A model for the interaction." *J. Mol. Biol.* **192**, 577-591.
- Freedman, L.P., Luisi, B.F., Richard-Korzun, Z, Basavappa, R., Sigler, P.B. & Yamamoto, K.R. (1988) "The function & structure of the metal-coordination sites within the glucocorticoid receptor DNA-binding domain." *Nature* **334**, 543-546.
- Gaunt, R. in *Handbook of Physiology : Endocrinology*. R.Greep, E. Astwood, Eds. (American Physiological Society, Washington, D.C., 1975) vol. 6, 1-12.
- Giguère, V., Hollenberg, S.M., Rosenfeld, M.G. & Evans, R.M. (1986) "Functional domains of the glucocorticoid receptor." *Cell* **46**, 645-652.
- Giguère, V., Ong, E.S., Segui, P. & Evans, R.M. (1987) "Identification of a receptor for the morphogen retinoic acid." *Nature* **330**, 624-631.
- Giguère, V., Yang, N., Segui, P. & Evans, R.M. (1988) "Identification of a new class of steroid hormone receptor." *Nature* **331**, 91-94.
- Glass, C.K., Franco, R., Weinberger, C., Albert, V.R., Evans, R.M. & Rosenfeld, M.G. (1987) "A c-erb-A binding site in rat growth hormone gene mediates trans-activation by thyroid hormone." *Nature* **329**, 738-740.

- Godowski, P.J., Rusconi, S., Miesfeld, R. & Yamamoto, K.R. (1987) "Glucocorticoid receptor mutants that are constitutive activators of transcriptional enhancement." *Nature* **325**, 365-367.
- Goldfarb, D.S., Garipey, J., Schoolnik, G. & Kornberg, R.D. (1986) "Synthetic peptides act as nuclear localization signals." *Nature* **322**, 641-644.
- Govindan, M.V., Spiess, E. & Majors, J. (1982) "Purified glucocorticoid receptor-hormone complex from rat liver cytosol binds specifically to mouse mammary tumor virus long terminal repeats in-vitro." *Proc. Natl. Acad. Sci. USA* **79**, 5157-5165.
- Graves, B.J., Johnson, P.F. & McKnight, S.L. (1986) "Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene." *Cell* **44**, 565-576.
- Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.M., Argos, P. & Chambon, P. (1986) "Human estrogen receptor cDNA: sequence, expression and homology to v-erb-A." *Nature* **320**, 134-137.
- Greene, G.L., Gilna, P., Waterfield, M., Baker, A., Hort, Y. & Shine, J. (1986) "Sequence and expression of human estrogen receptor complementary DNA." *Science* **231**, 1150-1154.
- Gronemeyer, H., Turcotte, B., Quirin-Stricker, C., Bocquel, M.T., Meyer, M.E., Krozowski, Z., Jeltsch, J.M., Lerouge, T., Garnier, J.M. & Chambon, P. (1987) "The chicken progesterone receptor: sequence, expression and functional analysis." *EMBO J.* **6**, 3985-3994.
- Hager, J. & Palmiter, R.D. (1981) "Transcriptional regulation of mouse liver metallothionein-1 gene by glucocorticoids." *Nature* **291**, 340-342.
- Hall, M.N., Hereford, L. & Herskowitz, I. (1984) "Targeting of *E. coli* β -galactosidase to the nucleus in yeast." *Cell* **36**, 1057-1065.
- Hanas, J.S., Hazuda, D.J., Bogenhagen, B.F., Wu, F.Y.H. & Wu, C.W. (1983a) "Xenopus transcription factor A requires zinc for binding to the 5S RNA gene." *J. Biol. Chem.* **258**, 14120-14125.
- Hanas, J.S., Bogenhagen, B.F. & Wu, C.W. (1983b) "Cooperative model for the binding of Xenopus transcription factor A to the 5S RNA gene." *Proc. Natl. Acad. Sci. USA* **80**, 2142-2145.

- Härd, T., Kellenbach, E., Boelens, R., Maler, B.A., Dahlman, K., Freedman, L.P., Carlstedt-Duke, J., Yamamoto, K.R., Gustafsson, J.-A. & Kaptein, R. (1990) "Solution structure of the glucocorticoid receptor DNA-binding domain." *Science* **249**, 157-160.
- Helbecque, N., Mocquin, C., Bernier, J.L., Morel, E., Guyot, M. & Henichart, J.P. (1987) "Grossularine-1 and grossularine-2, α -carboline alkaloids from *Dendrodoa grossularia*, as possible intercalative agents." *Cancer Biochem. Biophys.* **9**, 271-279.
- Hol, W.G.J., vanDuijnen, P.T. & Berendsen, H.G.J. (1978) "The α -helix dipole and the properties of proteins." *Nature* **273**, 443-447.
- Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M.G. & Evans, R.M. (1985) "Primary structure and expression of a functional human glucocorticoid receptor cDNA." *Nature* **318**, 635-638.
- Hollenberg, S.M., Giguere, V., Segui, P. & Evans, R.M. (1987) "Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor." *Cell* **49**, 39-47.
- Huxley, J.S. (1929) "Thyroid and temperature effects in cold-blooded vertebrates." *Nature* **123**, 712.
- Ivarie, R.D. & O'Farrell, P.H. (1978) "The glucocorticoid domain: Steroid mediated changes in the rate of synthesis of rat hepatoma proteins." *Cell* **13**, 41-49.
- Jantzen, H.M., Strahle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R. & Schutz, G. (1987) "Cooperativity of glucocorticoid response element located far upstream of the tyrosine aminotransferase gene." *Cell* **49**, 29-38.
- Jeltsch, J.M., Krozowski, Z., Quirin-Stricker, G., Gronemeyer, H., Simpson, R.J., Garnier, J.M., Krust, A., Jacob, F. & Chambon, P. (1986) "Cloning of the chicken progesterone receptor." *Proc. Natl. Acad. Sci. USA* **83**, 5424-5428.
- Jensen, E.V. & De Sombre, E.R. (1972) "Mechanism of action of the female sex hormones." *Annu. Rev. Biochem.* **41**, 203-284.
- Johnson, P.F., Landschultz, W.H., Graves, B.J. & McKnight, S.L. (1987) "Identification of a rat liver protein that binds to the enhancer core element of three animal viruses." *Genes Dev.* **1**, 133-146.

- Kalderon, D., Roberts, B.L., Richardson, W.D. & Smith, A.E. (1984) "A short amino acid sequence able to specify nuclear location." *Cell* **39**, 499-509.
- Kammerer, B., Guyonvarch, A. & Hubert, J.C. (1984) "Yeast regulatory gene PPR1. I. Nucleotide sequence, restriction map and codon usage." *J. Mol. Biol.* **180**, 239-250.
- Karin, M., Andersen, R.D., Slater, E., Smith, K. & Herschman, H.R. (1980) "Metallothionein mRNA induction in Hela cells in response to zinc or dexamethasone is a primary induction response." *Nature* **286**, 295-298.
- Kenani, A., Lohez, M., Houssin, R., Helbecque, N., Bernier, J.L., Lemay, J.P. & Henichart, J.P. (1987) "Chelating, DNA-binding and DNA-cleaving properties of a synthetic model for bleomycin." *Anti-Cancer Drug Design* **2**, 47-59.
- Kendall, E.C. (1915) "The isolation in crystalline form of the compound containing iodine, which occurs in the thyroid." *J. Am. Med. Assoc.* **64**, 2042-2063.
- Klein-Hitpaß, L., Schorpp, M., Wagner, U. & Ryffel, G.U. (1986) "An estrogen responsive element derived from 5'-flanking region of the vitellogenin A2 gene functions in human transfected cells." *Cell* **46**, 1053-1062.
- Krieger, D. (1981) in *Endocrinology and Metabolism*, P. Felig, D. Baxter, A.E. Broadus, L.A. Frohman, Eds. (McGraw-Hill, New-York), pp.125-149.
- Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J.M. & Chambon, P. (1986) "The chicken estrogen receptor sequence: homology with v-erb-A and the human estrogen and glucocorticoid receptors." *EMBO J.* **5**, 891-897.
- Kumar, V., Green, S., Staub, A. & Chambon, P. (1986) "Localisation of the estradiol binding and putative DNA-binding domains of the human estrogen receptor." *EMBO J.* **5**, 2231-2241.
- Kurtz, D.T. & Feigelson, P. (1977) "Multi-hormonal induction of hepatic α_2 -globulin mRNA as measured by hybridization to complementary DNA." *Proc. Natl. Acad. Sci. USA* **74**, 4791-4797.
- Laimins, L.A., Khoury, G., Gorman, C., Howard, B. & Gruss, P. (1982) "Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus." *Proc. Natl. Acad. Sci. USA* **79**, 6453-6460.

- Landschulz, W.H., Johnson, P.F. & McKnight, S.L. (1988) "The leucine-zipper: a hypothetical structure common to a new class of DNA-binding proteins." *Science* **240**, 1759-1764.
- Lanford, R.E. & Butel, J.S. (1984) "Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen." *Cell* **37**, 801-813.
- Lanford, R.E., White, R.G., Dunham, R.G. & Kanda, P. (1988) "Effect of basic and non-basic aminoacid substitutions on transport induced by SV40 T-antigen synthetic peptides nuclear transport signals." *Mol. Cell. Biol.* **8**, 2722-2729.
- Laughon, A. & Gesteland, R.F. (1984) "Primary structure of the *Saccharomyces cerevisiae* GAL4 gene." *Mol. Cell. Biol.* **4**, 260-267.
- Loosfelt, H., Atger, M., Misrahi, M., Guiochon-Mantel, A., Meriel, C., Logeat, F., Benarous, R. & Milgrom, E. (1986) "Cloning and sequence analysis of rabbit progesterone receptor complementary DNA." *Proc. Natl. Acad. Sci. USA* **83**, 9045-9053.
- Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R. & Sigler, P.B. (1991) "Crystallographic analyses of the interaction of the glucocorticoid receptor with DNA." *Nature* **352**, 497-505.
- MacDonnell, D.R., Mangelsdorf, D.J., Wesley-Pike, J., Haussler, M.R. & O'Malley, B.W. (1987) "Molecular cloning of complementary DNA encoding the avian receptor for vitamin D." *Science* **235**, 1214-1216.
- McKay, D.B. & Steitz, T.A. (1981) "Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA." *Nature* **290**, 744-749.
- Mack, D., Iverson, B.L. & Dervan, P.B. (1988) "Design and chemical synthesis of a sequence-specific DNA-cleaving molecule." *J. Am. Chem. Soc.* **110**, 7572-7574.
- Mader, S., Kumar, V., de Verneuil, H. & Chambon, P. (1989) "Three aminoacids from the estrogen receptor are essential to its ability to distinguish an estrogen from a glucocorticoid responsive element." *Nature* **338**, 271-274.
- Menino, G., Farrace, M.G., Ceru, M.P. & Piacentini, M. (1988) "Correlation between transglutaminase activity and polyamine levels in human neuroblastoma cells." *Exp. Cell Res.* **179**, 429-445.

- Messenguy, R., Dubois, E. & Descamps, F. (1986) "Nucleotide sequence of the ARGRII regulatory gene and aminoacid sequence homologies between ARGRII, PPRI and GAL4 regulatory proteins." *Eur. J. Biochem.* **157**, 77-81.
- Miesfeld, R., Rusconi, S., Godowski, P.J., Maler, B.J., Okret, S., Wikstrom, A.C., Gustafsson, J.-A. & Yamamoto, K.R. (1986) "Genetic complementation of a glucocorticoid receptor deficiency by expression of cloned receptor cDNA." *Cell* **46**, 389-399.
- Miller, J., McLachlan, D. & Klug, A. (1985) "Repetitive zinc-binding domains in protein transcription factor IIIA from *Xenopus* oocytes." *EMBO J.* **4**, 1609-1614.
- Misrahi, L., Atger, M., d'Auriol, L., Loosfelt, H., Meriel, C., Fridlanski, F., Guiochon-Mantel, A., Galibert, F. & Milgrom, E. (1987) "Complete aminoacid sequence of the human progesterone receptor deduced from cloned cDNA." *Biochem. Biophys. Res. Commun.* **143**, 740-745.
- Morier-Teissier, E., Bailly, C., Houssin, R., Bernier, J.L., Helbecque, N., Catteau, J.P., Colson, P., Houssier, C. & Henichart, J.P. (1989) "Synthesis, biological activity and DNA interaction of anilinoacridine and bithiazole peptide derivatives related to the antitumor drugs m-amsa and bleomycin." *Anti-Cancer Drug Design* **4**, 37-52.
- Moser, H. & Dervan, P.B. (1987) "Sequence-specific cleavage of double helical DNA by triple helix formation." *Science* **238**, 645-648.
- Needham, J. (1983) *Science and Civilisation in China* (Cambridge Univ. Press, Cambridge, MA.).
- Nervi, C., Grippo, J.F., Sherman, M.I., George, M.D. & Jetten, A.M. (1989) "Identification and characterization of nuclear retinoic acid binding activity in human myeloblastic leukemia HL60 cells." *Proc. Natl. Acad. Sci. USA* **86**, 5854-5858.
- Newmeyer, D.D. & Forbes, D.J. (1988) "Nuclear import can be separated in two steps *in-vitro*; nuclear pore binding and translocation." *Cell* **52**, 641-653.
- Okret, S., Carlstedt-Duke, J., Wrangé, O., Carlstrom, K. & Gustafsson, J.-A. (1982) "Characterization of an antiserum against the glucocorticoid receptor." *Biochim. Biophys. Acta* **677**, 205-212.
- O'Neil, K., Hoes, R.H. & DeGrado, W.F. (1990) "Design of DNA-binding peptides based on the leucine-zipper motif." *Science* **249**, 774-778.

- Oppenheimer, J.H., Korner, D., Schwartz, H.L. & Saules, M.I. (1972) "Specific nuclear triiodothyronine binding sites in rat liver and kidney." *J. Clin. Endocrinol. Metab.* **35**, 330-347.
- Ostrowski, M.C., Huang, A.L., Kessel, M., Wolford, R.G. & Hager G.L. (1984) "Modulation of enhancer activity by the hormone responsive regulatory element from mouse mammary tumor virus." *EMBO J.* **3**, 1891-1899.
- Pabo, C.O. & Lewis, M. (1982) "The operator binding domain of λ repressor: structure and DNA recognition." *Nature* **298**, 443-447.
- Pabo, C.O. & Sauer, R.T. (1984) "Protein-DNA recognition." *Annu. Rev. Biochem.* **53**, 293-321.
- Parks, W.P., Scolnick, E.M. & Kozikowski, E.M. (1974) "Dexamethasone stimulation of mouse mammary tumor virus expression: a tissue culture source of virus." *Science* **184**, 158-161.
- Peisach, J. & Blumberg, W.E. (1974) "Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins." *Arch. Biochem. Biophys.* **165**, 691-708.
- Petkovich, M., Brand, N.J., Krust, A. & Chambon, P. (1987) "A human retinoic acid receptor which belongs to the family of nuclear receptors." *Nature* **330**, 444-447.
- Pfahl, M. (1982) "Specific binding of the glucocorticoid receptor complex to the mouse mammary tumor proviral promoter region." *Cell* **31**, 475-482.
- Picard, D. & Yamamoto, K.R. (1987) "Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor." *EMBO J.* **6**, 3333-3340.
- Plouvier, B., Bailly, C., Houssin, R. & Henichart, J.P. (1991) "Synthesis of two new thiazole-containing oligopeptides as potential DNA minor groove binding analogs of netropsin." *Heterocycles* **32**, 691-701.
- Pullman, A. & Pullman, B. (1981) "Molecular electrostatic potentials of nucleic acids." *Quart. Rev. Biophys.* **14**, 289-380.
- Rhodes, D. & Klug, A. (1986) "An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA." *Cell* **46**, 123-132.

- Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A. & Dingwall, C. (1988) "Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores." *Cell* **52**, 655-664.
- Ringold, G., Yamamoto, K.R., Tomkins, G.M., Bishop, J.M. & Warmus, H.E. (1975) "Dexamethasone-mediated induction of mouse mammary tumor virus RNA: a system for studying glucocorticoid action." *Cell* **6**, 299-306.
- Ringold, G.M. (1985) "Steroid hormone regulation of gene expression." *Annu. Rev. Pharmacol. Toxicol.* **25**, 529-599.
- Robertson, M. (1987) "Towards a biochemistry of morphogenesis." *Nature* **330**, 420.
- Sakonju, S., Bogenhagen, D.F. & Brown, D.D. (1980) "A control region in the center of the 5S RNA gene directs specific initiation of transcription. I. The 5' border region." *Cell* **19**, 13-25.
- Sakonju, S., Brown, D.D., Engelke, D., Ng, S.Y., Shastry, B.S. & Roeder, R.G. (1981) "The binding of a transcription factor to deletion mutants of a 5S ribosomal RNA gene." *Cell* **23**, 665-669.
- Sanchez, E.R., Toft, D.O., Schlesinger, M.J. & Pratt, W.B. (1985) "Evidence that the 90-kDa phosphoprotein associated with the untransformed L-cell glucocorticoid receptor is a murine heat-shock protein." *J. Biol. Chem.* **260**, 12398-12401.
- Sap, J., Munoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H. & Vennstrom, B. (1986) "The c-erb-A protein is a high affinity receptor for thyroid hormone." *Nature* **324**, 635-640.
- Scheidereit, C., Geisse, S., Westphal, H.M. & Beato, M. (1983) "The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of mouse mammary tumor virus." *Nature* **304**, 749-751.
- Schultz, P., Taylor, J.S. & Dervan, P.B. (1982) "Design and synthesis of a sequence-specific DNA-cleaving molecule. (Distamycin-EDTA) Iron(II)." *J. Am. Chem. Soc.* **104**, 6861-6863.
- Schwabe, J.W.R., Neuhaus, D. & Rhodes, D. (1990) "Solution structure of the DNA-binding domain of the estrogen receptor." *Nature* **348**, 458-461.

- Segall, J., Matsui, T. & Roeder, R.G. (1980) "Multiple factors are required for the accurate transcription of purified genes by RNA polymerase III." *J. Biol. Chem.* **255**, 11986-11991.
- Silver, P.A., Chiang, A. & Sadler, I. (1988) "Mutation that alter both localization and production of yeast nuclear protein." *Genes Dev.* **2**, 707-717.
- Shastry, B.S., Ng, S.Y. & Roeder, R.G. (1982) "Multiple factors involved in the transcription of class III genes in *Xenopus laevis*." *J. Biol. Chem.* **257**, 12979-12986.
- Sluka, J., Horvath, S.J., Bruist, M.F., Simon, M.I. & Dervan, P.B. (1987) "Synthesis of a sequence-specific DNA-cleaving peptide." *Science* **238**, 1129-1132.
- Spindler, S.R., Mellon, S.H. & Baxter, J.D. (1982) "Growth hormone gene transcription is regulated by thyroid and glucocorticoid hormones in cultured rat pituitary tumor cells." *J. Biol. Chem.* **257**, 11627-11633.
- Steitz, T.A., Ohlendorf, D.H., McKay, D.B., Anderson, W.F. & Matthews, B.W. (1982) "Structural similarities in the DNA-binding domains of catabolite gene activator and cro repressor proteins." *Proc. Natl. Acad. Sci., USA.* **79**, 3097-3100.
- Sung, M.T. & Dixon, G.M. (1970) "Modification of histones during spermiogenesis in trout: a molecular mechanism for altering histone binding to DNA." *Proc. Natl. Acad. Sci. USA* **67**, 1616-1623.
- Thompson, C.C., Weinberger, C., Lebo, R. & Evans, R.M. (1987) "Identification of a novel thyroid hormone receptor expressed in the central nervous system." *Science* **237**, 1610-1614.
- Troen, P. & Oshima, H. in *Endocrinology and Metabolism*, P. Felig, D. Baxter, A.E. Broadus, L.A. Frohman, Eds. (McGraw-Hill, New-York), pp. 627-668.
- Umesono, K. & Evans, R.M. (1989) "Determinants of target gene specificity for steroid/thyroid hormone receptors." *Cell* **57**, 1139-1146.
- Venkataram Prasad, B.V. & Balaram, P. (1984) "The stereochemistry of peptides containing α -aminoisobutyric acid." *CRC Crit. Rev. Biochem.* **16**, 307-348.
- Vogt, P.K., Bos, T.J. & Doolittle, R.F. (1987) "Homology between the DNA-binding domain of the GCN4 regulatory protein of yeast and the carboxy-terminal region of a protein encoded by the oncogene jun." *Proc. Natl. Acad. Sci., USA.* **84**, 3316-3319.

- von der Ahe, D., Janish, S., Scheidereit, C., Renkawitz, R., Schutz, G. & Beato, M. (1985) "Glucocorticoid and progesterone receptors bind to the same sites in two hormonally regulated promoters." *Nature* **313**, 706-709.
- Vrana, K., Churchill, M.E.A., Tullius, T.D. & Brown, D.D. (1988) "Mapping functional regions of transcription factor IIIA." *Mol. Cell. Biol.* **8**, 1684-1696.
- Warrant, R.W. & Kim, S.H. (1978) " α -helix-double-helix interaction shown in the structure of protamine-transfer RNA complex and a nucleoprotamine model." *Nature* **271**, 130-135.
- Weinberger, C., Hollenberg, S.M., Rosenfeld, M.G. & Evans, R.M. (1985) "Domain structure of human glucocorticoid receptor and its relationship to the v-erb-A oncogene product." *Nature* **318**, 670-674.
- Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J. & Evans, R.M. (1986) "The c-erb-A gene encodes a thyroid hormone receptor." *Nature* **324**, 641-644.
- Wrangé, O. & Gustafsson, J.-Å. (1978) "Separation of the hormone- and DNA-binding sites of the hepatic glucocorticoid receptor by means of proteolysis." *J. Biol. Chem.* **253**, 856-876.
- Wrangé, O., Carlstedt-Duke, J. & Gustafsson, J.-A. (1979) "Purification of the glucocorticoid receptor from rat liver cytosol." *J. Biol. Chem.* **254**, 9284-9290.
- Wrangé, O., Okret, S., Radojic, M., Carlstedt-Duke, J. & Gustafsson, J.-A. (1984) "Characterization of the purified activated glucocorticoid receptor from rat liver cytosol." *J. Biol. Chem.* **259**, 4534-4545.
- Wray, L.V., Witte, M.M., Dickson, R.C. & Riley, M.I. (1987) "Characterization of a positive regulatory gene Lac9 that controls induction of the lactose-galactose regulon of *Kluyveromyces lactis*: structural and functional relationships to GAL4 of *Saccharomyces cerevisiae*." *Mol. Cell. Biol.* **7**, 1111-1121.
- Wüthrich, K. (1990) "Protein structure determination in solution by NMR spectroscopy." *J. Biol. Chem.* **265**, 22059-22062.
- Yamamoto, K.R. (1985) "Steroid receptor regulated transcription of specific genes and gene network." *Annu. Rev. Genet.* **19**, 209-252.



Zubay, G. & Doly, P. "The isolation and properties of deoxyribonucleoprotein particles containing single nucleic acid molecules." (1959) *J. Mol. Biol.* **1**, 1-20.

- SUMMARY -

Proteins from the steroid/thyroid hormone receptors superfamily are ligand-dependent transcription factors. The cellular pathway mediated by these nuclear receptors results in a regulation of transcription of target genes. Two essential steps in which molecular recognition plays a dominant role are required to achieve this activity : (i) **nuclear import** of the receptor and (ii) **DNA specific binding** to a definite hormone responsive element. For each step, biological studies have inferentially involved a short domain of the conserved structure among receptors.

The work presented here describes the chemical synthesis of peptides, of which sequences correspond to the putative functional domains and the demonstration of their efficiency out of receptor context : (i) **a nuclear localization signal** is able to direct to the nucleus a nonnuclear protein on which it has been previously coupled and (ii) **a "zinc-finger" motif** (or its structural reduction to an alpha-helix) shows the same selectivity in DNA-binding as the whole purified receptor. In any case, structural features of the structure-function relationship are discussed.

In a last part, we made use of the zinc-finger motif as a recognition "tool" for the elaboration of a very high specific **artificial endonuclease**.

- RESUME -

Les protéines de la superfamille des récepteurs des hormones stéroïdiennes / thyroïdiennes sont des facteurs de transcription dont l'activité est sous la dépendance d'un ligand. Le mécanisme d'action de ces récepteurs hormonaux nucléaires résulte *in-fine* en une régulation du taux de transcription de gènes cibles et fait intervenir pour cela deux étapes où la reconnaissance moléculaire est prépondérante : **(i) la translocation du récepteur vers le compartiment nucléaire**, et **(ii) la reconnaissance spécifique d'une séquence d'ADN**. Pour chacune de ces étapes, il est possible d'identifier dans la formule du récepteur une séquence en acides aminés inférentiellement impliquée par des études biochimiques et de biologie moléculaire.

Les études rapportées dans ce mémoire décrivent la synthèse chimique de peptides correspondant à ces séquences et la démonstration de leur fonctionnalité hors du contexte de la protéine : **(i) un peptide signal de localisation nucléaire** est suffisant pour diriger vers le compartiment nucléaire une protéine sur laquelle il est chimiquement greffé, et **(ii) un motif en "doigt de zinc"** (ou sa réduction structurale en une hélice alpha) est capable de reconnaître spécifiquement la même séquence d'ADN que le récepteur entier. Pour chaque peptide, l'aspect structural de la relation structure-fonction est discuté.

Une dernière partie traite de l'utilisation du motif de reconnaissance de l'ADN comme "outil" pour l'élaboration d'une **endonucléase artificielle** de très haute spécificité.

MOTS - CLES

Récepteur

Stéroïde

Doigt de zinc

Adressage nucléaire

Interactions protéines/ADN

Endonucléase

Reconnaissance