Mécanismes d'échappement des cellules tumorales mammaires au contrôle hormonal : étude de l'impact du biais d'usage des codons

Léa Clusan

▶ **To cite this version:**

Léa Clusan. Mécanismes d'échappement des cellules tumorales mammaires au contrôle hormonal : étude de l'impact du biais d'usage des codons. Médecine humaine et pathologie. Université de Rennes, 2022. Français. NNT : 2022REN1B053 . tel-04290727

HAL Id: tel-04290727

https://theses.hal.science/tel-04290727v1

Submitted on 17 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE RENNES 1

THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1

ECOLE DOCTORALE N° 605
Biologie Santé
Spécialité : « *Biologie cellulaire, biologie du développement* »

Par
Léa CLUSAN

Mécanismes d'échappement des cellules tumorales mammaires au contrôle hormonal : étude de l'impact du biais d'usage des codons

Thèse présentée et soutenue à Rennes, le 16 novembre 2022
Unité de recherche : Irset – Inserm UMR_S 1085

Rapporteurs avant soutenance :

Anne-Catherine PRATS	Directrice de recherche Inserm (HDR), Inserm/UPS UMR 1297, Toulouse
Jérôme EECKHOUTE	Directeur de recherche CNRS (HDR), Inserm U1011, Université de Lille

Composition du Jury :

Président :	
Reynald GILLET	Professeur des Universités UR1 (HDR), UMR 6290 CNRS-UR1, Rennes
Examinateurs :	
Anne-Catherine PRATS	Directrice de recherche Inserm (HDR), Inserm/UPS UMR 1297, Toulouse
Jérôme EECKHOUTE	Directeur de recherche CNRS (HDR), Inserm U1011, Université de Lille
Coralie FONTAINE	Chargée de recherche INSERM (HDR), Inserm/UPS UMR 1297, Toulouse
Dir. de thèse :	
Farzad PAKDEL	Directeur de recherche CNRS (HDR), Irset – Inserm UMR_S 1085, Rennes
Gilles FLOURIOT	Chargé de recherche CNRS, Irset – Inserm UMR_S 1085, Rennes

Remerciements

Ce manuscrit est l'aboutissement de trois années riches en travail et apprentissage, dans un contexte sanitaire sans précédent, qui m'auront autant appris sur le plan personnel que professionnel. J'ai ainsi croisé le chemin de nombreuses personnes que je souhaite remercier ici.

Tout d'abord, merci beaucoup à Anne-Catherine Prats et Jérôme Eeckhoute, qui prendront le temps d'évaluer ces travaux. Merci également à Coralie Fontaine et Reynald Gillet d'avoir accepté de prendre part à mon jury de thèse. Il me tarde de discuter avec vous de mon projet de thèse et des perspectives que celui-ci annonce !

Je remercie également Jean-François Arnal d'avoir participé à mon comité de suivi individuel aux côtés de Reynald Gillet. Vos conseils scientifiques m'ont été très bénéfiques au cours de l'avancée de la thèse.

Surtout, je remercie tout particulièrement mes directeurs de thèse, Farzad Pakdel et Gilles Flouriot, pour m'avoir fait confiance dans la réalisation de ce projet et m'avoir accompagnée lors de ces trois années. Vous avez tous deux toujours su vous rendre disponibles pour répondre à mes questions et prodiguer vos conseils, tant techniques que rédactionnels. Je salue notamment Farzad pour ton recul et tes commentaires constructifs le long du projet, et Gilles pour ta passion et tes idées originales, me faisant découvrir la « vraie » recherche. Je vous suis très reconnaissante également de m'avoir accordé la flexibilité qui m'a permis de réaliser des missions d'enseignement et de m'investir dans différentes initiatives de médiation scientifique en parallèle de mes travaux de recherche.

Je souhaite également remercier le directeur de l'Irset, Michel Samson, pour m'avoir accueillie ici et pour œuvrer au développement de cet institut très accueillant et engagé pour la santé publique et le partage des connaissances avec le grand public.

Je remercie également chaleureusement tous les membres de l'équipe, que j'ai eu plaisir à côtoyer lors de ces trois années. Quelques mots supplémentaires pour les personnes avec lesquelles j'ai le plus particulièrement travaillées :

- merci à Fred, de m'avoir formée à certaines techniques au laboratoire, ta rigueur dans le soin apporté en salle de culture cellulaire et à nos chères lignées de cellules, mais aussi pour tout le travail réalisé pour générer tant de plasmides synonymes et délétés ;
- merci à Catherine, pour ta disponibilité dès que je ne trouve pas quelque chose au laboratoire, ta gestion des commandes, et les midipreps réalisées pour refaire nos stocks de plasmides ;
- merci à Yann, pour l'entretien de la salle de Western Blot et tes conseils avisés dans leur réalisation, mais surtout pour ton aide dans la mise au point et la réalisation de toutes ces expériences de polysome profiling (et toutes ces plaques de qPCR) ; j'ai toujours apprécié ton caractère entier, qui a rendu ses aller-retours au campus de Beaulieu beaucoup plus plaisants !

- merci à Pascale, pour tes corrections et ta participation à l'écriture de ma première revue, et ta bonne humeur qui complète notre bureau ;
- merci également à Christophe, pour ta bonne humeur lorsque l'on se croise aux premières lueurs du matin !

Une grande partie de mon projet de thèse a nécessité la réalisation d'ultracentrifugations en gradients de sucrose, rendues possibles grâce à l'équipe de Reynald Gillet à l'IGDR. Je remercie alors encore une fois Reynald pour nous avoir permis d'utiliser leurs ultracentrifugeuses et matériel spécifique, et Claire Heichette pour la formation à l'utilisation de ces machines. Merci énormément à Fanny Demay pour m'avoir formée à ces expériences, tes réponses à mes questions, et tous tes conseils concernant les ribosomes. Merci également à Sylvie Georgeault-Daguenet pour ton aide dans la logistique de ces expériences ! J'ai également eu l'occasion de voir « mes » ribosomes par microscopie électronique, grâce à Sophie Chat que je remercie beaucoup pour cette expérience et la session de formation dont j'ai pu profiter.

Merci aussi énormément à Raphaël Métivier pour tout le travail réalisé dans l'analyse du séquençage de l'ARN, ta pédagogie pour m'expliquer les dessous de ces résultats, et ta disponibilité pour répondre à toutes mes questions sur le sujet.

Tout aussi pédagogue, je remercie Emmanuelle Com pour sa formation sur la spectrométrie de masse, c'était un plaisir de travailler avec toi lors de cette journée !

Je remercie également Christine Baysse et Renan Goude pour leur accompagnement en TD et TP de microbiologie, qui m'ont permis de vivre mon expérience d'enseignement avec autant de sérénité que possible.

Enfin, je remercie sincèrement tout le personnel de soutien à la recherche dans l'institut, œuvrant à la laverie et logistique, à l'administration, à l'hygiène et sécurité ; avec un clin d'œil particulier à Jeannick, dont j'aurai croisé le sourire tous les jours de bon matin.

Des pensées également pour les autres doctorants de l'Irset, en particulier Hà et Morgane, que je remercie pour nos discussions et conseils échangés avec plaisir tout au long de nos parcours respectifs.

Pour finir, je suis extrêmement reconnaissante du soutien dont j'ai bénéficié de la part de ma famille et amis. Merci notamment à Marius, le meilleur des témoins, à qui je dois ma toute première figure de revue. A ma grand-mère, dont la force de caractère m'a inspirée et qui m'a accompagnée en pensées jusqu'à la fin. Et surtout merci, merci, à mon (tout nouveau) mari, dont le soutien et les encouragements n'ont jamais faibli au cours de ces trois années de travail et de remise en question. Cette aventure se termine avec de nouvelles perspectives professionnelles que j'ai hâte d'explorer, en continuant ma vie à tes côtés.

Sommaire

Liste des abréviations

Ψ	pseudouridine
AF1/2	transcriptional activation function 1/2
AI	aromatase inhibitor
AIB1	amplified in breast cancer 1
AMPc	adénosine monophosphate cyclique
AP1	activator protein 1
AREG	amphiregulin
ARNm	ARN messager
ARNr	ARN ribosomique
ARNt	ARN de transfert
AT3	A ou T à la troisième position d'un codon
BCL2	B-cell lymphoma 2
BRCA1/2	breast cancer associated gene 1/2
CARM1	coactivator associated arginine methyltransferase 1
CBP	CREB binding protein
CCND1/E1	cyclin D1/E1
CDK4/6	cyclin-dependent kinase 4/6
CFTR	cystic fibrosis transmembrane conductance regulator
CXCL12	C-X-C motif chemokine ligand 12
DBD	DNA-binding domain
DDX6	DEAD-box helicase 6
DKC1	dyskerin
E2	estrogen
EGFR	epidermal growth factor receptor
EMT	epithelial-mesenchymal transition
ERα	estrogen receptor alpha
ERβ	estrogen receptor beta
ERBS	estrogen receptor binding site
ERE	estrogen responsive element
ERK	extracellular signal-regulated kinase
ESR1/2	estrogen receptor 1/2
FBL	fibrillarin
FOXA1/M1	forkhead box A1/M1
GATA3	GATA binding protein 3
GC3	G ou C à la troisième position d'un codon
GF	growth factor
GPR30	G protein-coupled receptor 30
GREB1	growth regulating estrogen receptor binding 1
HER2	human epidermal growth factor receptor 2
HSP	heat shock protein
IGF1	insulin like growth factor 1
IGF1R	insulin like growth factor 1 receptor
IRES	internal ribosome entry site
LBD	ligand-binding domain
LBP	ligand-binding pocket
MAPK	mitogen-activated protein kinase
MCF7	Michigan cancer foundation-7
mESCs	mouse embryonic stem cells

NCoR nuclear receptor corepressor

NLS nuclear localization signal

PDX patient derived xenograft

PELP1 proline-, glutamate- and leucine-rich protein 1

PFAR protein folding activity of ribosomes

PI3K phosphoinositide 3-kinase

PR progesterone receptor

PROTAC proteolysis-targeting chimera

PTEN phosphatase and tensin homolog

PTM post-translational modification

RANKL receptor activator of nuclear factor-κB ligand

RBM47 RNA binding motif protein 47

RFLP restriction fragment length polymorphism

SERCA selective ER covalent antagonist

SERD selective estrogen receptor downregulator

SERM selective estrogen receptor modulator

siRNA small interfering RNA

SMRT silencing mediator for retinoid and thyroid hormone receptor

SP1 specificity protein 1

SRC1 steroid receptor coactivator 1

SRE serum responsive element

TF transcription factor

TP53 tumor protein p53

UTR untranslated region

Liste des figures et tableaux

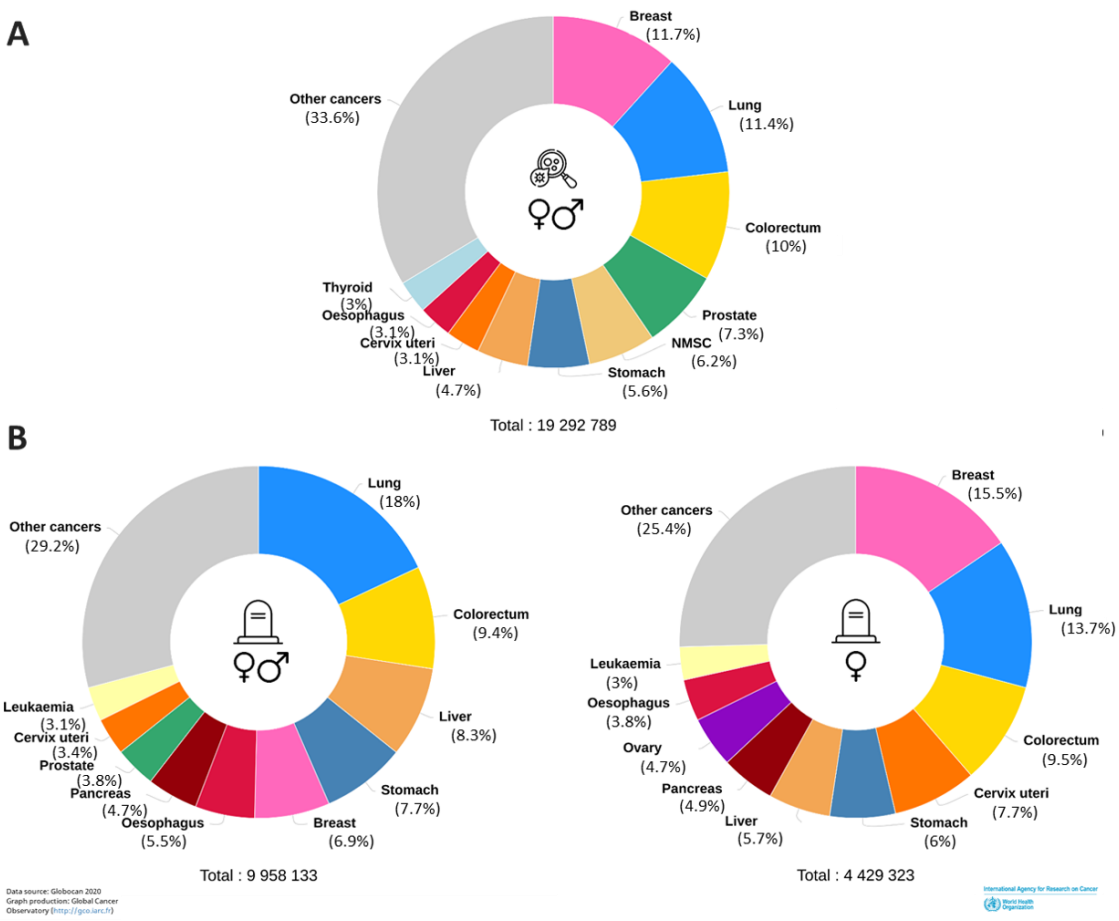


Figure 1 : Epidémiologie du cancer dans le monde.

A) Proportion de nouveaux cas de chaque cancer estimée en 2020, tout âge et sexe confondus. B) Proportion de décès dus à chaque cancer estimée en 2020, tout âge et sexe confondus (à gauche), ou chez la femme (à droite). Figures issues puis modifiées à partir du Global Cancer Observatory (https://gco.iarc.fr/today/home, consulté le 24 janvier 2022). NMSC : non-melanoma skin cancer.

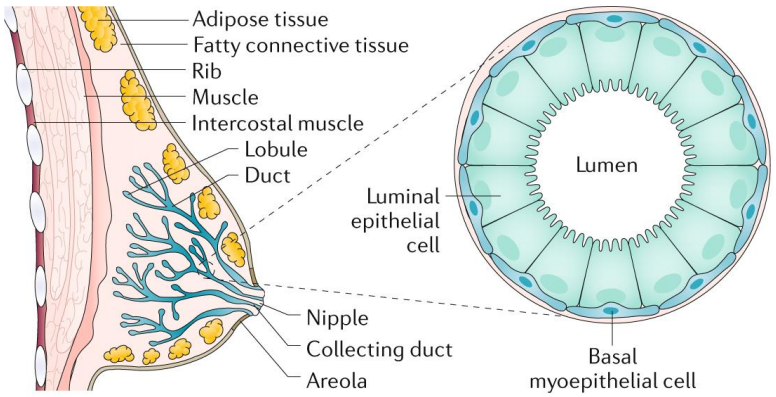


Figure 2 : Anatomie du tissu mammaire.

L'épithélium des lobules et canaux galactophores est constitué d'une couche de cellules luminales puis une couche de cellules basales. Figure issue de Harbeck et al. [1].

Introduction

1. LE CANCER DU SEIN

a) Epidémiologie, développement et prise en charge

Parmi les 19,3 millions de nouveaux cas de cancer recensés dans le monde en 2020, le cancer du sein est devenu le plus diagnostiqué, juste devant le cancer du poumon (représentant 11,7% et 11,4% des nouveaux cas de cancer, respectivement ; Figure 1, A). Le cancer du poumon reste à l'origine de la majorité des décès tout sexe confondu (18% des 10 millions de décès dus aux cancers dans le monde) tandis que le cancer du sein est le plus meurtrier chez la femme (6,9% des décès, tout sexe confondu, mais 15,5% des décès chez la femme ; Figure 1, B) [2]. En France, le cancer du sein est effectivement le cancer le plus fréquent chez la femme, avec près de 59 000 nouveaux cas et 12 000 décès estimés en 2018 d'après l'Institut National du Cancer.

Le sein est un organe qui contient la glande mammaire, dont la fonction est de produire et excréter du lait. Le tissu mammaire est alors constitué de lobules, qui vont produire le lait, et de canaux connectant les lobules au mamelon, le tout étant entouré de tissus conjonctifs, adipeux, et lymphatique. L'épithélium des lobules et des canaux galactophores est constitué d'une première couche de cellules luminales et d'une seconde couche de cellules basales (Figure 2). Le cancer du sein désigne alors la prolifération anormale de ces cellules mammaires, débutant en général au niveau des lobules [3]. Après s'être développé localement, le cancer peut devenir invasif en propageant des métastases à d'autres organes tels que les os, les poumons, le foie ou encore le cerveau [4]. Il n'existe ainsi pas un mais des cancers du sein, différant de par leur localisation, type histologique, et signature moléculaire. La combinaison de ces facteurs est ainsi à l'origine de cancers du sein dont l'agressivité et donc le pronostique varie grandement, dont la prise en charge thérapeutique sera alors différente [1]. La classification des cancers du sein suit ainsi les recommandations de l'OMS (Organisation Mondiale de la Santé), révisées régulièrement pour permettre une meilleure prise en charge des patientes [5].

L'un des principaux paramètres permettant la classification des cancers du sein repose sur leur profil moléculaire tel que décrit en 2000 par Perou et al. [6]. Cette étude a en effet mis en évidence l'hétérogénéité des cancers du sein au niveau moléculaire via l'expression différentielle d'un panel de gènes clefs. Les profils d'expression obtenus permettent ainsi de diviser le cancer du sein en quatre principaux groupes :

- luminal A (60% des cas)
- luminal B (10% des cas)
- HER2-positif (20% des cas)
- triple négatif (10% des cas)

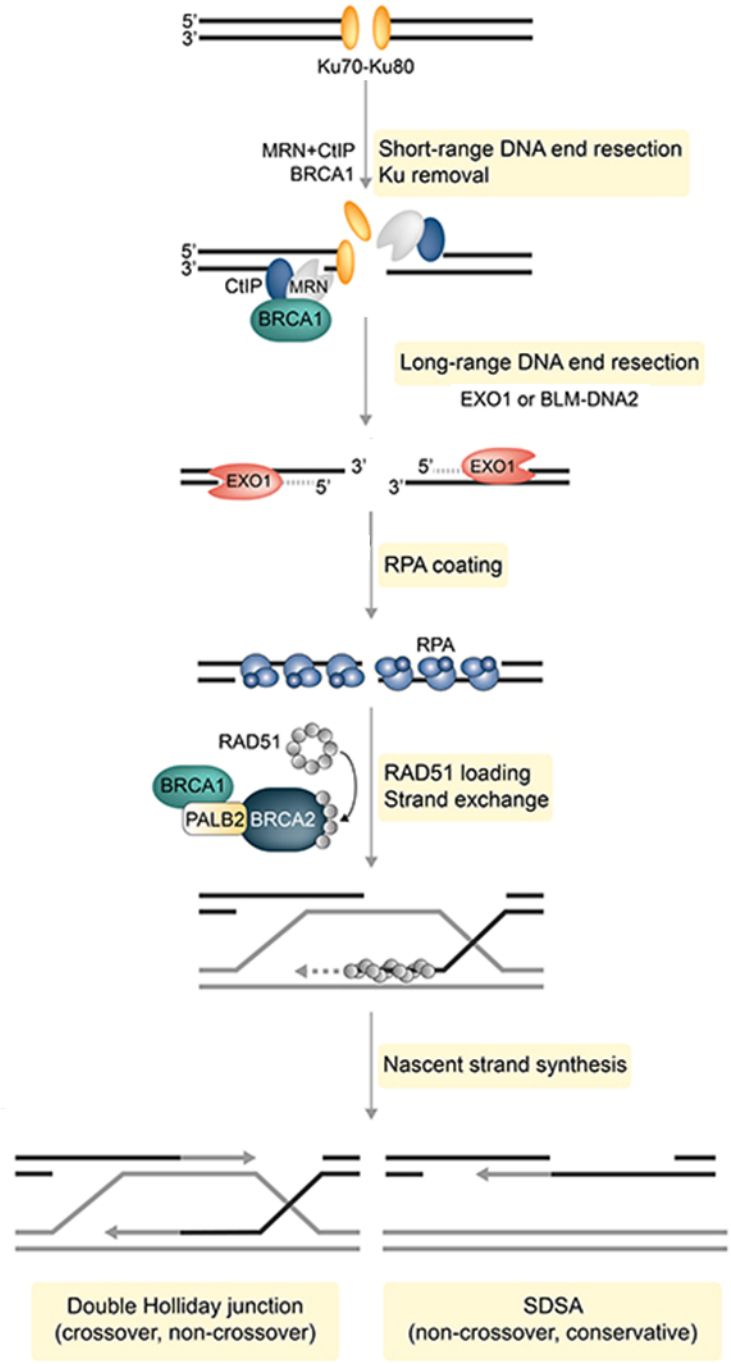


Figure 3 : Mécanisme de réparation des cassures d'ADN double-brin par recombinaison homologue. Le complexe MRN reconnaît d'abord la cassure double-brin, puis BRCA1 et CtIP sont recrutés et induisent l'excision des brins 5' par EXO1 ou BLM-DNA2. Les extrémités 3' simple-brin sont alors recouvertes d'hétérotrimères RPA. Le complexe BRCA1-PALB2-BRCA2 permet ensuite l'assemblage de monomères de RAD51 sur l'ADN simple-brin, permettant l'invasion du brin d'ADN homologue. La synthèse du brin d'ADN manquant a alors lieu, et la réparation s'achève généralement par SDSA (synthesis-dependent strand annealing) en cellules somatiques, ou la résolution d'une double jonction de Holliday peut avoir lieu. Figure issue et modifiée de Trenner and Sartori [7]. BLM : Bloom's syndrome helicase ; BRCA1/2 : breast cancer associated gene 1/2 ; CtIP : CtBP interacting protein ; DNA2 : DNA replication helicase/nuclease 2 ; EXO1 : exonuclease 1 ; MRN : MRE11-RAD50-NBS1 ; PALB2 : partner and localizer of BRCA2 ; RPA : replication protein A.

Les cancers luminaux sont caractérisés par l'expression du récepteur aux œstrogènes ERα et sont les moins agressifs. Ils peuvent exprimer le récepteur à la progestérone PR, et seul le type B exprime également le récepteur HER2 (Human Epidermal growth factor Receptor 2). Les cancers de type luminal B sont alors plus agressifs que le luminal A, de par une plus grande expression de gènes impliqués dans la prolifération cellulaire [8].

Dans le cas de cancers HER2-positif, les récepteurs ERα et PR ne sont pas exprimés. La sur-expression du récepteur HER2 active différentes voies de signalisation telles que les voies RAS/MAPK ou PI3K/AKT, résultant en une augmentation de la prolifération et survie cellulaires. Ceci favorisant le développement de métastases, ce type de cancer du sein est plus agressif que les cancers luminaux [1].

Les cancers du sein triple négatif (n'exprimant aucun des récepteurs ERα, PR, HER2) sont les plus agressifs et les plus hétérogènes. Ils peuvent alors être divisés en plusieurs sous-catégories selon des caractéristiques telles que l'expression de gènes clefs, la présence de certaines mutations, ou encore la composition du microenvironnement tumoral. Différentes classifications ont ainsi été publiées mais s'accordent à mettre en évidence un phénotype principal qui est le basal-like. Ce type de cancer du sein triple négatif est caractérisé par des marqueurs de prolifération cellulaire et d'altération de la réparation de l'ADN, tels que l'amplification de MYC, CDK6, CCNE1 ; la délétion de BRCA2, PTEN, MDM2 ; ou encore un fort taux de mutation de TP53 [9].

Comme pour tout cancer, les mécanismes biologiques menant au développement du cancer du sein sont complexes. La dérégulation de nombreux acteurs cellulaires entre en effet en jeu pour mener à la prolifération et survie non contrôlée des cellules cancéreuses. Ainsi peuvent être combinés [10]:

- la surexpression d'oncogènes tels que MYC : un facteur de transcription impliqué dans l'expression de nombreux gènes de prolifération et survie cellulaire. Son expression est augmentée dans de nombreux cancers en plus du cancer du sein, via une augmentation de sa transcription ou de sa stabilité notamment [11,12].
- des pertes d'expression ou fonction de gènes suppresseur de tumeur tels que :
    - BRCA1/2

Les gènes BRCA1 et BRCA2 (Breast Cancer Associated gene 1 et 2) codent pour des protéines impliquées notamment dans la réparation par recombinaison homologue des cassures double-brin de l'ADN (Figure 3). La mutation de ces gènes altère alors ces mécanismes de réparation de l'ADN, ce qui augmente les risques d'acquisition de mutations et participe au développement de différents cancers, en particulier ceux du sein et de l'ovaire [13].

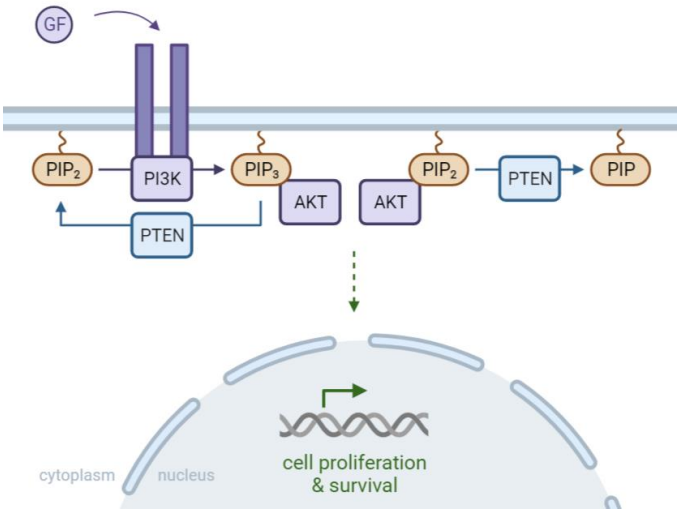


Figure 4 : Régulation de la voie PI3K/AKT par PTEN.

La stimulation de récepteurs aux facteurs de croissance (EGFR, IGFR...) active PI3K qui va phosphoryler PIP_2 en PIP_3. PIP_3 peut alors recruter AKT à la membrane et mener à son activation, favorisant la prolifération et survie cellulaires. La phosphatase PTEN inhibe cette voie de signalisation en déphosphorylant PIP_3 en PIP_2, puis PIP_2 en PIP ; PIP_2 pouvant également interagir avec AKT. Figure adaptée de Csolle et al. [14]. GF : facteur de croissance ; PIP : phosphatidylinositol phosphate.

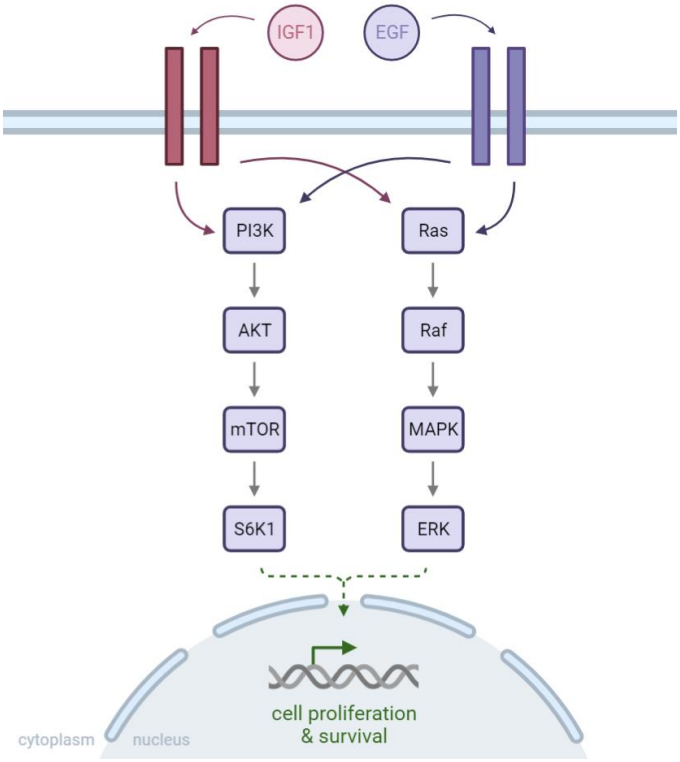


Figure 5 : Principales voies de signalisation activées par les récepteurs à tyrosine kinase IGF1R et EGFR, favorisant la prolifération et survie cellulaires.

Figure adaptée de Christopoulos et al. et Steelman et al. [15,16].

- TP53

La protéine p53 codée par le gène TP53 est quant à elle un facteur de transcription activant l'expression de gènes impliqués dans l'arrêt du cycle cellulaire ou l'apoptose, en réponse à une variété de stress cellulaires (dommages à l'ADN, hypoxie, stress oxydatif…). Cette activité antiproliférative est alors primordiale pour empêcher le développement de cellules cancéreuses, expliquant pourquoi la mutation de TP53 joue un rôle important dans de nombreux cancers dont le cancer du sein [17].

- PTEN

Le gène PTEN (Phosphatase and Tensin homolog) code pour une phosphatase qui participe à l'inhibition de la voie PI3K/AKT (Figure 4). La perte d'expression ou d'activité de PTEN participe alors à l'activation dérégulée de cette voie de signalisation, qui joue un rôle pro-prolifératif et métastatique prépondérant dans le développement du cancer du sein [14].

- la surexpression ou activation de récepteurs à tyrosine kinase et leurs voies de signalisation tels que :
  - EGFR1 ou HER2

La famille des récepteurs EGFR (Epidermal Growth Factor Receptor) désigne des récepteurs à tyrosine kinase dont l'activation induit de nombreuses voies de signalisation telles que RAS/MAPK, PI3K/AKT ou encore PLC/PKC, impliquées dans la régulation de la prolifération cellulaire (Figure 5). Au sein de cette famille, la surexpression de EGFR1 et HER2 notamment joue alors un rôle dans le développement du cancer du sein [16].

- IGF1R

IGF1R (Insulin like Growth Factor 1 Receptor) est le récepteur de la protéine IGF1 qui joue un rôle important dans le développement et la fonctionnalité de la glande mammaire. La liaison d'IGF1 à son récepteur entraîne l'activation d'IGF1R par phosphorylation, qui va activer les voies de signalisation PI3K/AKT et RAS/MAPK ayant une activité proliférative et anti-apoptotique (Figure 5). Une suractivation de cet axe de signalisation peut notamment être provoquée par une augmentation du taux d'IGF1 circulant ou une surexpression d'IGF1R, participant à une prolifération et survie dérégulées des cellules cancéreuses mammaires [15].

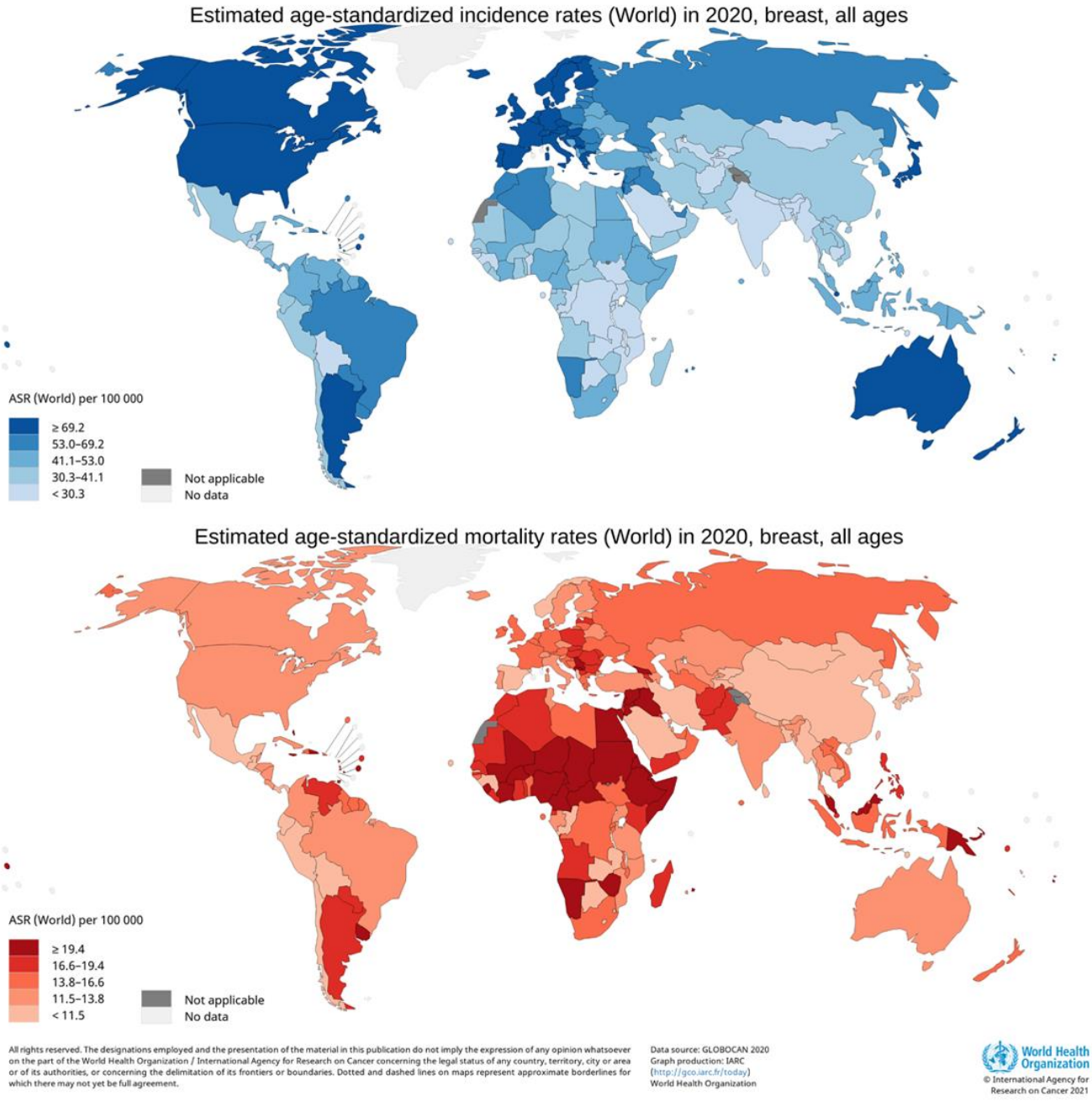


Figure 6 : Taux d'incidence (bleu) et de mortalité (rouge) du cancer du sein au niveau mondial, en 2020.

Alors que l'incidence est la plus forte dans les pays développés, le taux de mortalité y est au contraire plus faible que dans les pays en développement. Figures issues du Global Cancer Observatory (https://gco.iarc.fr/today/home, consulté le 24 janvier 2022).

Ces différentes altérations de mécanismes biologiques menant au développement du cancer du sein peuvent être provoquées par un certain nombre de facteurs de risques, résumés en Figure 7. Ceux-ci sont aussi bien liés à la génétique qu'au mode de vie et à l'environnement [8,18–21].

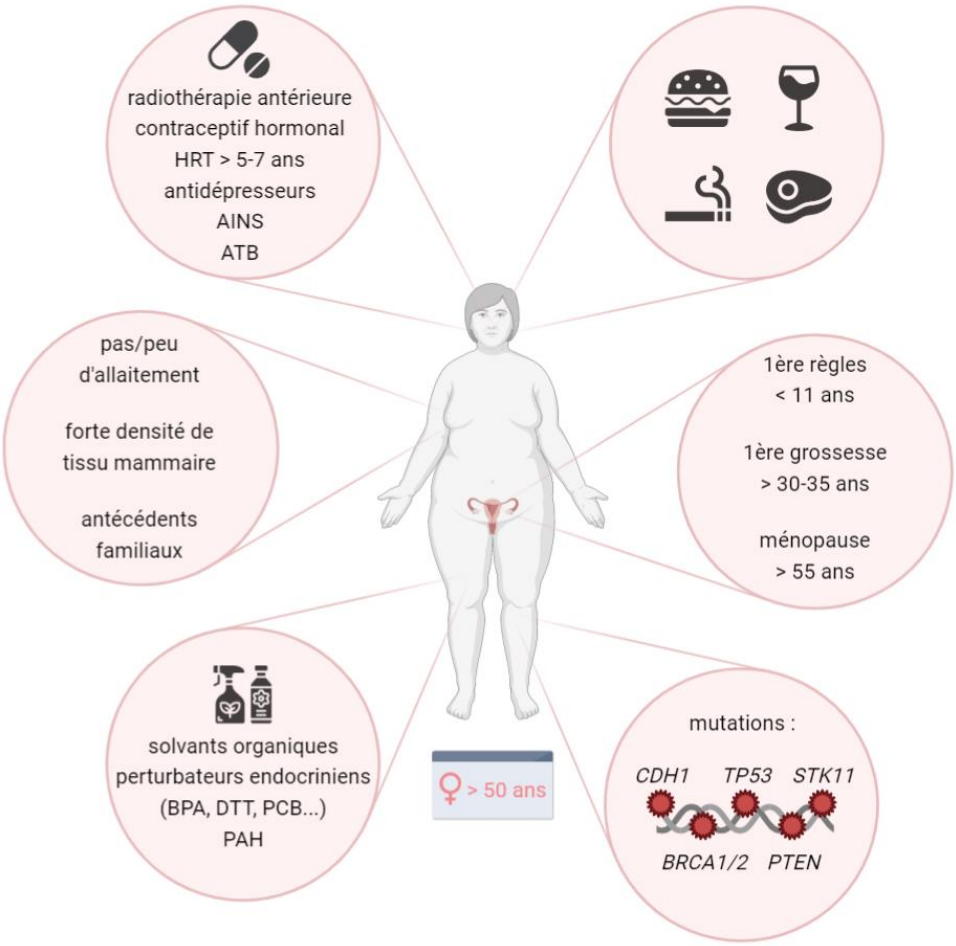


Figure 7 : Facteurs de risques de développer un cancer du sein.

AINS : anti-inflammatoires non stéroïdiens ; ATB : antibiotiques ; BPA : bisphénol A ; DTT : dichlorodiphényltrichloroéthane ; HRT : traitement hormonal de la ménopause ; PAH : hydrocarbures aromatiques polycycliques ; PCB : biphényles polychlorés.

Ces facteurs de risques sont alors plus présents au sein des pays développés, où la première grossesse est plus tardive, l'allaitement moins répandu, les traitements hormonaux plus fréquemment utilisés et le mode de vie plus sédentaire. Ceci explique que le taux d'incidence du cancer du sein soit plus élevé que dans les pays en développement, où les campagnes de dépistage sont par ailleurs moins répandues (Figure 6). L'évolution des modes de vie dans ces pays en développement et en Asie mène cependant à une augmentation de l'incidence du cancer du sein, due aux facteurs de risques mentionnés ci-dessus. La mortalité due au cancer du sein augmente alors particulièrement dans les pays en développement, où la prise en charge des patientes est plus tardive et moins efficace que dans les pays développés (Figure 6). La mise en place de campagnes de prévention et dépistage précoce est donc particulièrement importante pour contrer la progression de ce cancer, dans les pays en développement notamment [2].

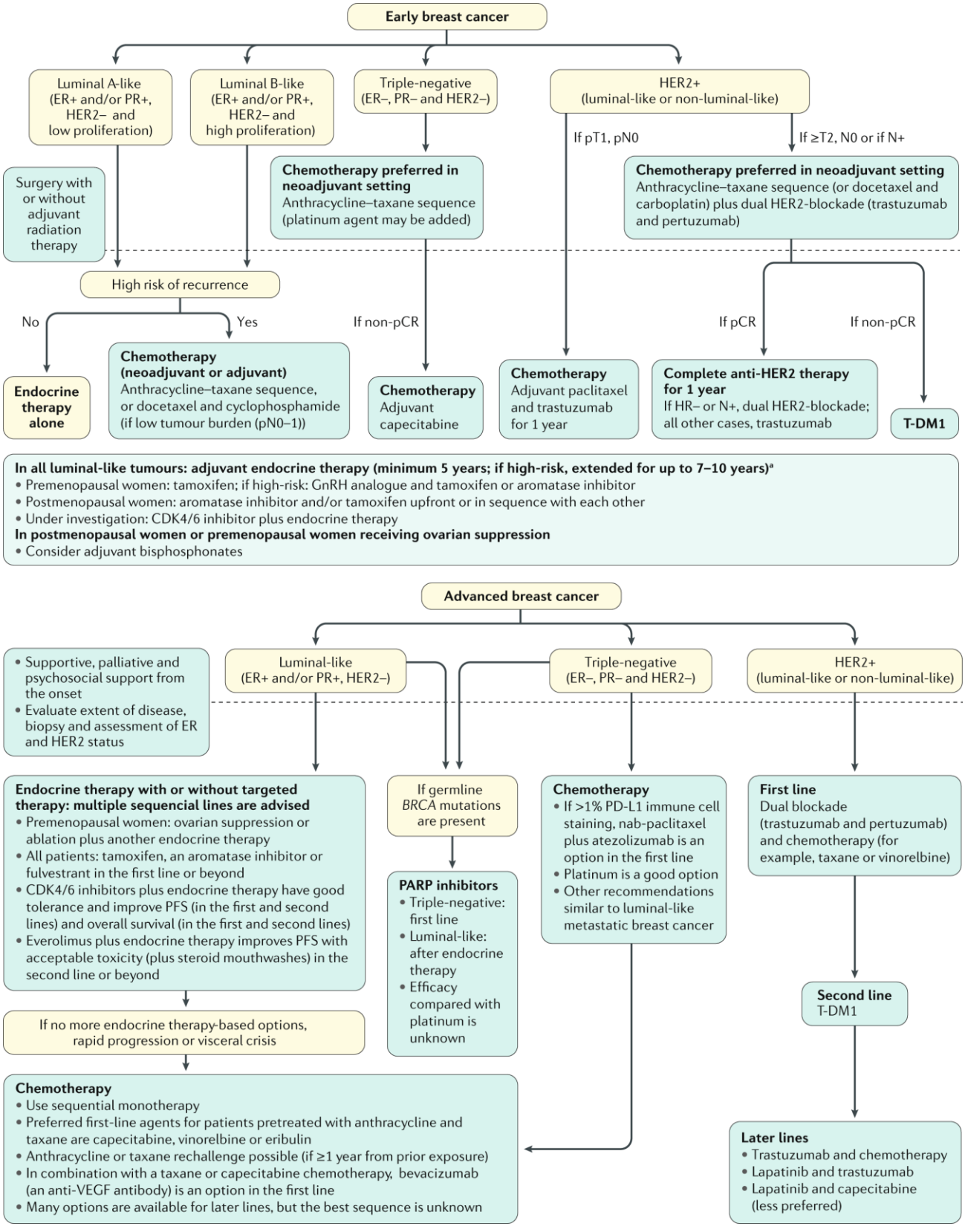


Figure 8 : Stratégies thérapeutiques en cas de cancer du sein.

Figures issue de Harbeck et al. [1]. ER : estrogen receptor ; GnRH : gonadotropin- releasing hormone ; HER2 : human epidermal growth factor receptor 2 ; N : node status ; p : pathological ; PARP : poly(ADP-ribose) polymerase ; pCR : pathological complete response ; PD-L1 : programmed death-ligand 1 ; PFS : progression-free survival ; PR : progesterone receptor ; T : tumour grade ; T-DM1 : ado-trastuzumab emtansine ; VEGF : vascular endothelial growth factor.

En cas de suspicion de cancer du sein, le diagnostic est réalisé en combinant l'examen clinique, l'imagerie et l'analyse de biopsies. Après avoir déterminé le type de cancer du sein et son stade de développement, différentes options thérapeutiques existent et sont combinées en fonction du profil de chaque patiente, comme indiqué en Figure 8 [1] :

- chirurgie : ablation de la tumeur en préservant au maximum la glande mammaire (75% des cas), ou ablation totale de la glande mammaire, selon l'étendue du cancer ;
- radiothérapie : irradiation ciblée visant à détruire d'éventuelles cellules tumorales résiduelles suite à un traitement chirurgical ;
- chimiothérapie : traitement non ciblé menant à la destruction des cellules en division, pouvant être réalisé préalablement à une opération afin de réduire la taille de la tumeur, et/ou après l'opération si le risque de rechute est élevé ;
- hormonothérapie : traitement spécifique des cancers du sein luminaux ;
- autres traitements spécifiques : ciblant HER2 ou d'autres voies de signalisation cellulaires.

Le suivi des patientes (généralement par imagerie) est alors primordial pour s'assurer de la réussite du traitement choisi, ajuster la prise en charge thérapeutique, puis vérifier l'absence de rechute.

Afin de sélectionner au mieux le traitement adéquat à chaque patiente selon l'évolution de son cancer, une autre stratégie développée ces dernières années consiste en l'analyse de l'ADN tumoral circulant. Cette méthode non invasive permet la détection de mutations oncogéniques, dont la présence peut orienter la prise en charge thérapeutique vers le choix de traitements spécifiques [22]. Des cellules tumorales circulantes peuvent également être détectées dans le sang, et servir de biomarqueur attestant de l'efficacité ou non du traitement d'un cancer métastatique [23]. L'intérêt de recourir à ces approches nécessite cependant d'être validé par des essais cliniques supplémentaires avant de pouvoir être appliqué en routine.

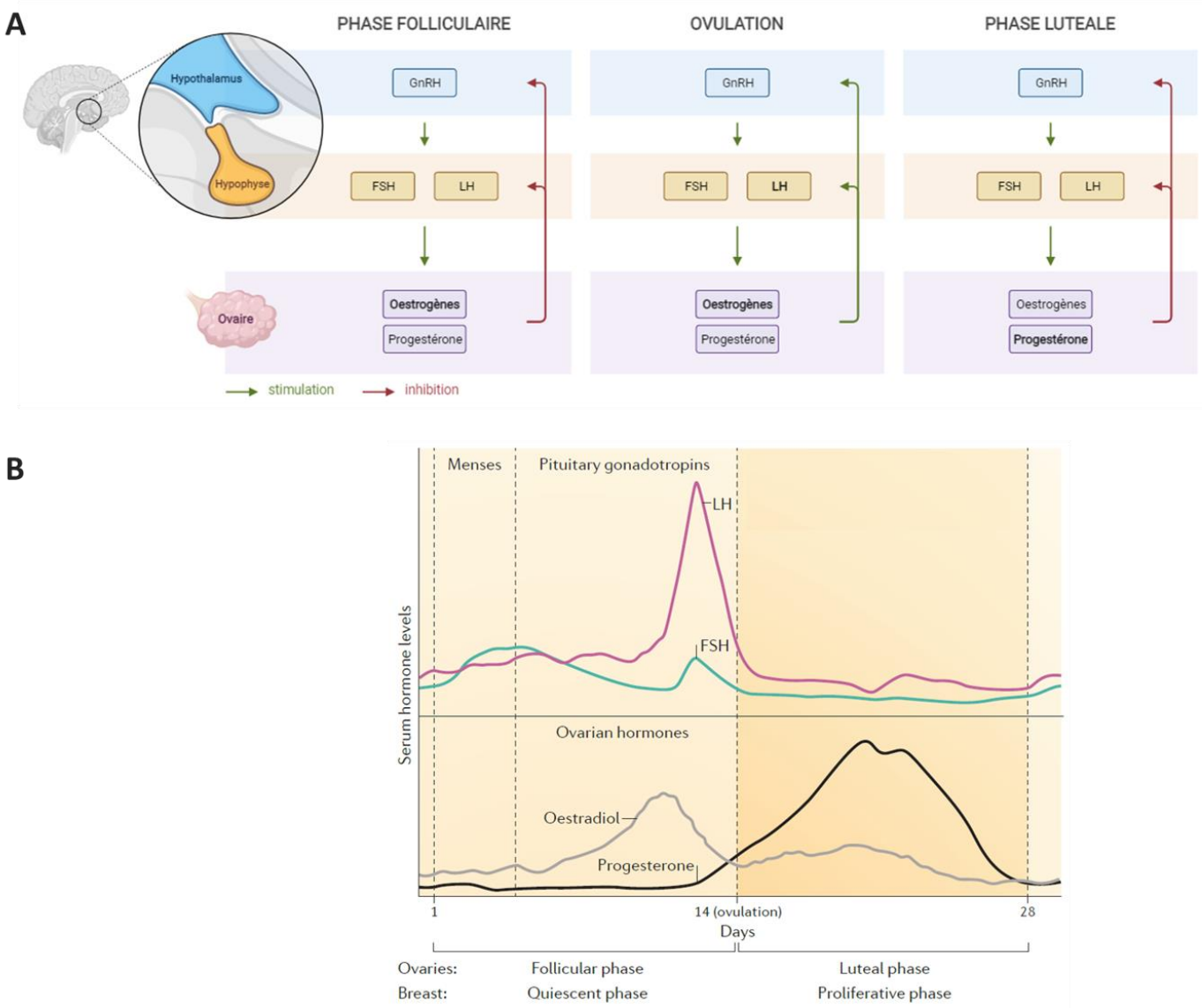


Figure 9 : Régulation hormonale des cycles menstruels chez la femme.

A) Les hormones de l'axe hypothalamo-hypophysaire stimulent la production d'œstrogènes et de progestérone par les follicules ovariens, qui exercent en retour un rétrocontrôle négatif sur la production de GnRH, FSH et LH. Lorsque la quantité d'œstrogènes atteint un certain seuil, en fin de phase folliculaire, les hormones stéroïdiennes stimulent au contraire l'axe hypothalamo-hypophysaire, résultant en un pic de LH qui déclenche l'ovulation. En phase lutéale, la production de GnRH, FSH et LH est de nouveau inhibée par les œstrogènes et la progestérone, cette dernière devenant l'hormone majoritairement produite par les ovaires (plus précisément par le corps jaune) [24]. B) L'augmentation progressive du niveau d'œstrogènes en phase folliculaire aboutit à un pic de LH permettant l'ovulation, suite à laquelle le taux de progestérone devient majoritaire. Figure issue et modifiée de Brisken [25]. FSH : follicle-stimulating hormone ; GnRH : gonadotropin-releasing hormone ; LH : luteinizing hormone.

b) Cancer du sein ERα+

Dans 70% des cas, le cancer du sein est de type luminal, donc caractérisé par l'expression du récepteur aux œstrogènes ERα. La prévalence de ce type de cancer du sein est liée à l'imprégnation hormonale du tissu mammaire. En effet, le développement de la glande mammaire est sous contrôle de deux principales hormones stéroïdiennes qui sont les œstrogènes et la progestérone. Ces hormones sont produites par les ovaires, en réponse à la stimulation par l'axe hypothalamo-hypophysaire (Figure 9, A) [26,27]. Les œstrogènes stimulent le développement mammaire lors de la puberté et de la grossesse essentiellement, via l'élongation et la ramification des canaux galactophores, respectivement. Au cours des cycles menstruels, les œstrogènes stimulent la prolifération des cellules mammaires de manière plus indirecte : le taux d'œstrogènes est d'abord élevé en phase folliculaire (pré-ovulatoire), permettant l'expression du récepteur à la progestérone dans les cellules épithéliales mammaires. Après l'ovulation, en phase lutéale, c'est le taux de progestérone qui devient supérieur à celui des œstrogènes. Cette hormone stimule alors la prolifération des cellules mammaires via l'activation de PR, au niveau des lobules terminaux de la glande mammaire (Figure 9, B) [25,28,29].

ERα et PR ne sont exprimés que par 30 à 50% des cellules luminales. Il apparaît que ces cellules prolifèrent peu, contrairement aux cellules environnantes. Les cellules ERα et/ou PR-positives régulent en effet la prolifération de l'épithélium mammaire via une communication paracrine. L'activation des récepteurs ERα et PR permet alors l'expression de facteurs tels que l'amphiréguline et IGF1, ou WNT4 et RANKL (Receptor Activator of Nuclear factor-κB Ligand), respectivement, qui seront sécrétés pour activer la prolifération des cellules proches (Figure 10) [25]. L'activation de ces voies de signalisation présente cependant un potentiel oncogénique, qui augmente le risque de provoquer des dommages à l'ADN et la survenue de mutations au cours de la vie des femmes, de par l'accumulation des cycles menstruels. Les cellules cancéreuses continuent ensuite à proliférer en réponse à la stimulation par les œstrogènes [1].

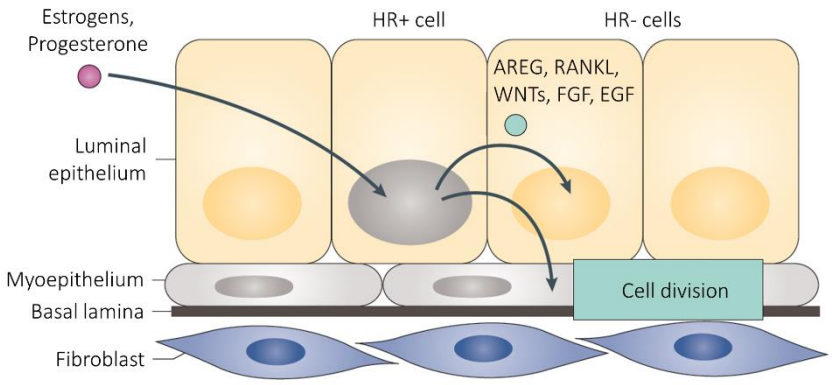


Figure 10 : Régulation paracrine de la prolifération cellulaire de l'épithélium mammaire.
Seules certaines cellules luminales expriment les récepteurs aux œstrogènes et à la progestérone (HR+ cell). Après stimulation par ces hormones stéroïdiennes, ces cellules sécrètent un panel de facteurs qui vont induire la prolifération des cellules environnantes. Figure issue et modifiée de Brisken [25].

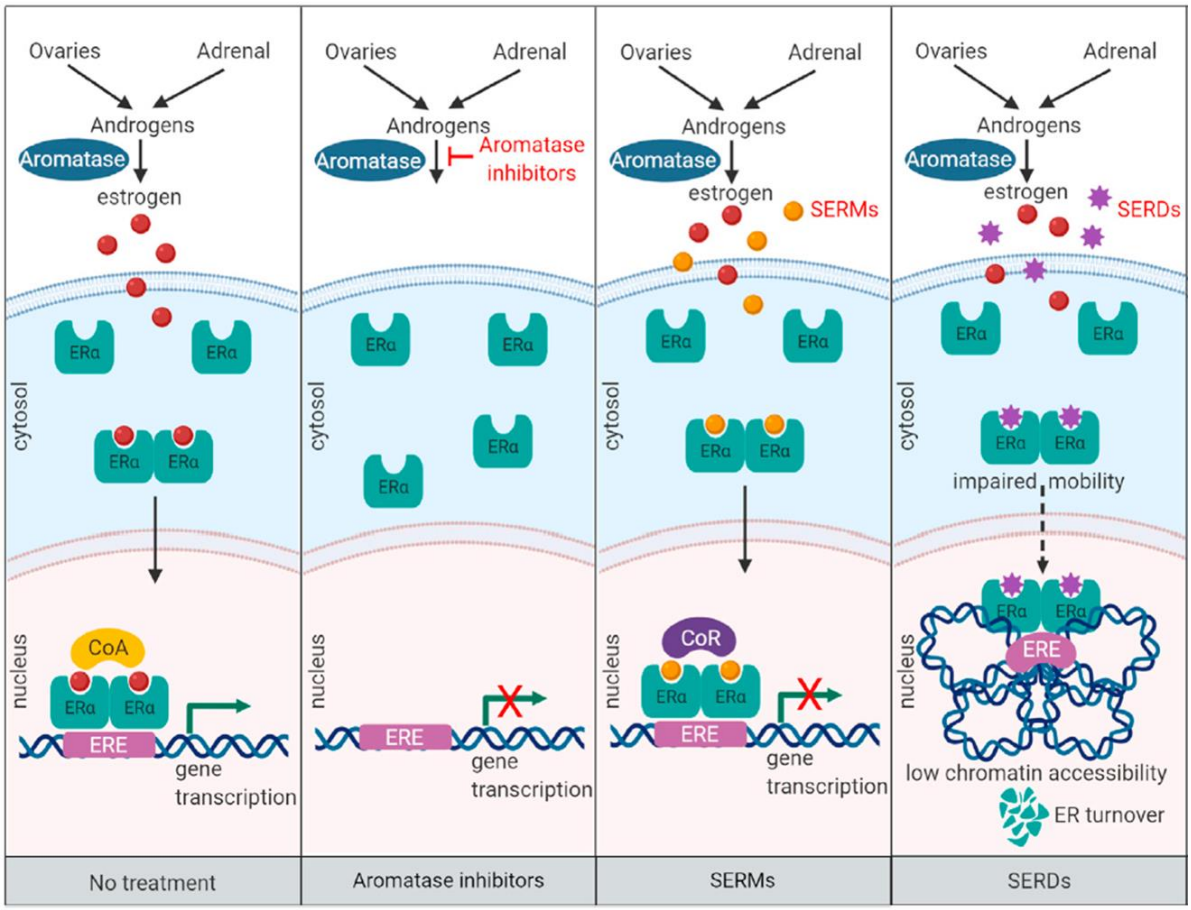


Figure 11 : Mécanismes d'action des principaux traitements hormonaux.

Tandis que les inhibiteurs d'aromatase visent à priver les cellules en œstrogènes, les SERMs et SERDs interagissent avec ERα. Les SERMs permettent d'inhiber l'activité du récepteur en favorisant le recrutement de corépresseurs, tandis que les SERDs provoquent la dégradation d'ERα. Figure issue de [30]. CoA : coactivator ; CoR : corepressor ; ERα : estrogen receptor α ; ERE : estrogen responsive element ; SERD : selective estrogen receptor downregulator ; SERM : selective estrogen receptor modulator.

La prolifération des cellules cancéreuses étant dépendante de l'activation d'ERα par les œstrogènes, le traitement de choix est alors l'hormonothérapie. Cette thérapie consiste à priver la tumeur en œstrogènes, ou bloquer l'activité d'ERα. La privation en œstrogènes est plutôt réalisée chez les femmes ménopausées, via l'utilisation d'inhibiteurs de l'aromatase (AIs), voire l'ablation des ovaires [31]. L'activité d'ERα peut quant à elle être bloquée par l'utilisation de SERMs (Selective Estrogen Receptor Modulator) ou SERDs (Selective Estrogen Receptor Downregulator). Les SERMs (le plus commun étant le Tamoxifène) sont des inhibiteurs compétitifs des œstrogènes, dont la fixation au récepteur ERα favorise la répression de son activité dans les cellules mammaires. Dans d'autres tissus comme les os ou l'utérus en revanche, les SERMs peuvent induire une activité partiellement agonistique, limitant les risques d'ostéoporose mais participant au risque de développer un cancer de l'utérus par exemple. La fixation des SERDs à ERα quant à elle induit la dégradation du récepteur, empêchant alors toute signalisation médiée par ERα (Figure 11) [32]. Le seul SERD autorisé pour le moment est le Fulvestrant mais, son administration devant être réalisée par injection intramusculaire, de nouveaux SERDs pouvant être prescrits par voie orale sont en cours de développement et d'essai clinique [33]. L'utilisation du Fulvestrant est alors généralement réservée aux cancers du sein luminaux métastatiques, ou aux patientes devenues résistantes à un premier traitement par AIs ou SERMs [32].

Dans 30% des cas de cancers du sein ERα+, les cellules tumorales mammaires peuvent en effet échapper au contrôle hormonal, c'est-à-dire acquérir la capacité à proliférer en absence de stimulation par les œstrogènes. L'hormonothérapie n'est alors plus efficace pour contrer le développement cancéreux. Une telle résistance est rarement due à une perte d'expression d'ERα (10% des cas seulement), mais plutôt liée à une dérégulation de l'activation de ce récepteur, qui peut être stimulé indépendamment de la liaison aux œstrogènes. Différents mécanismes peuvent mener à cette activité ligand-indépendante, les principaux étant [30] :

- L'acquisition de mutations rendant le récepteur constitutivement actif ;
- Une augmentation de l'interaction d'ERα avec des coactivateurs, au détriment de ses corépresseurs ;
- L'activation d'ERα par des voies de signalisation intracellulaires oncogéniques (PI3K/AKT, RAS/MAPK) elles-mêmes dérégulées dans les cellules cancéreuses.

Pour les patientes devenues résistantes, une piste thérapeutique est alors de cibler les voies de signalisation impliquées dans l'activation ligand-indépendante d'ERα. Il est ainsi possible de combiner l'hormonothérapie avec des inhibiteurs de CDK4/6 (tel que le Palbociclib) ou de la voie PI3K (Alpelisib, Evérolimus), des essais cliniques étant en cours pour tester l'efficacité de différentes combinaisons thérapeutiques en fonction du profil des patientes (Figure 12, A) [34].

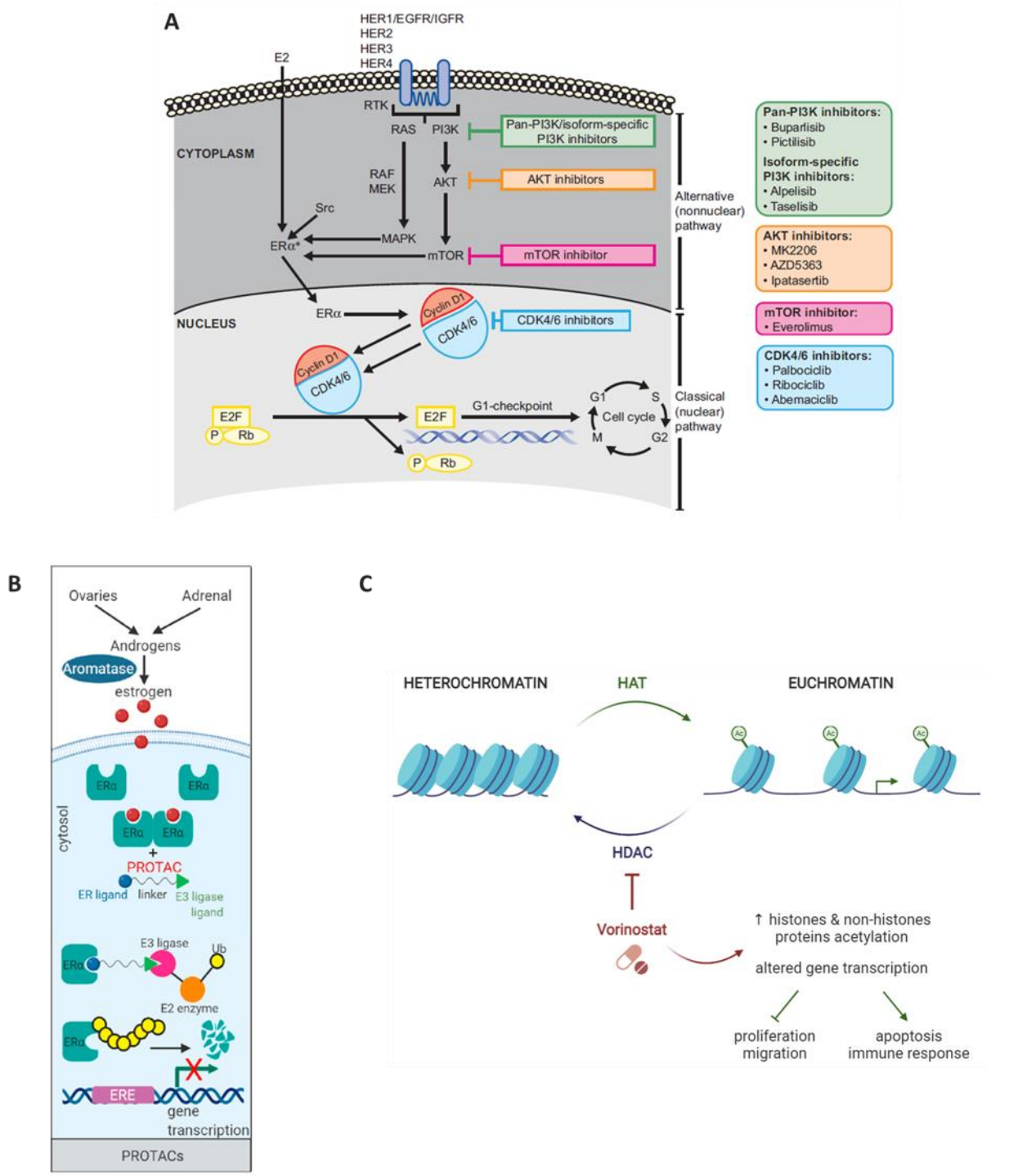


Figure 12 : Mécanismes d'action de nouvelles stratégies thérapeutiques contre le cancer du sein résistant à l'hormonothérapie.

A) Des inhibiteurs de la voie PI3K (PI3K, AKT, mTOR) peuvent être utilisés pour inhiber l'activation ligand-indépendante d'ERα. Des inhibiteurs de CDK4/6 permettent quant à eux d'empêcher l'entrée des cellules dans le cycle cellulaire. Figure issue et modifiée de Brufsky et Dickler [34]. B) Les PROTACs (Proteolysis-Targeting Chimeras) sont des molécules interagissant avec ERα et une ubiquitine ligase E3 afin de provoquer la dégradation du récepteur. Figure issue de Hanker et al. [30]. C) L'utilisation d'inhibiteurs d'histone désacétylases (HDAC) est envisagée afin de promouvoir les programmes transcriptionnelles aboutissant à la mort des cellules cancéreuses. Figure adaptée de Herrera-Martínez et al. [35].

Un nouveau type d'agent thérapeutique a également été développé pour permettre la dégradation d'ERα malgré l'acquisition de potentielles mutations réduisant la liaison de SERMs ou SERDs. Il s'agit de molécules nommées PROTAC (Proteolysis-Targeting Chimera) qui vont interagir à la fois avec ERα et une ubiquitine ligase E3 afin de provoquer la dégradation d'ERα par le protéasome (Figure 12, B). Un essai clinique avec une telle molécule (ARV-471) est actuellement en cours [36].

Un autre mécanisme qui participerait au développement de résistance à l'hormonothérapie serait lié aux modifications épigénétiques, altérant in fine l'expression d'acteurs clefs dans l'échappement au contrôle hormonal. Des molécules ciblant cet axe de régulation (essentiellement des inhibiteurs d'histone désacétylases, tel que le Vorinostat, Figure 12, C) sont alors également en essai clinique mais des développements supplémentaires semblent nécessaires pour améliorer cette option thérapeutique [37,38].

En cas de résistance à l'hormonothérapie, l'utilisation d'autres thérapies ciblées peut elle aussi mener à des phénomènes de résistance. Dans ce cas, la chimiothérapie peut être utilisée, mais le recours à de tels traitements diminue grandement la qualité de vie des patientes [1].

Touchant de plus en plus de femmes chaque année, le cancer du sein reste un enjeu de santé publique majeur. Dans la majorité des cas, son développement est lié au récepteur aux œstrogènes ERα et, bien que des traitements ciblés existent pour contrer ce type de cancer, des phénomènes de résistance ont fréquemment lieu. Comprendre les mécanismes permettant aux cellules tumorales mammaires d'échapper au contrôle hormonal est alors nécessaire pour mieux prendre en charge ces patientes, et permettre le développement de nouvelles options thérapeutiques.

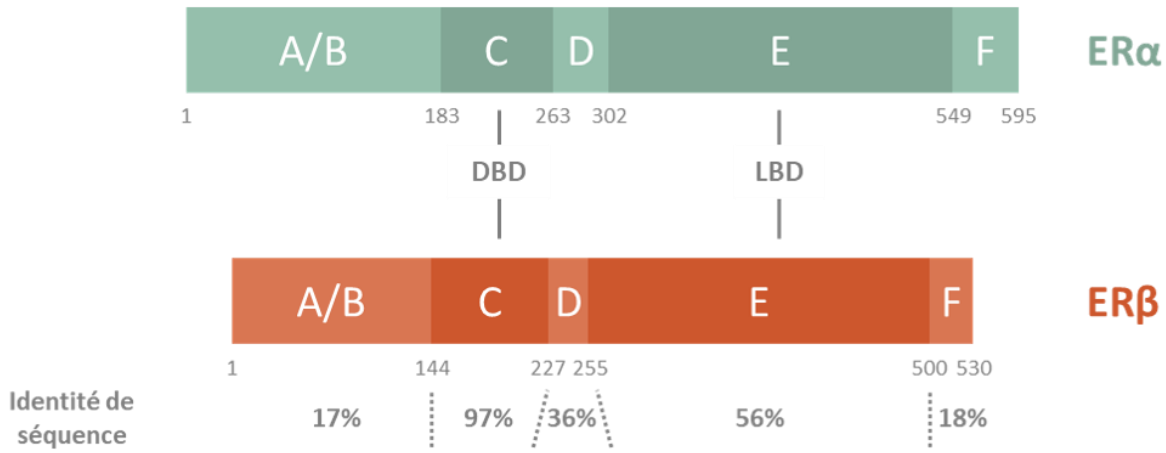


Figure 13 : Comparaison de l'organisation des récepteurs ERα et ERβ.

Les domaines A à F composant chaque récepteur sont représentés, avec les acides aminés numérotés sous chaque séquence. Le pourcentage d'identité entre les deux récepteurs est indiqué pour chaque domaine, d'après Mal et al. [39]. DBD : DNA-binding domain ; LBD : ligand-binding domain.

2. LE RECEPTEUR AUX ŒSTROGENES ERα

Le récepteur aux œstrogènes ERα est un récepteur nucléaire, dont l'activation mène à la régulation de la transcription de nombreux gènes impliqués dans la prolifération cellulaire. La dérégulation de son expression et/ou activité joue alors un rôle prépondérant dans le développement de la majorité des cancers du sein, dit ERα+ ou luminaux. La signalisation médiée par ERα joue également un rôle dans la physiologie non-reproductive qui ne sera pas détaillée dans ce manuscrit, de par l'expression du récepteur dans d'autres tissus tels que les os [40], le cerveau [41,42], le tissu adipeux [43], ou le système vasculaire [44].

Il est à noter que deux autres récepteurs peuvent être stimulés par les œstrogènes : le récepteur aux œstrogènes ERβ, et le récepteur couplé aux protéines G GPR30.

ERβ est un récepteur nucléaire homologue d'ERα, codé par le gène ESR2 et présentant une structure proche d'ERα (Figure 13). Comme ERα, ERβ est exprimé dans de nombreux tissus en plus des cellules épithéliales mammaires et des organes reproductifs, tels que le tissu adipeux ou le cerveau. Ce récepteur partage certains ligands avec ERα, tels que les œstrogènes et les SERMs, mais avec une affinité différente. Contrairement à ERα, l'activation d'ERβ résulte généralement en une inhibition de la prolifération et l'induction de l'apoptose, mais ces effets dépendent du tissu étudié et de la co-expression ou non d'ERα. L'expression d'ERβ dans les cellules tumorales mammaires a alors tendance à être corrélée à un pronostic favorable, mais certaines études indiquent le contraire. Des travaux supplémentaires restent ainsi nécessaires pour mieux comprendre le rôle physiologique de ce récepteur et son implication dans le cancer du sein [45,46].

GPR30 est lui aussi exprimé dans différents organes tels que le foie et le tissu adipeux en plus de la glande mammaire. Sa stimulation par les œstrogènes induit la production d'AMPc et l'activation de cascades de signalisation telles que PI3K/AKT et RAS/MAPK, régulant in fine la transcription de gènes de prolifération et survie cellulaire. L'activation de ce récepteur semblerait alors avoir un effet pro-tumoral et pourrait jouer un rôle dans l'acquisition de résistance à l'hormonothérapie, les SERMs et SERD agissant comme des agonistes de ce récepteur. Comme pour ERβ, des travaux supplémentaires sont nécessaires pour mieux identifier les fonctions de ce récepteur, voire l'envisager comme une cible thérapeutique dans le cancer du sein [47].

ERα reste ainsi le récepteur aux œstrogènes le mieux caractérisé, et le plus étudié lorsque l'on s'intéresse au cancer du sein. Les récepteurs ERβ et GPR30 ne seront alors pas plus détaillés dans la suite de ce manuscrit.

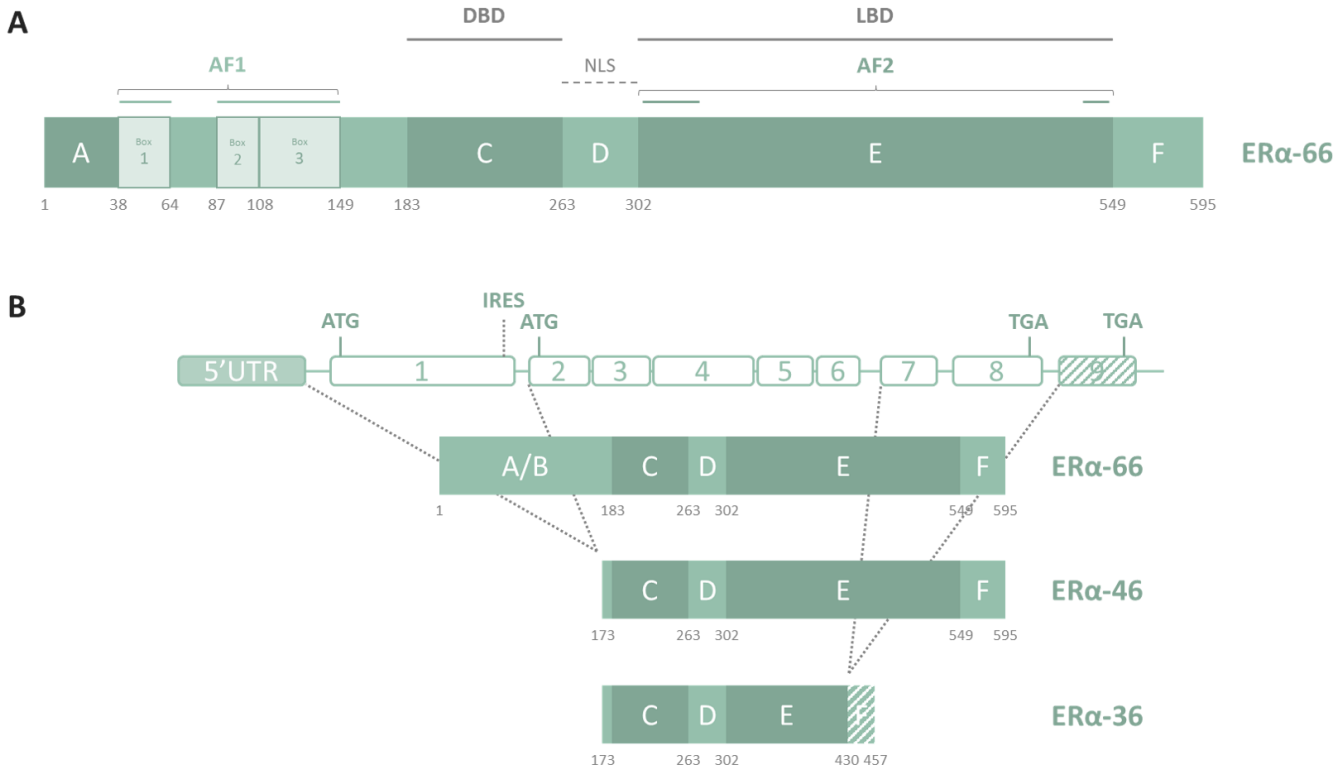


Figure 14 : Structures du récepteur ERα et de ses isoformes.

A) Représentation détaillée des domaines composants ERα, dont les acides aminés sont numérotés en gris. La fonction de transactivation 1 (AF1) est constituée de trois boîtes fonctionnelles au sein du domaine B du récepteur. Le domaine C contient le domaine de liaison à l'ADN (DBD), un signal de localisation nucléaire est situé au sein du domaine D, et le domaine E contient le domaine de liaison au ligand (LBD) ainsi que la seconde fonction de transactivation (AF2). Celle-ci est formée lors du rapprochement tridimensionnel des acides aminés présents dans les premières et la dernière hélice du domaine E. B) Représentation schématique du gène ESR1 codant les isoformes naturels d'ERα. La forme entière, de 66 kDa, est codée par les exons 1 à 8. L'isoforme de 46 kDa est obtenu suite à l'épissage de l'exon 1 ou l'initiation de la traduction via un IRES contenu dans l'exon 1, et manque ainsi les domaines A et B du récepteur. L'isoforme le plus court, de 36 kDa, est quant à lui obtenu lorsque la transcription est initiée via un promoteur alternatif dans le premier intron du gène. Le récepteur ainsi produit manque également les domaines A et B, les exons 7 et 8 ne sont pas transcrits, et un domaine F alternatif est codé par l'exon 9. Figure adaptée de Rusidzé et al. [29]. UTR : untranslated region ; IRES : internal ribosome entry site ; ATG : codons d'initiation de la transcription ; TGA : codons de terminaison de la transcription.

a) Structure, variants et mutations d'ERα

Codé par le gène ESR1, ERα est une protéine de 66kDa, composée de 595 acides aminés formant 6 domaines structurels, nommés A à F (Figure 14, A).

ERα peut lier ses ligands (œstrogènes, SERMs, SERD) grâce à un domaine de liaison au ligand (LBD) localisé au niveau du domaine E. La structure de cette région C-terminale a pu être déterminée : elle serait formée de douze hélices α et deux feuillets β, formant une poche hydrophobe pour accueillir le ligand. La liaison d'un ligand induit alors un réarrangement des hélices primordial dans l'activité du récepteur, notamment via le positionnement de l'hélice 12. La structure adoptée permet alors l'interaction avec des coactivateurs ou au contraire des corépresseurs. Ce domaine E serait également impliqué dans la dimérisation du récepteur, via les hélices 8 à 11 ; cette dimérisation étant ensuite stabilisée par le DBD des récepteurs [48].

Etant un récepteur nucléaire, ERα contient en effet un domaine de liaison à l'ADN (DBD) constitué de deux doigts de zinc, permettant son interaction directe avec l'ADN. Cette liaison a lieu entre le domaine C de la protéine, et des séquences dites ERE (Estrogen Responsive Element) suivant le motif : GGTCAnnnTGACC [48].

Le domaine D est quant à lui une région charnière conférant une certaine flexibilité structurelle entre les parties N- et C-terminales de la protéine. Cela permet par exemple au domaine A d'interagir avec la partie C-terminale du récepteur pour en réprimer l'activité en absence de ligand [29]. Le domaine D contient également un signal de localisation nucléaire (NLS) pour l'adressage d'ERα au noyau et est le site d'interaction avec d'autres protéines telles que les facteurs de transcription c-jun ou SP1. Aussi, ce domaine contient des résidus pouvant être modifiés de manière post-traductionnelle, jouant un rôle dans l'activité du récepteur [48].

Enfin, l'activité transcriptionnelle d'ERα repose sur deux régions distinctes : les domaines A et B pour la fonction AF1, le domaine E pour la fonction AF2. La fonction AF1 est constituée de 3 boîtes fonctionnelles parmi lesquelles se situent des résidus dont la phosphorylation permet le recrutement de coactivateurs et l'activation d'ERα en absence de son ligand. Aucune structure tridimensionnelle n'a pu être déterminée pour cette partie N-terminale qui apparaît structurellement désordonnée. La fonction AF2, localisée au domaine E, repose quant à elle sur le rapprochement tridimensionnel d'acides aminés présents dans les hélices 1, 3 et 12 [48].

ERα peut également être retrouvé dans les cellules mammaires sous la forme de deux isoformes plus courts, de 46 ou 36kDa (Figure 14, B). ERα-46 serait obtenu par épissage alternatif de l'exon 1 d'ESR1, ou lors de l'initiation de la traduction via un IRES (Internal Ribosome Entry Site) à un ATG présent dans le domaine B. La protéine produite manque ainsi les 173 premiers acides aminés, correspondant à la fonction AF1. ERα-36 serait quant à lui obtenu suite à la transcription d'ESR1 via un promoteur alternatif situé dans le premier intron du gène, et manque les exons 7 et 8.

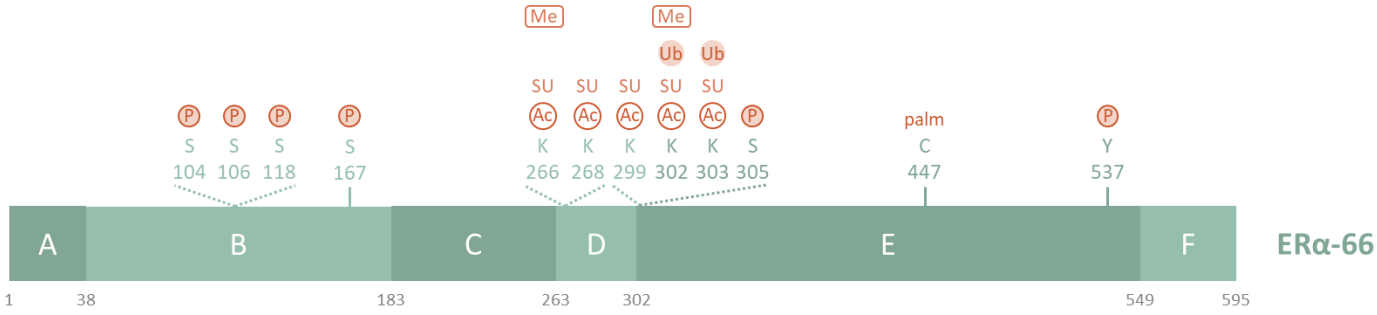


Figure 15 : Représentation schématique des principales modifications post-traductionnelles d'ERα. Les acides aminés modifiés sont indiqués en vert, et les différentes modifications possibles en orange. Les acides aminés bordant les domaines A à F d'ERα sont numérotés en gris. Figure adaptée de Le Romancer et al. et Jeffreys et al. [49,50]. Ac : acétylation ; Me : méthylation ; P : phosphorylation ; palm : palmitoylation ; SU : sumoylation ; Ub : ubiquitination.

Le récepteur ainsi obtenu ne contient alors plus ses domaines fonctionnels AF1 et AF2, mais conserve ses domaines de liaison à l'ADN et au ligand ainsi que les motifs susceptibles d'être myristoylés. La myristoylation consiste en l'ajout d'un groupe myristoyle à une glycine dans la partie N-terminale des protéines, induisant leur ancrage à la membrane plasmique. ERα-36 serait alors majoritairement localisé à la membrane cellulaire, permettant l'induction d'une activité non-génomique suite à la stimulation par les œstrogènes. ERα-36 pourrait également entrer en compétition avec ERα-66 au niveau de la liaison à l'ADN. Ces deux caractéristiques permettraient à ERα-36 de jouer un rôle dans l'acquisition de résistance à l'hormonothérapie. Le rôle joué par ERα-46 quant à lui n'est pas exactement caractérisé [29].

Comme mentionné précédemment, ERα est la cible de nombreuses modifications post-traductionnelles (PTMs) impactant l'activité, la stabilité et la localisation cellulaire du récepteur. 22 résidus d'ERα ont été identifiés comme étant la cible de PTMs et sont listés dans la revue de Le Romancer [49]. Parmi les modifications les mieux caractérisées (représentées sur la Figure 15), l'on peut citer :

- la phosphorylation de sérines au sein de la fonction AF1 (S104, 106, 118, 167 notamment) ainsi que celle de la sérine 305. Ces phosphorylations sont réalisées par des kinases intracellulaires, suite à l'activation de voies de signalisation de facteurs de croissance (IGF1, EGF), et mènent au recrutement de coactivateurs et l'activation d'ERα en absence de son ligand. Ces phosphorylations semblent alors associées au développement de résistance à l'hormonothérapie lors de cancers du sein ERα+ [49].
  La tyrosine 537 peut également être phosphorylée suite à la liaison d'un ligand à ERα, ce qui stimulerait son activité transcriptionnelle mais favoriserait également une localisation cytoplasmique du récepteur et son activité non-génomique [50].
- Différentes lysines peuvent être acétylées par p300, de manière constitutive (K299, 302, 303) ou suite à la stimulation par les œstrogènes (K266, 268). Ces résidus constitutivement acétylés semblent réprimer l'activité transcriptionnelle d'ERα tandis que ceux dont l'acétylation est induite par les œstrogènes stimuleraient cette activité transcriptionnelle, mais les mécanismes régulant ces PTMs sont encore mal connus [49].
- Alors que la mono-ubiquitination d'ERα semble le stabiliser et jouer un rôle dans la stimulation de son activité transcriptionnelle, la poly-ubiquitination permet sa dégradation par le protéasome. Seules les lysines 302 et 303 ont été identifiées comme étant la cible d'ubiquitination, mais d'autres résidus sont certainement impliqués également.
- Ces lysines présentes dans la région charnière du récepteur peuvent également être sumoylées et ainsi participer à l'activation de l'activité transcriptionnelle d'ERα [49].
- Enfin, la lysine 302 peut également être méthylée, ce qui augmente la stabilité d'ERα en entrant en compétition avec son ubiquitination notamment. D'autres lysines peuvent être méthylées dans cette région charnière, alternativement avec leur acétylation en fonction du niveau d'œstrogènes, menant à l'activation ou répression de l'activité transcriptionnelle d'ERα.

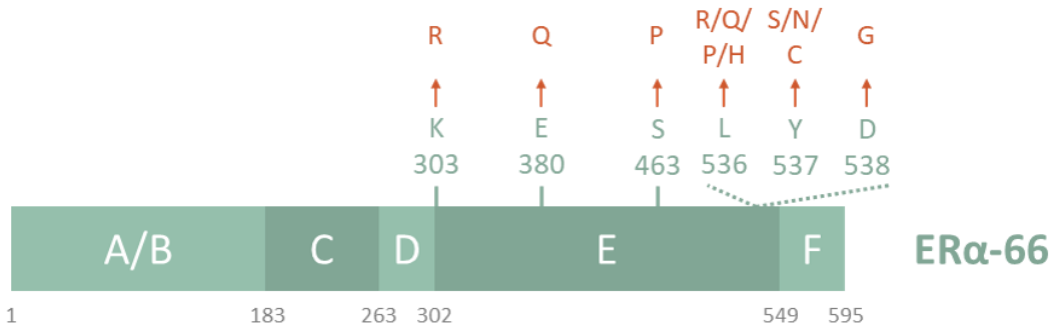


Figure 16 : Représentation schématique des principales mutations d'ERα.

Les acides aminés touchés sont indiqués en vert, et leur substitution en orange. Les acides aminés bordant les domaines A à F d'ERα sont numérotés en gris. Données issues de Clusan et al. [51].

- Récemment, la glycosylation de la sérine 573 a été identifiée comme favorisant la localisation nucléaire du récepteur et son activité génomique [50].
- Une autre PTM importante dans l'activité d'ERα est la palmitoylation de la cystéine 447, qui permet son adressage à la membrane plasmique grâce à l'interaction avec la cavéoline-1. Cette localisation cellulaire permet l'activation d'ERα par les œstrogènes directement à la membrane, provoquant la dépalmitoylation du récepteur et dissociation de la cavéoline-1. ERα libéré interagit alors avec des kinases intracellulaires, résultant en une activité non-génomique (voir partie 2. b) ii.) [49].

La dérégulation de nombreuses voies de signalisation intracellulaires impliquées dans le cancer du sein peut ainsi mener à la dérégulation des PTMs d'ERα. Ces modifications jouant un rôle important dans la stabilité, l'activité et la localisation du récepteur, leur dérégulation est alors impliquée dans le développement du cancer et l'échappement au contrôle hormonal.

Un autre mécanisme menant à l'échappement au contrôle hormonal, en lien avec la structure et activité d'ERα, est l'acquisition de mutations. Il a en effet été montré que des mutations d'ERα étaient fréquemment acquises suite à un traitement par hormonothérapie (20 à 50% des cancers du sein métastatiques). Ces mutations sont majoritairement localisées au sein du LBD, en touchant principalement les acides aminés Y537, D538 et E380, mais aussi les résidus 534 à 536 ou encore S463 (Figure 16). L'impact de ces mutations sur la structure d'ERα n'a été déterminé que pour celles touchant les acides aminés Y537 et D538, qui résultent en une stabilisation de la forme active du récepteur, menant donc à une activité constitutive. D'une manière générale, toutes ces mutations induisent une augmentation de l'activité ligand-indépendante d'ERα, menant à une hausse de la prolifération des cellules tumorales. Les œstrogènes n'étant plus nécessaires à l'activation d'ERα, l'utilisation d'inhibiteurs de l'aromatase n'est alors plus efficace. Le LBD du récepteur étant modifié par ces mutations, la liaison des SERMs s'en retrouve généralement altérée et explique la résistance des cellules tumorales mammaires. L'utilisation du SERD Fulvestrant reste cependant possible [51].

Moins fréquentes, des mutations ont également été recensées en dehors LBD. La principale concerne l'acide aminé K303, dans la région charnière du récepteur qui est la cible de nombreuses PTMs. La mutation K303R provoque en l'occurrence un dérèglement des PTMs des résidus environnants, menant là aussi à une augmentation de l'activité ligand-indépendante d'ERα. Cette mutation induit également une meilleure sensibilité du récepteur aux œstrogènes, augmentant l'activité ligand-dépendante et permettant une stimulation même en cas de très faibles taux d'œstrogènes dus à l'utilisation d'AIs. L'utilisation de SERMs résulte quant à elle en une activité agoniste, et seuls les SERDs restent efficaces [51].

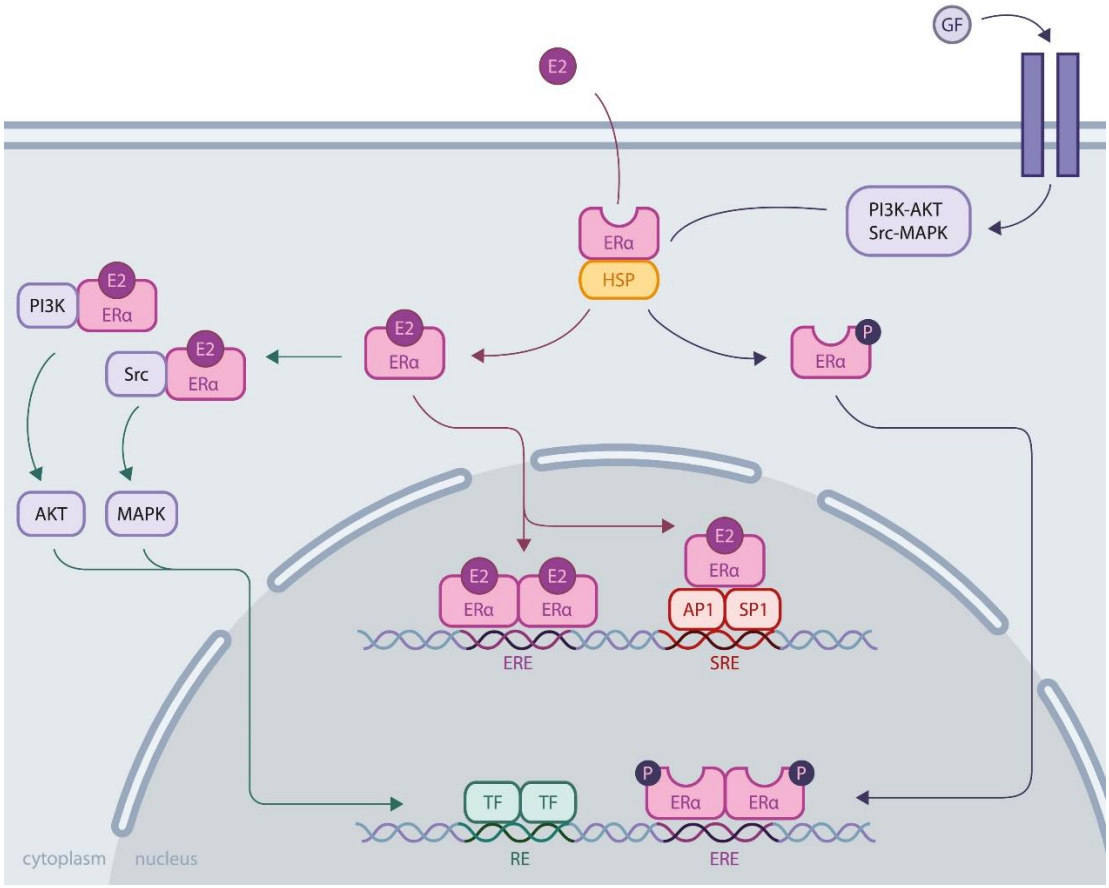


Figure 17 : Activités d'ERα.

Suite à son activation par les œstrogènes (E2) ou via phosphorylation après l'activation de récepteurs aux hormones de croissance (GF), ERα exerce son activité via des voies génomique et non-génomique. La voie génomique consiste en la régulation de la transcription de gènes cibles suite à la translocation d'ERα dans le noyau. Le récepteur peut alors lier l'ADN de manière directe, au niveau de motifs ERE, ou indirecte, via l'interaction avec d'autres facteurs de transcription (comme AP1 et SP1). La voie non-génomique consiste quant à elle en l'interaction entre ERα activé et des kinases dans le cytoplasme, résultant en leur activation et la régulation de l'expression d'autres gènes cibles. Figure issue de Clusan et al. [51].

La caractérisation de la structure d'ERα est primordiale pour la compréhension de son activité et le rôle joué dans le développement du cancer du sein puis l'échappement au contrôle hormonal. L'identification de mutations impactant la structure du récepteur et sa réponse à l'hormonothérapie permet en l'occurrence de mieux comprendre certains phénomènes de résistance. Cela met en lumière la nécessité de développer de nouvelles molécules pouvant cibler les formes mutées d'ERα, qui pourraient être combinées à des traitements ciblant les voies de signalisation dérégulées dans le cancer du sein. La détection de ces mutations chez les patientes pourrait notamment permettre d'adapter au mieux la stratégie thérapeutique, vers une prise en charge plus personnalisée.

b) Activités d'ERα

L'activation d'ERα par les œstrogènes a un effet mitogène sur les cellules environnantes, lui conférant un rôle primordial dans le développement mammaire. Celui-ci a pu être particulièrement étudié grâce à des modèles murins exprimant différentes mutations du récepteur ou de ses partenaires, détaillés dans la revue de Rusidzé et al. [29]. Cet effet est médié par deux activités du récepteur : une activité génomique, et une non-génomique. La première fait suite à la translocation du récepteur activé au noyau et sa liaison, directe ou indirecte, à l'ADN pour réguler la transcription de gènes cibles. L'activité non-génomique démarre quant à elle dans le cytoplasme, où la stimulation d'ERα peut mener à sa liaison avec des kinases et les activer, résultant in fine en la régulation de gènes supplémentaires (Figure 17). Ces deux voies permettent ainsi l'activation de l'expression de gènes impliqués dans la prolifération et survie cellulaire. La caractérisation moléculaire de ces activités est quant à elle permise par des études in vitro, menant aux connaissances détaillées ci-après.

i. Activité génomique

En absence de ligand, ERα est séquestré sous forme inactive dans le cytoplasme via l'interaction avec des protéines HSP90/70. La liaison des œstrogènes à ERα induit des modifications conformationnelles, notamment au niveau de l'hélice 12 du LBD, induisant sa dissociation des protéines chaperonnes. ERα activé va alors former un dimère avec un autre récepteur activé avant d'être transférés au noyau, où aura lieu le recrutement de cofacteurs pour permettre la liaison du dimère d'ERα à l'ADN. Cette liaison peut être directe, au niveau de séquences ERE, ou indirecte, via l'interaction avec d'autres facteurs de transcription tels que SP1 ou AP1. Dans ce cas, les complexes formés lient l'ADN au niveau des motifs propres à ces facteurs de transcription. ERα peut également être activé de manière ligand-indépendante, via la phosphorylation de résidus présents dans la fonction AF1 notamment. Ces phosphorylations sont réalisées par les effecteurs de différentes voies de signalisation des facteurs de croissance, suite à l'activation d'IGF1R et EGFR par exemple (Figure 17) [29].

ERα régule la transcription de gènes cibles grâce au recrutement d'une panoplie de cofacteurs, qui vont activer ou au contraire réprimer la transcription, de par leur activité enzymatique au niveau de la chromatine ou à différentes étapes de la transcription (initiation, élongation, terminaison, épissage…).

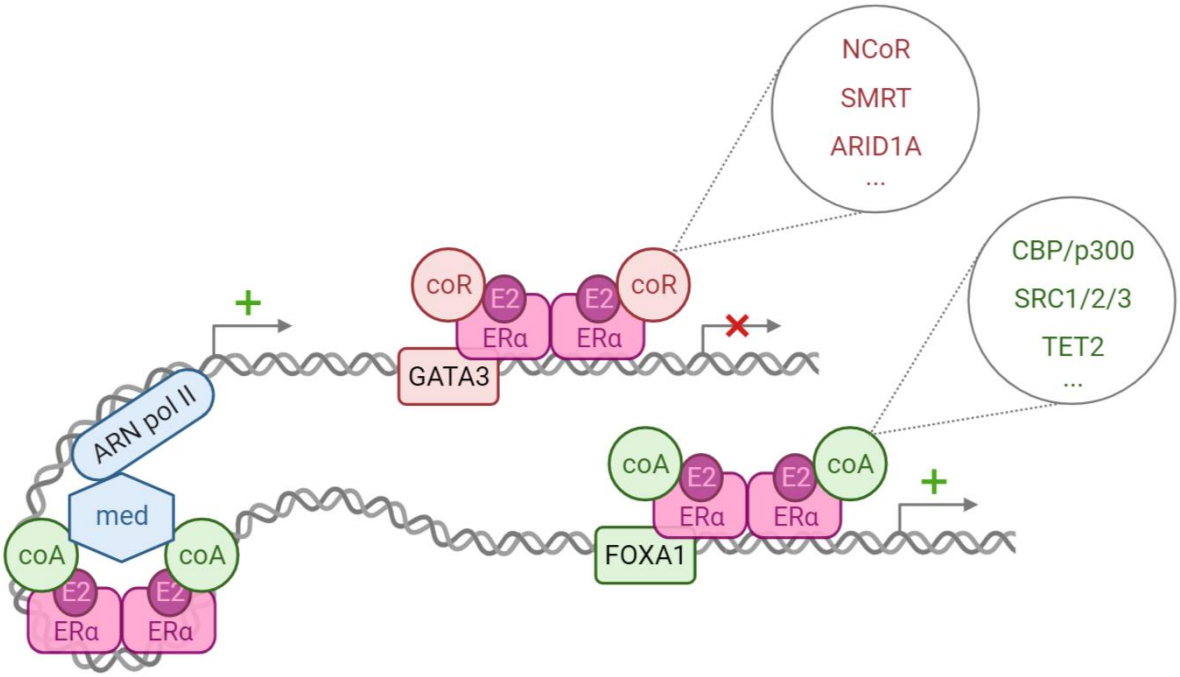


Figure 18 : Activité génomique directe d'ERα.

Après liaison de facteurs pionniers tels que GATA3 ou FOXA1 à l'ADN, les dimères d'ERα activé peuvent lier leurs éléments de réponse et recruter des corépresseurs (coR) ou coactivateurs (coA), menant à la répression ou activation de la transcription. La liaison du récepteur à l'ADN peut avoir lieu à plus de 100 kb de distance du gène cible, dont la régulation de l'expression est permise par un remodelage de la chromatine rapprochant ERα et la machinerie transcriptionnelle recrutée du promoteur ciblé. Figure adaptée de Manavathi et al. et Farcas et al. [52,53]. E2 : œstrogènes ; med : complexe médiateur.

Tout d'abord, ERα lie des facteurs pionniers tels que FOXA1 ou GATA3, qui ont la capacité de lier leurs éléments de réponse dans un contexte d'hétérochromatine pour ensuite permettre la liaison d'autres facteurs de transcription comme ERα. L'interaction avec des coactivateurs tels que la famille de protéines p160 (SRC1, AIB1) ou les protéines CBP/p300 est généralement permise au niveau des domaines C-terminaux des dimères d'ERα liant un ligand. Il semblerait dans ce cas qu'un dimère d'ERα interagisse d'abord avec SRC1 et AIB1, qui vont alors permettre le recrutement de CBP et p300, dont l'activité d'acétylation des histones stimule alors la transcription. D'autres modifications de la chromatine peuvent être provoquées par le recrutement d'autres coactivateurs par ERα, tels que CARM1 ou le complexe SWI/SNF. ERα peut aussi plus directement permettre le recrutement de l'ARN polymérase II en interagissant avec des acteurs du complexe médiateur tel que TRAP220 [48]. ERα peut à l'inverse recruter des corépresseurs, tels que NCoR et SMRT qui vont permettre la répression de la transcription via le recrutement d'histone désacétylases. Le recrutement de corépresseurs a lui aussi lieu au niveau du domaine E du récepteur, après liaison d'un ligand permettant le réarrangement conformationnel exposant les hélices nécessaires à l'interaction avec le corépresseur. A noter que les ERE permettant la liaison d'ERα pour réguler l'expression de gènes cibles ne se situent qu'en minorité au niveau des promoteurs des gènes en question. Ces motifs ERE sont en effet plus fréquemment situés à des sites distants des gènes cibles, impliquant l'existence d'importantes modifications de la chromatine pour former des boucles rapprochant ERα et la machinerie transcriptionnelle recrutée du promoteur ciblé (Figure 18) [48]. La dérégulation de l'expression de ces cofacteurs, avec notamment une augmentation de la quantité de coactivateurs au détriment des corépresseurs, participe alors à la progression du cancer du sein en privilégiant une activité transcriptionnelle activatrice de la part d'ERα vis-à-vis de ses gènes cibles impliqués dans la prolifération et migration cellulaire. La surexpression du coactivateur AIB1 ou la diminution de l'expression de NCoR1 ont notamment été associées avec la résistance au tamoxifène [52,54]. Outre des variations d'expression de cofacteurs, l'acquisition de résistance à l'hormonothérapie et plus généralement le développement cancéreux mammaire sont marqués par des modifications globales de la liaison d'ERα à l'ADN (cistrome). En effet, dix mille sites de liaison du récepteur ont été identifiés dans des cellules de cancer du sein luminal MCF7, pour la régulation de quelques centaines de gènes seulement. La majorité de ces sites resteraient alors inoccupés suite à l'activation d'ERα, mais permettrait une certaine flexibilité de réponse transcriptionnelle en fonction du contexte cellulaire. Il apparaît par exemple que l'activation ligand-indépendante d'ERα résulte en un cistrome différent de celui observé suite à une activation par les œstrogènes, avec plus de motifs liés par AP1 par exemple. Le cistrome d'ERα serait également différent entre des tissus sains et cancéreux, puis avant ou après hormonothérapie, avec un basculement vers des programmes transcriptionnels pro-tumoraux. La plasticité du cistrome d'ERα serait alors un paramètre clef de son activité génomique, menant à la reprogrammation de l'expression de gènes pour le développement du cancer du sein et la résistance aux traitements hormonaux [53,55].

Outre la caractérisation du cistrome d'ERα dans différentes conditions, le développement des techniques de séquençage a permis d'identifier un panel de plusieurs centaines de gènes régulés par la stimulation d'ERα par les œstrogènes. Les résultats diffèrent cependant selon les études, aussi bien d'un point de vue quantitatif que qualitatif, malgré l'utilisation de la même lignée de cellules et de traitements similaires. Ces différences seraient expliquées par la moindre variation en terme de manipulation des cellules et de paramètres utilisés pour l'analyse des données de séquençage [56]. Les gènes œstrogéno-régulés identifiés de manière récurrente sont notamment impliqués dans la régulation du cycle cellulaire, la prolifération ou encore l'apoptose tels que : MYC, CCND1, FOXM1, GREB1, BCL2, ou encore AREG et CXCL12 [52,57].

Cette activité transcriptionnelle est médiée par les fonctions de transactivation AF1 et AF2 d'ERα, qui participent cependant de manière inégale à l'activité du récepteur. Il a notamment été montré par des travaux réalisés au sein de l'équipe que la prédominance d'une activité AF1 ou AF2 dépend de l'état de différenciation de la cellule. En l'occurrence, l'activité génomique est principalement permise par la fonction AF1, lorsque la cellule est bien différenciée, alors que cette activité est perdue dans les cellules dédifférenciées, où l'activité transcriptionnelle est globalement réduite et ne repose plus que sur la fonction AF2 [58]. La prévalence de la fonction AF1 dans les cellules différenciées est notamment liée à l'activité de sa boîte 1, stimulée par la présence de jonctions cellulaires, expliquant pourquoi la perte de ces jonctions en cas de transition épithélio-mésenchymateuse (EMT) mène à la suppression d'activité d'AF1 [59].

ii. Activité non-génomique

ERα présente également une activité dite non-génomique. Dans ce cas, l'activation du récepteur (par les œstrogènes ou de manière ligand-indépendante) mène à l'interaction d'ERα avec des kinases intracellulaires et leur activation, telles que ERK, AKT et Src. Le rôle de certaines PTMs d'ERα a été identifié dans ces interactions, comme la méthylation de l'arginine 260 pour la liaison avec PI3K et Src, ou la phosphorylation de la tyrosine 537 pour l'interaction avec Src également [48]. Des cofacteurs impliqués dans l'activité génomique d'ERα peuvent également jouer un rôle dans son activité non-génomique, en liant ERα dans le cytoplasme et participant à son interaction avec des kinases. Le coactivateur PELP1 sert par exemple d'adaptateur entre ERα et Src. La dérégulation de l'expression de ces cofacteurs impacte alors l'activité non-génomique d'ERα en plus de son activité génomique, participant au phénomène de résistance à l'hormonothérapie [54]. ERα peut aussi interagir plus en amont avec les récepteurs aux facteurs de croissance (EGFR, IGFR…) ou des protéines G, dont l'activation induit là aussi les voies de signalisation jouant un rôle dans le développement du cancer du sein [60]. L'activité non-génomique d'ERα est caractérisée par sa rapidité d'action (quelques secondes à minutes), les cascades de signalisation ainsi activées menant in fine à la régulation de la transcription de gènes cibles différents de ceux régulés via l'activité génomique d'ERα. L'identification de gènes régulés suite à la stimulation de cellules MCF7 avec des composés ne pouvant activer que la voie non-génomique d'ERα a en effet mis en évidence que seulement 25% de ces gènes étaient également régulés par une stimulation aux œstrogènes, permettant elle une activité-génomique en plus de l'activité non-génomique du récepteur [61].

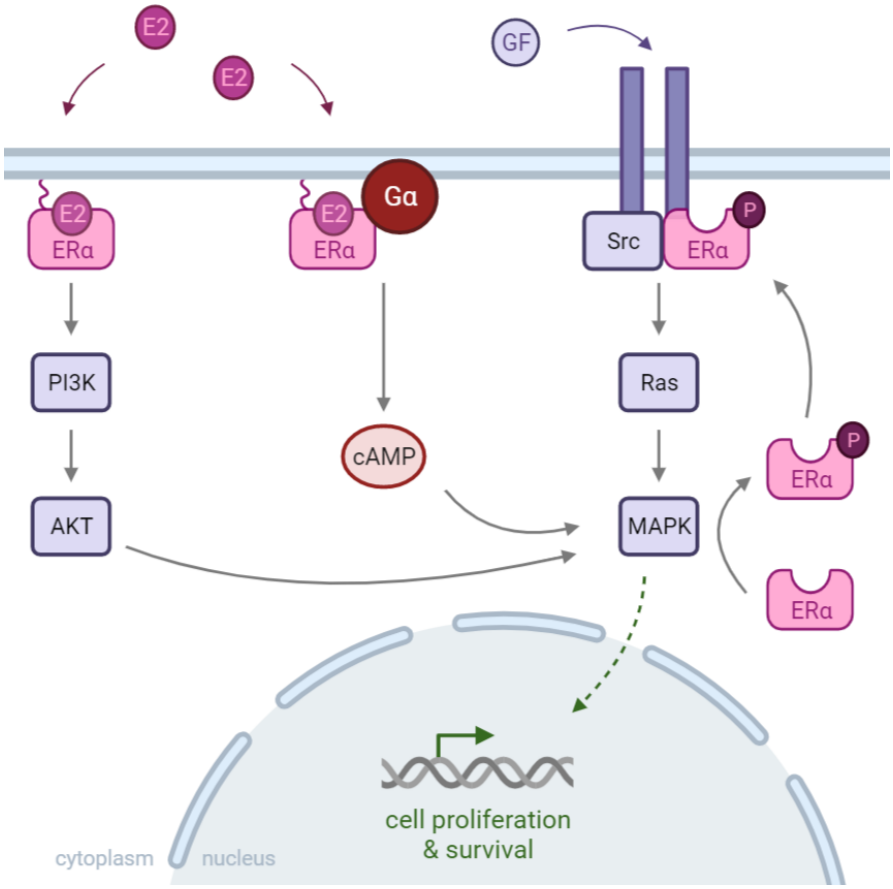


Figure 19 : Activité non-génomique d'ERα.

ERα peut être ancré à la membrane plasmique ou interagir avec des protéines membranaires telles que des protéines G ou des récepteurs aux facteurs de croissance (EGFR, IGFR...). Son activation peut alors entraîner la production de seconds messagers comme l'AMP cyclique (cAMP) et l'induction de voies de signalisation telles que PI3K/AKT ou RAS/MAPK. Une fois activées, les kinases intracellulaires peuvent elles-mêmes induire l'activation d'ERα par phosphorylation, ce qui amplifie son activité non-génomique, menant in fine à la régulation de l'expression de gènes impliqués dans la prolifération et survie cellulaire. E2 : œstrogènes ; GF : facteurs de croissance.

Cette activité non-génomique est particulièrement induite lors de la stimulation d'ERα localisé à la membrane plasmique dans les cellules de l'endothélium vasculaire (suite à la palmitoylation de la cystéine 447 ou la myristoylation du récepteur). Dans ce cas, la stimulation d'ERα membranaire par les œstrogènes mène à la production de monoxyde d'azote permettant la relaxation des vaisseaux sanguins par exemple, d'où un effet vasoprotecteur des œstrogènes [29,48].

L'activité non-génomique d'ERα est alors interconnectée avec des cascades de signalisation intracellulaires telles que PI3K/AKT et RAS/MAPK, elles-mêmes capables d'activer ERα de manière ligand-indépendante (Figure 19). L'activation de ces voies étant oncogénique, l'activité non-génomique d'ERα joue ainsi un rôle important dans la prolifération, survie et migration des cellules cancéreuses mammaires [52].

ERα est donc un récepteur nucléaire pouvant être activé de manière ligand-dépendante ou indépendante. Dans les deux cas, l'activité de ce récepteur repose sur une voie génomique directe ou indirecte mais aussi une voie non-génomique. Chacune de ces activités entraîne la régulation de l'expression d'un panel de gènes impliqués dans la prolifération et survie cellulaire. De nombreux mécanismes peuvent alors prendre part à la modification d'activité d'ERα, menant au développement du cancer du sein puis à l'échappement au contrôle hormonal. Ceux-ci incluent notamment [30] :

- des mutations activatrices d'ERα ;
- la dérégulation de ses PTMs ;
- la suractivation des voies de signalisation des facteurs de croissance ;
- un déséquilibre dans l'expression de ses cofacteurs ;
- une dérégulation de l'expression d'ERα.

Tous ces mécanismes favorisent une activité ligand-indépendante d'ERα et augmentent ses capacités activatrices au détriment de l'effet répressif que présente le récepteur dans un épithélium mammaire sain. Des modifications fonctionnelles d'ERα sont donc l'une des conséquences du profond remaniement cellulaire subi par les cellules cancéreuses mammaires. Cela peut mener à un échappement au contrôle hormonal, et donc la résistance à l'hormonothérapie chez 30% des patientes atteintes de cancer du sein luminal.

Afin de comprendre la survenue des cancers, le développement des techniques de biologie moléculaire a d'abord mis en évidence le rôle joué par les mutations avant de mettre en lumière les reprogrammations épigénétiques à l'œuvre dans les cellules cancéreuses. Ces modifications épigénétiques sont notamment apparues comme des facteurs clefs permettant l'adaptation de ces cellules à leur environnement, en affectant l'expression de protéines pro- ou anti-tumorales. La dérégulation des voies de signalisation en résultant permet ainsi la prolifération et dédifférenciation des cellules cancéreuses, menant à des phénotypes plus agressifs [62,63]. Ce phénomène d'adaptation est nommé plasticité cellulaire, et est primordial dans le développement des cellules cancéreuses qui doivent faire face à de profonds changements d'environnements tels que l'hypoxie, un moindre approvisionnement en nutriments, ou encore l'usage de molécules anti-cancéreuses. En plus de ces mécanismes touchant l'ADN, il est apparu que la traduction jouerait elle aussi un rôle important dans la plasticité des cellules cancéreuses. En adaptant la quantité mais aussi la nature des protéines synthétisées, la régulation de la traduction permet en effet de privilégier des programmes pro-tumoraux [64,65].

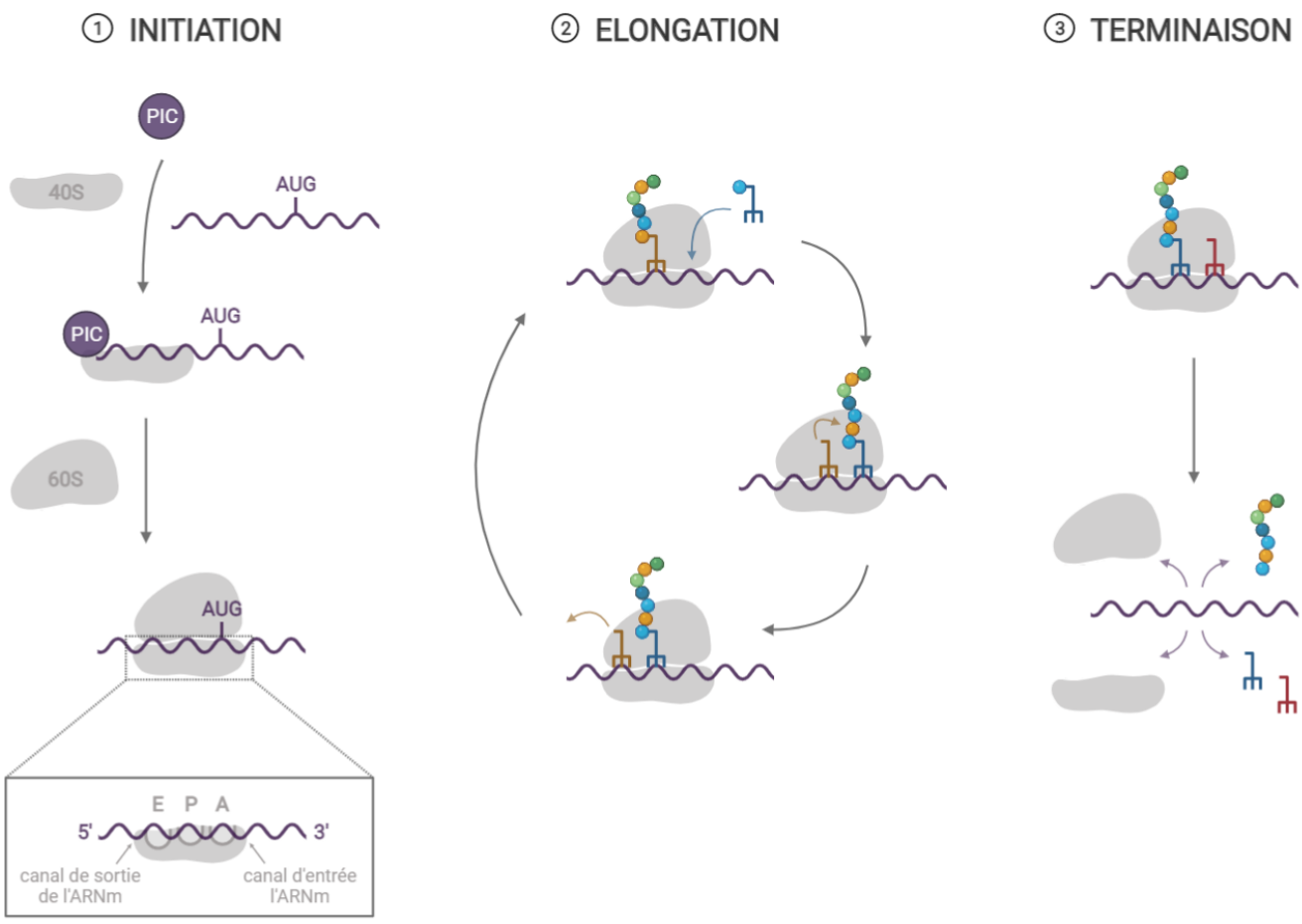


Figure 20 : Mécanisme de la traduction.

1) Lors de l'initiation de la traduction, la petite sous-unité ribosomale (40S) s'associe avec le complexe de pré-initiation (PIC) pour former un complexe d'initiation avant d'être recruté par l'ARNm. Ce complexe d'initiation balaye alors la région 5' non traduite de l'ARNm jusqu'à reconnaître le codon d'initiation AUG. Cette reconnaissance permet ensuite le recrutement de la grande sous-unité ribosomale (60S), formant ainsi le ribosome 80S fonctionnel prêt à passer à l'étape d'élongation de la traduction. 2) L'élongation consiste à recruter l'ARNt chargé correspondant au codon situé dans le site A du ribosome puis former une liaison peptidique entre l'acide aminé apporté au site A et le peptide en cours de traduction au niveau du site P. La translocation du ribosome a ensuite lieu, permettant la sortie de l'ARNt déchargé au site E, et la libération du site A pour le recrutement d'un nouvel ARNt chargé. 3) Suite à l'enchaînement de cycles d'élongation, la terminaison de la traduction a lieu lorsqu'un codon stop entre dans le site A, provoquant la dissociation du ribosome 80S et libération de la protéine traduite. Figure adaptée de Fabbri et al. [66].

3. LA TRADUCTION ET LE REPLIEMENT CO-TRADUCTIONNEL

La traduction désigne le phénomène biologique permettant la production des protéines à partir de l'ARNm au sein des cellules. Il s'agit d'un processus complexe, se déroulant en trois étapes que sont l'initiation [67], l'élongation [68], et la terminaison [69] (Figure 20). De nombreuses revues détaillent chaque étape, chacune étant finement modulée par différents facteurs afin d'assurer la production optimale des protéines, d'un point de vue qualitatif comme quantitatif, et ainsi permettre le bon fonctionnement des cellules.

Ces dernières années, la mise au point de nouvelles méthodes de protéomique et transcriptomique a permis d'acquérir de plus en plus d'informations sur la régulation de la traduction. Il est notamment apparu que la phase d'élongation n'est pas uniforme : tous les ARNm ne sont pas traduits à la même vitesse, et un même ARNm n'est pas traduit à une vitesse constante le long de sa séquence. De plus, la cinétique de traduction d'un ARNm pourrait être différente selon les conditions cellulaires. En effet, les acteurs de la machinerie traductionnelle sont multiples et peuvent être modifiés en réponse à une variété de stimuli, impactant in fine la régulation de la traduction. Dans le cas de cancers par exemple, il est connu que la traduction est fortement remaniée afin d'adapter la production de protéines impliquées dans la tumorigenèse ainsi que le développement de résistances aux traitements [66,70].

Cette hétérogénéité dans la vitesse d'élongation jouerait un rôle fonctionnel dans la production des protéines. En effet, tandis qu'une vitesse de traduction élevée est associée à une production massive de protéines, un certain ralentissement est nécessaire pour assurer la fidélité de la traduction (sélection de l'ARN de transfert correct pour chaque codon afin d'incorporer le bon acide aminé à la chaîne peptidique), l'adressage des protéines synthétisées, leur association avec des protéines partenaires, ou encore leur repliement. En revanche, une vitesse d'élongation trop faible peut entraîner la dégradation de l'ARNm et du peptide en cours de traduction [71]. La régulation de la vitesse de traduction est donc critique pour assurer la production des protéines nécessaires au bon fonctionnement de la cellule ainsi que leur fonctionnalité. En effet, au-delà de produire une certaine quantité de protéines, l'élongation de la traduction joue un rôle majeur dans un phénomène nommé repliement co-traductionnel. Ce processus désigne la capacité de certaines protéines à se replier et adopter leur conformation directement en cours de traduction, indispensable à leur bonne structuration et donc l'acquisition de leurs propriétés fonctionnelles. Ce repliement peut débuter au sein du tunnel de sortie du ribosome, puis une fois sorti via des interactions électrostatiques avec la surface du ribosome, ainsi que des interactions avec des protéines partenaires (Figure 21). Ces protéines peuvent être différentes chaperonnes ou directement des sous-unités dont l'assemblage en oligomères a lieu en cours de traduction [72,73]. A noter que les interactions entre le peptide et la surface du ribosome pourraient également inhiber le repliement co-traductionnel en entrant en compétition avec des protéines partenaires [74].

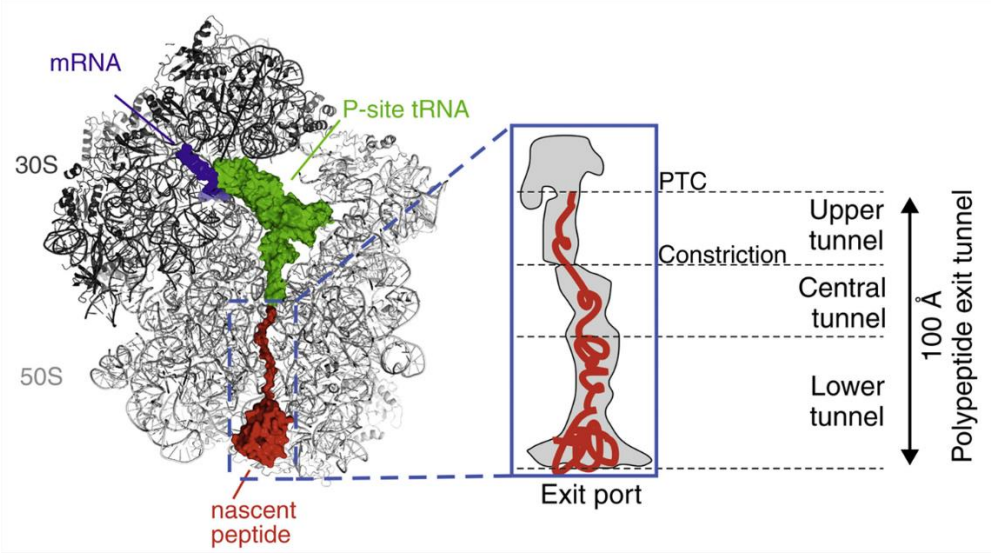


Figure 21 : Représentation structurelle d'un peptide en cours de traduction.

Le tunnel de sortie du ribosome débute au niveau du centre de transfert peptidique (PTC) et est divisé en trois régions (haute et centrale séparées par une constriction, puis région basse). Le repliement co-traductionnel débuterait à partir de la partie basse du tunnel. Ribosome d'*E.coli*, PDB 3J7Z32 ; figure issue de Thommen et al. [72].

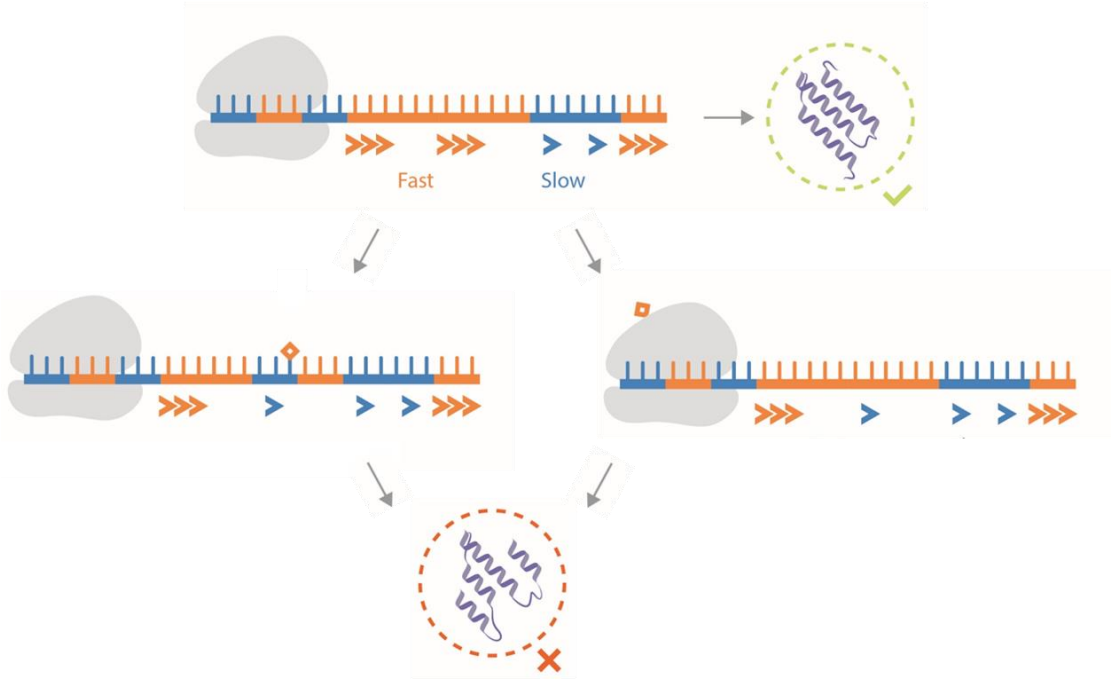


Figure 22 : Rôle de la vitesse de traduction dans le repliement co-traductionnel des protéines.

Des modifications de la séquence traduite (telles que des mutations synonymes) ou des ribosomes (telles que la phosphorylation de protéines ribosomales) peuvent modifier la vitesse de traduction d'un même ARNm et résulter en une conformation différente de la protéine. Figure modifiée à partir de Clusan et al. [51].

La conformation des protéines étant essentielle à leur fonction, la régulation de la vitesse de traduction est alors critique dans la fonctionnalité de nombreuses protéines. Modifier la vitesse de traduction pourrait alors altérer le repliement de ces protéines et ainsi être à l'origine de modifications fonctionnelles (Figure 22). Différents paramètres régulent cette vitesse de traduction, notamment liés au ribosome lui-même ou à la séquence traduite.

a) Rôle du ribosome dans le repliement co-traductionnel

Parmi les acteurs dont la modification peut jouer un rôle dans la régulation de la traduction, des travaux réalisés ces dernières années ont révélé que le ribosome lui-même est susceptible d'être modifié. Suite à la découverte des ribosomes dans les années 1950-60, il a longtemps été pensé que les ribosomes représentaient une population uniforme assurant une même fonctionnalité qui est la traduction des ARNm en protéines. La fin du XXème siècle a cependant vu émerger des travaux mettant en évidence des différences de composition des ribosomes chez les invertébrés et, depuis 2010, de plus en plus d'études s'intéressent à l'hétérogénéité des ribosomes et leur rôle fonctionnel chez l'homme [75].

Le ribosome eucaryote est constitué de 80 protéines, 4 ARN ribosomaux (ARNr) ainsi que des segments d'expansion d'ARNr (Tableau 1, Figure 23) [76].

Tableau 1 : Composition du ribosome eucaryote en protéines et ARN.

Sous-unité ribosomale	Nombre de protéines	ARNr	Nombre de segments d'expansion
Petite (40S)	33	18S	5
Grande (60S)	47 (46 chez la levure)	5S 5,8S 28S (25S chez la levure)	16

La dynamique des interactions entre ces différents composants confère la capacité des ribosomes à traduire l'ARNm en protéine. Des modifications touchant les protéines et/ou ARNr pourraient impacter leurs interactions et ainsi mener à des différences de capacité traductionnelle. De récentes études ont en effet mis en évidence des variations en termes de composition en protéines et ARNr, ceux-ci pouvant être modifiés de manière post-traductionnelle ou -transcriptionnelle, résultant en un éventail de combinaisons possibles qu'il reste à identifier et caractériser.

i. Diversité des protéines ribosomales

Des différences de composition en protéines ribosomales ont d'abord été mises en évidence par l'étude de ribosomopathies (maladies génétiques altérant la fabrication et/ou fonction des ribosomes) et l'identification de protéines ribosomales mutées à partir de 1999 [77,78].

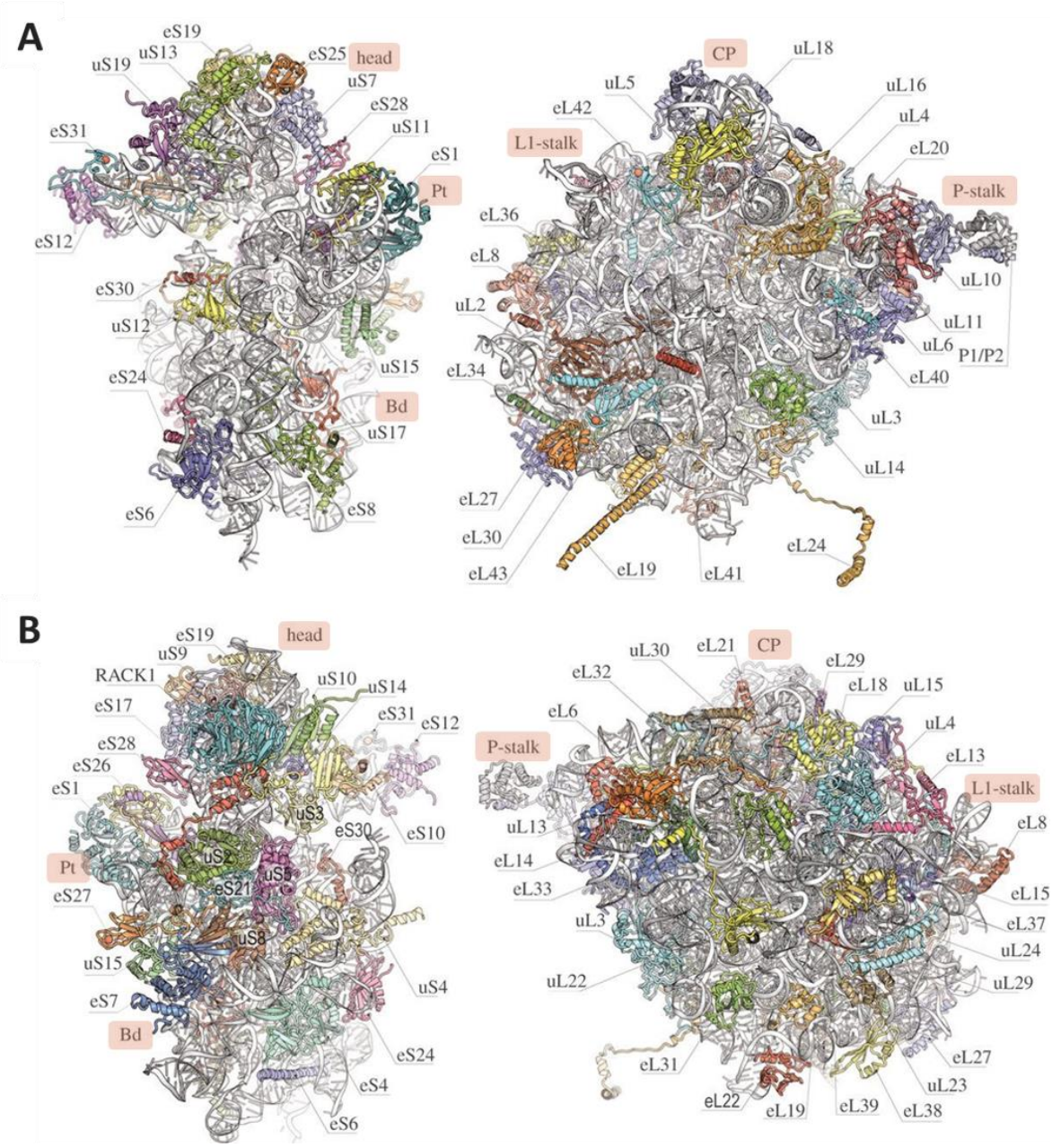


Figure 23 : Structure du ribosome eucaryote.

A gauche : petite sous-unité 40S ; à droite : grande sous-unité 60S. A) Vue de l'interface entre les sous-unités ; B) vue des faces exposées. Les protéines ribosomales universelles (retrouvées chez les eucaryotes, bactéries et archaebactéries) portent le préfixe « u » tandis que les protéines spécifiques aux eucaryotes portent le préfixe « e ». Figure issue de Yusupova et Yusupov [79]. Bd : body ; CP : central protuberance ; Pt : platform.

En dehors de ce contexte pathologique, plusieurs études se sont depuis intéressées aux différences de composition des ribosomes mais n'ont pas encore établi de consensus. En effet, des travaux publiés en 2015 mettent en évidence un enrichissement différent en protéines ribosomales (essentiellement de surface) entre des monosomes et polysomes (populations composées d'un seul ou plusieurs ribosomes traduisant un même ARNm, respectivement) au sein de mêmes cellules (cellules souches embryonnaires murines, mESCs) [80]. Cette observation n'a cependant pas été reproduite dans une étude de 2018 réalisée dans d'autres lignées cellulaires (HEK293 et HeLa) [81]. Cette dernière étude a cependant montré que la phosphorylation d'une protéine ribosomale particulière (RPL12) serait enrichie en monosomes, et que les ribosomes portant cette modification traduiraient des ARNm spécifiques (enrichis en A ou U à la troisième position des codons, codant des protéines ayant un rôle en mitose) [81].

Les protéines ribosomales peuvent en effet porter différentes modifications post-traductionnelles qui vont impacter la traduction. Des marques de phosphorylation et ubiquitination ont notamment été identifiées, mais peu de travaux ont pu élucider leurs conséquences fonctionnelles pour le moment [82].

D'autres études ont mis en évidence des traductions d'ARNm spécifiques liées à certaines protéines ribosomales, telles que RPL38 chez la souris [83], RPS26 chez la levure [84], RPS25 et RPL10A en mESCs [85]. Il est à noter que cette dernière étude a également mis en évidence un appauvrissement en certaines protéines ribosomales au sein des polysomes.

ii. Diversité des ARN ribosomaux

Tandis que la séquence des ARNr diffère peu, plus de 130 modifications chimiques ont été recensées en 2017 [86,87]. La majorité des modifications identifiées touche les ARNr 18S et 28S et est retrouvée au niveau de sites fonctionnels du ribosome, tels que les centres de décodage ou de transfert peptidique, les sites de liaison à l'ARNm et ARN de transfert (ARNt), ou encore le tunnel de sortie du peptide, suggérant un rôle dans la fonctionnalité des ribosomes. Les modifications situées proches des sites A ou P peuvent par exemple impacter l'élongation de par leur rôle dans le maintien du cadre de lecture ou la formation de la liaison peptidique [87]. 95% de ces modifications consistent en des méthylations du ribose (2'-O-Me) ou pseudouridylations (isomérisation de l'uridine en pseudouridine, Ψ), les 5% restant étant des méthylations ou acétylations de bases [88]. Les événements de 2'-O-Me et Ψ se produisent lors de la synthèse des ribosomes mais ne sont pas indispensables à leur fabrication. Ils modifient la structure des ARNr et confèrent ainsi une hétérogénéité au sein des populations de ribosomes, dont l'impact dans la régulation de la traduction n'est pas encore complètement élucidé (Figure 24) [89,90]. Les études réalisées ces dernières années mettent en évidence un rôle dans la fidélité de la traduction ou la traduction d'ARNm spécifiques (notamment liée à la présence d'IRES) mais des recherches supplémentaires sont nécessaires pour mieux comprendre les mécanismes liant la modification des ARNr au développement de certaines pathologies telles que les cancers, dont le cancer du sein [91,92].

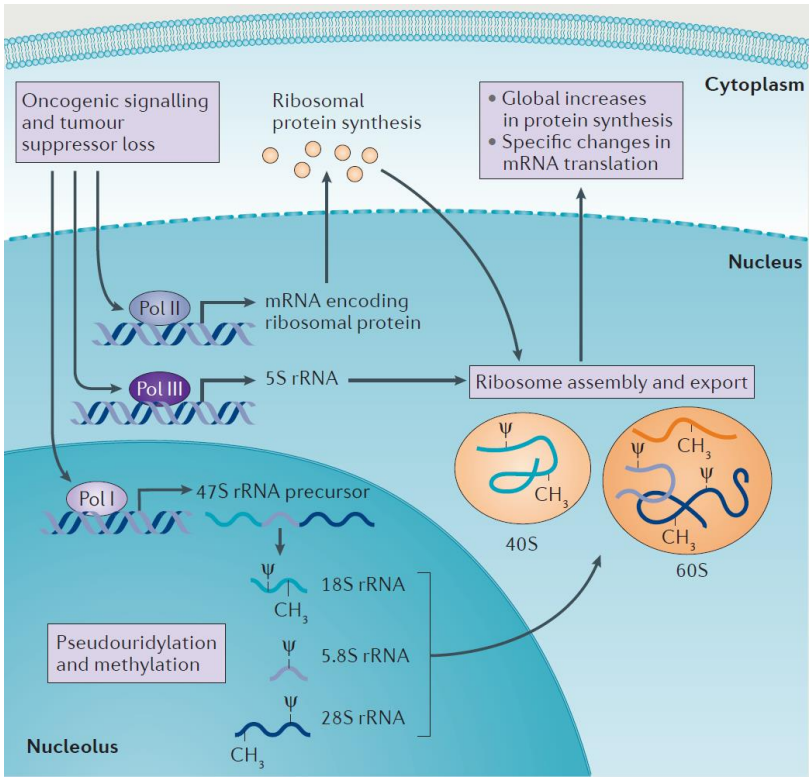


Figure 24 : Potentiel rôle des modifications des ARNr dans la cancérogénèse.

La méthylation (CH_3) et la pseudouridylation (Ψ) des ARNr ont lieu lors de leur synthèse dans le nucléole. Ces modifications pourraient être induites par des signaux oncogéniques, résultant en la production de ribosomes aux propriétés fonctionnelles différentes, participant alors au développement de cancers. Figure issue de Truitt et Ruggero [70]. Pol : ARN polymerase.

1ère base	2nd base U	2nd base C	2nd base A	2nd base G	3ème base
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA STOP	UGA STOP	A
	UUG Leu	UCG Ser	UAG STOP	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met (start)	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

Figure 25 : Table du code génétique.

Mis à part la méthionine et le tryptophane, chaque acide aminé peut être codé par deux à six codons différents.

Bien que de plus en plus de travaux mettent en évidence une hétérogénéité des ribosomes au sein des cellules, la composition et le rôle fonctionnel associé de ces populations reste à être élucidés. Des ribosomes différant de par leurs protéines et/ou ARNr pourraient ainsi être impliqués dans des variations de vitesse de traduction, mais cette régulation reste à être caractérisée.

b) Rôle de l'usage des codons dans le repliement co-traductionnel

Parallèlement à la population de ribosomes, d'autres paramètres jouant un rôle dans les variations de vitesse de traduction sont liés à la séquence de l'ARNm traduit.

i. Biais d'usage des codons

La dégénérescence du code génétique permet d'utiliser des codons différents pour coder une même protéine (Figure 25). Cependant, l'utilisation de ces codons synonymes n'est pas uniforme à l'échelle du génome, mais est au contraire dit « biaisée » : certains codons sont plus utilisés que d'autres pour coder un même acide aminé. Or, bien que codant pour le même acide aminé, l'utilisation de codons synonymes peut modifier la structure de l'ARNm et être plus ou moins adaptée au stock d'ARNt disponible pour traduire la séquence. Le biais d'usage des codons participerait alors à la régulation de la production de protéines en impactant l'efficacité de transcription [93], la stabilité des messagers, et l'initiation de la traduction (Figure 26) [94].

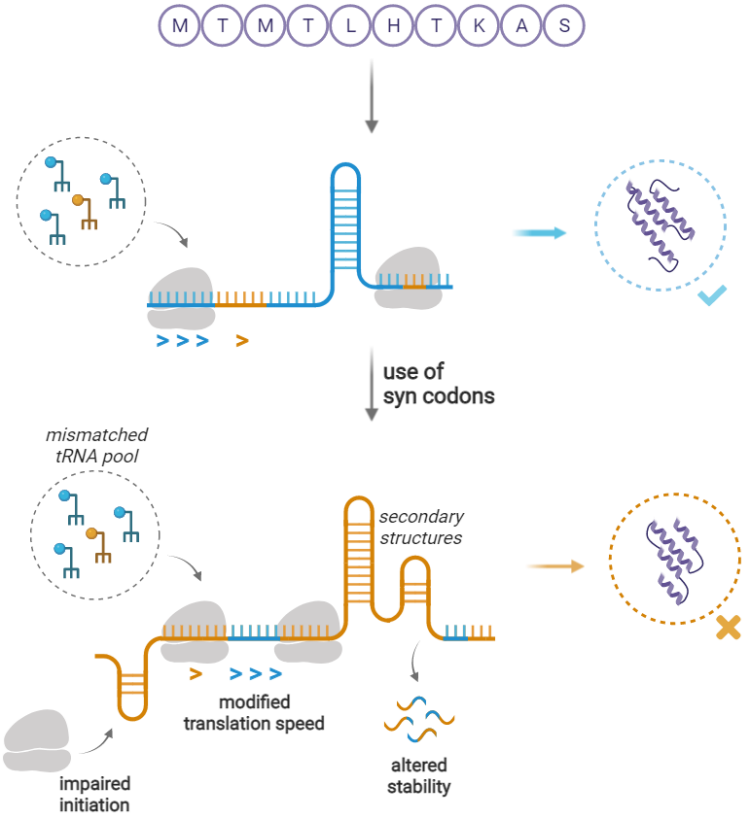


Figure 26 : Conséquences de l'utilisation de codons synonymes sur la production d'une même protéine. Utiliser des codons différents pour coder une même séquence d'acides aminés peut modifier la structure de l'ARNm et être moins adapté au stock d'ARNt disponibles. Cela peut impacter l'initiation de la traduction, la vitesse d'élongation, et la stabilité de l'ARNm, modifiant alors la quantité de protéine produite ainsi que sa conformation. Figure adaptée de Nieuwkoop et al. [94].

Cette caractéristique des codons composant les ARNm serait également impliquée dans la régulation de la vitesse d'élongation de la traduction. Il est alors question de codons fréquents vs rares mais aussi de codons optimaux vs non-optimaux, cette optimalité étant usuellement déterminée en fonction de l'adéquation avec le pool d'ARNt. Il est ainsi supposé que les codons optimaux seraient traduits plus rapidement que les codons non-optimaux, car les ARNt correspondant étant en quantité majoritaire dans la cellule, les ribosomes réaliseraient l'association codon-anticodon correcte plus rapidement que dans le cas où les ARNt adéquats sont plus rares. Cette hypothèse a été vérifiée chez la levure [95,96], mais le rôle des codons synonymes dans la vitesse de traduction ne se jouerait pas seulement à l'échelle codon à codon. Différents travaux pointent en effet l'importance de considérer les codons environnant celui en cours de lecture dans le site A, générant un environnement qui influe sur la vitesse d'élongation, de par l'enchaînement de couples codons-anticodons particuliers ou encore la nature des acides aminés codés [97,98]. L'enchaînement d'acides aminés chargés positivement ralentirait la traduction par exemple [98].

Il semblerait que l'usage des codons soit similaire pour les gènes formant certains groupes fonctionnels [99]. Modifier le stock d'ARNt disponible pourrait alors être une réponse cellulaire pour adapter le protéome à certaines conditions, en favorisant la traduction de certains gènes dans le cas de cancers par exemple [100–102]. La fréquence des codons semblant corrélée à leur optimalité d'après ces études, l'utilisation des qualificatifs « optimal » et « fréquent » vs « non-optimal » et « rare » est alors confondue pour définir l'usage des codons, mais il semblerait plutôt que le caractère « optimal » d'un codon soit dépendant de plusieurs paramètres bien plus complexes qu'une simple corrélation avec la disponibilité des ARNt [94].

ii. Corrélation entre usage des codons et vitesse de traduction

Quel que soit le qualificatif employé, le rôle du biais d'usage des codons dans la vitesse d'élongation a un temps été débattue, mais il apparaît bien à présent que ce paramètre participe aux variations de vitesse de traduction le long des ARNm [103]. Certains travaux ont alors montré que les codons dits optimaux seraient traduits plus rapidement, mais cette régulation reste à être démontrée pour la traduction de gènes endogènes (d'autant plus lorsque l'on questionne l'attribution du caractère optimal aux différents codons synonymes) [94]. Aussi, l'utilisation de différents codons peut mener à la formation de structures secondaires de l'ARNm, ce qui pourrait ralentir la traduction, mais ce mécanisme de régulation reste à être investigué [94].

En régulant la vitesse de traduction, l'usage des codons joue alors un rôle dans le repliement co-traductionnel des protéines. Différentes études ont ainsi montré que la modification de codons synonymes résulte en des protéines de conformation différente (généralement attesté de manière indirecte par des expériences de digestion ménagée, dénaturation thermique ou encore capacité à former des agrégats) suite à la modification de la vitesse de traduction [103–105].

Pour mieux comprendre cette relation entre usage des codons et repliement co-traductionnel, des analyses bio-informatiques ont été réalisées et suggèrent que les codons optimaux seraient préférentiellement retrouvés au sein de séquences codant des régions très structurées, tandis que les codons non-optimaux composeraient plutôt les régions plus flexibles voire intrinsèquement désordonnées. Les protéines possédant des domaines intrinsèquement désordonnés seraient d'ailleurs plus sensibles à l'impact de l'usage des codons sur leur repliement co-traductionnel [106]. Il est à noter que ces protéines ont tendance à être plus riches en GC [107], et que la partie N-terminale d'ERα est elle-même intrinsèquement désordonnée [108]. Des travaux réalisés par Fernàndez-Calero et al. ont par ailleurs montré qu'une mutation synonyme dans cette partie du récepteur entraîne des modifications fonctionnelles suggérant une modification de sa conformation [109]. Des prédictions réalisées chez des procaryotes (*E.coli* et *B.subtilis*) ainsi qu'une analyse effectuée chez la levure ont quant à elles recensé des codons rares à la jonction des domaines constituant certaines protéines, suggérant un rôle de ralentissement de la traduction pour terminer le repliement d'un domaine avant de traduire le suivant [110–112]. D'autres études indiquent cependant que ces codons rares se retrouvent en réalité tout au long des messagers, avec une certaine conservation entre procaryotes et eucaryotes soulignant le rôle fonctionnel probable du biais d'usage des codons (Figure 27) [113,114]. De telles observations suggèrent ainsi que les codons utilisés portent une information supplémentaire à la nature de l'acide aminé, pour permettre la structuration correcte des protéines en régulant la vitesse de traduction [103].

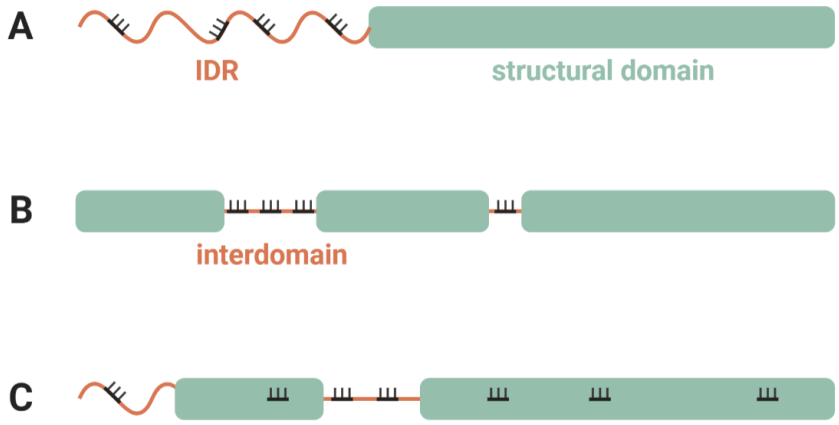


Figure 27 : Modèles proposés de position des codons rares au sein des séquences codant des protéines. A) Les codons rares (en noir) seraient préférentiellement utilisés pour coder des régions intrinsèquement désordonnées (IDR). B) Un autre modèle propose que les codons rares seraient localisés entre les domaines structurels composant la protéine. C) Selon un dernier modèle, les codons rares seraient retrouvés tout le long des séquences codantes, sans localisation préférentielle mais avec une certaine conservation entre procaryotes et eucaryotes.

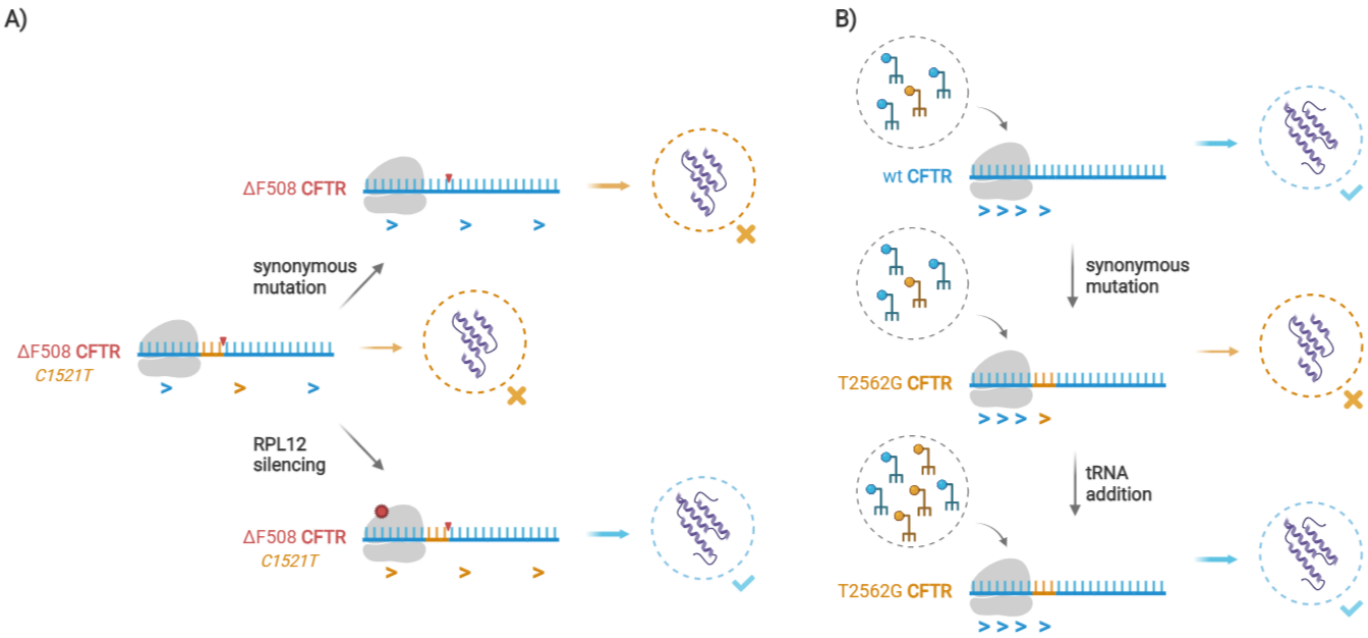


Figure 28 : Résumé des travaux réalisés sur le rôle des codons synonymes et la vitesse de traduction dans la production de la protéine CFTR.

A) La délétion Δ508 associée à la mutation synonyme C1521T modifie la production, conformation et fonction de la protéine CFTR. Tandis que la restauration du codon synonyme permet d'augmenter la production du mutant Δ508, la suppression de la protéine ribosomale RPL12 induit un ralentissement de la traduction menant à la restauration de la conformation et fonction de la protéine CFTR. B) La mutation synonyme T2562G modifie la production, conformation et fonction de la protéine CFTR dû à un ralentissement de sa vitesse de traduction. Cette conséquence est liée au manque d'ARNt complémentaire au codon synonyme ; l'augmentation de la concentration de cet ARNt permettant de restaurer la production et fonction de la protéine CFTR sauvage. Figure adaptée de Kirchner et al. [115].

Le rôle fonctionnel du biais d'usage des codons serait alors à prendre en compte en biotechnologie, pour la production ou l'étude fonctionnelle de protéines en système hétérologue [94,116]. Aussi, ces travaux remettent en question le rôle des mutations synonymes dans la survenue de pathologies. Il a en effet été montré que de telles mutations sont associées à différentes maladies en impactant la transcription et/ou la traduction des protéines [117]. Un exemple de protéine dont la composition en codon, la vitesse de traduction et le repliement co-traductionnel jouent un rôle dans le développement d'une pathologie concerne la protéine CFTR (cystic fibrosis transmembrane conductance regulator) impliquée dans la mucoviscidose. La principale mutation de CFTR menant à la mucoviscidose est la délétion de la phénylalanine 508 (ΔF508), qui génère par ailleurs une mutation synonyme du codon codant l'isoleucine 507 (C1521T). La protéine mutante CFTR-ΔF508 présente une modification de conformation et fonction, ainsi qu'une importante dégradation. Il a alors été montré que restaurer le codon sauvage de l'isoleucine 507 chez le mutant CFTR-ΔF508 permet d'en augmenter le niveau d'expression, soulignant le rôle joué par les mutations synonymes dans la production des protéines [118]. La conformation et fonctionnalité de la protéine CFTR-ΔF508 ont quant à elles pu être restaurées en diminuant la vitesse de traduction grâce à l'inhibition d'une protéine ribosomale particulière, RPL12 (Figure 28, A) [119]. Une autre mutation commune chez les patients atteints de mucoviscidose, bien que ne causant pas à elle seule cette pathologie, est une mutation synonyme de la thréonine 854 de CFTR (T2562G). Il a été montré que ce polymorphisme diminue le taux d'expression de la protéine CFTR et modifie sa stabilité, conformation et fonction. Ces conséquences seraient liées au manque d'ARNt correspondant au codon synonyme, induisant un ralentissement de la vitesse de traduction de CFTR. Augmenter la concentration cellulaire de l'ARNt manquant a en effet permis de restaurer la production, conformation et fonction de la protéine CFTR sauvage (Figure 28, B) [115]. De telles expériences mettent en lumière le potentiel rôle fonctionnel de mutations synonymes en pathologie humaine, à prendre en compte pour une meilleure compréhension des processus pathologiques et la prise en charge des patients.

Loin d'être un processus uniforme, l'élongation de la traduction apparaît donc être une phase dont la régulation est critique pour la production de protéines fonctionnelles. Les travaux réalisés ces dernières années mettent ainsi en lumière l'implication de ribosomes hétérogènes et le biais d'usage des codons dans la régulation de cette étape de la traduction. Les mécanismes exacts expliquant comment est régulée la traduction d'un ARNm en fonction de sa séquence par des ribosomes particuliers restent cependant à élucider. Leur compréhension permettrait d'améliorer la production de protéines in vitro et de développer nos connaissances de biologie cellulaire, notamment pour mieux comprendre le développement de pathologies telles que les cancers.

Contexte et objectifs de la thèse

Les travaux présentés dans ce manuscrit de thèse résultent de trois années de recherche au sein de l'équipe DREAM (Réparation d'ADN, expression d'ARN, signalisation d'œstrogène et carcinome mammaire) de l'Irset (Institut de recherche en santé, environnement et travail), UMR_S Inserm 1085, sous la supervision des docteurs Farzad Pakdel et Gilles Flouriot.

Les travaux entrepris au sein de l'équipe s'inscrivent dans l'étude et la lutte contre le cancer du sein hormono-dépendant, de par l'étude des mécanismes moléculaires impliqués dans la résistance aux traitements antiprolifératifs. Un des axes de recherche de l'équipe vise notamment à élucider les mécanismes contrôlant la conformation du récepteur aux œstrogènes ERα, et mieux comprendre la relation structure/fonction de cette protéine impliquée dans le cancer du sein. C'est dans ce contexte que s'inscrivent mes travaux de thèse, financés par un contrat doctoral de l'Université de Rennes 1. L'objectif du projet était d'élucider le rôle joué par la composition en codons dans la traduction de l'ARNm d'ERα en protéine fonctionnelle. En effet, de précédentes études ont montré qu'il existe un biais d'usage des codons au sein des séquences codantes, et que celui-ci diffère pour exprimer certaines protéines ou entre des tissus sains et cancéreux. La composition en codons d'un ARNm pourrait alors jouer un rôle fonctionnel dans la production des protéines.

Après avoir mis en évidence l'implication de la voie de signalisation Rho/actine/MRTFA dans la transition épithélio-mésenchymateuse des cellules cancéreuses mammaires, et les modifications fonctionnelles d'ERα que cela entraîne, l'intérêt du docteur Gilles Flouriot a en effet été porté sur le potentiel rôle de l'usage des codons dans la fonctionnalité des protéines. Mon projet a ainsi débuté en explorant le rôle de la fréquence d'usage des codons dans la fonctionnalité d'ERα en fonction de l'état cellulaire. Ces travaux ont été soumis en 2022 et constituent la première partie des résultats présentés dans ce manuscrit. Notre intérêt s'est ensuite plus spécifiquement porté sur le rôle joué par la 3ème base des codons dans la traduction des protéines en général et d'ERα en particulier. Un enrichissement en G/C ou A/T au niveau de cette troisième base semblait en effet participer à la régulation de l'expression des gènes et la production de protéines spécifiques au cours du cycle cellulaire. Ces travaux constituent le second chapitre des résultats présentés dans ce manuscrit, et s'inscrivent dans une plus grande étude réalisée avec différents collaborateurs, qui sera prochainement soumise à publication. Les résultats obtenus mettant en avant une régulation de la traduction dépendant de la composition en codon, j'ai par la suite développé un protocole de profilage de polysomes (« polysome profiling ») afin d'explorer la traduction des gènes selon leur enrichissement en G/C et l'état cellulaire. La composition de la machinerie traductionnelle a également été étudiée dans différentes conditions biologiques, afin d'identifier des candidats participant à cette régulation de la traduction

dépendant de l'usage des codons et de l'état cellulaire. Les résultats obtenus, présentés en chapitre trois, clôturent ainsi mes travaux de thèse, et seront étoffés au sein de l'équipe avant publication.

En parallèle de ces travaux de recherche, j'ai rédigé une revue au sujet des mutations faux sens et synonymes d'ERα et leur implication dans la prise en charge du cancer du sein hormono-dépendant. Cette revue, publiée en 2021 dans *International Journal of Molecular Sciences*, est accessible en annexe.

Résultats

1. L'ADAPTATION DES CODONS PAR DES MUTATIONS SYNONYMES MODIFIE LES PROPRIETES FONCTIONNELLES DU RECEPTEUR AUX ŒSTROGENES ALPHA DANS LES CELLULES CANCEREUSES MAMMAIRES

Le récepteur aux œstrogènes alpha (ERα) est impliqué dans une variété de processus physiologiques et pathologiques, le principal étant le cancer du sein. Le rôle joué par ERα est alors lié à l'effet mitotique de son ligand, les œstrogènes, qui stimulent la croissance tumorale. Dans le cas de cancers du sein ERα-positif (ERα+), le traitement de choix est alors l'hormonothérapie, qui vise à inhiber l'activité d'ERα en réduisant le taux d'œstrogènes circulants ou en liant directement le récepteur. Une résistance hormonale a cependant lieu pour 30% des patientes, chez qui les cellules tumorales échappent au contrôle hormonal et continuent de proliférer en absence de stimulation oestrogénique, rendant les anti-œstrogènes inefficaces. Cette résistance est associée à un basculement du phénotype des cellules cancéreuses mammaires vers un état moins différencié et plus prolifératif, pouvant mener à des processus métastatiques. Récemment, deux programmes traductionnels distincts ont été mis en évidence au cours de la prolifération ou différentiation cellulaire, basés sur des répertoires d'ARNt et fréquences d'usage des codons différents. Considérant le basculement de phénotype des cellules cancéreuses mammaires vers un état plus prolifératif et moins différencié, des modifications du stock d'ARNt et de l'usage des codons pourraient avoir lieu chez ces cellules, rendant la séquence nucléotidique d'ERα non adaptée à une traduction correcte dans ces conditions. En effet, des modifications de la disponibilité des ARNt peuvent mener à des modifications de la vitesse de traduction et du repliement co-traductionnel d'ERα, altérant sa conformation et alors ses propriétés fonctionnelles. Pour vérifier cette hypothèse, une séquence synonyme d'ERα a été créée, en optimisant la fréquence d'usage des codons pour que celle-ci soit plus proche des fréquences observées dans des cellules en prolifération. Les propriétés fonctionnelles du récepteur ont alors été étudiées, en mesurant son activité transcriptionnelle, ses interactions avec des partenaires génomiques et non-génomiques, et son impact sur le devenir cellulaire en générant une lignée stable de cellules MCF7. Nous avons montré que l'optimisation des codons restaure les propriétés fonctionnelles d'ERα observées dans des cellules différenciées, en améliorant l'activité transcriptionnelle médiée par la fonction AF1, en conférant de meilleures capacités répressives en présence de corépresseurs, et en diminuant les interactions avec des kinases. Concernant le devenir cellulaire, la version synonyme d'ERα réduit la prolifération induite par les œstrogènes des cellules MCF7 en abolissant l'effet anti-apoptotique des œstrogènes. Nous montrons également que ces modifications fonctionnelles sont associées à des modifications de la conformation d'ERα. Ensemble, ces résultats mettent en évidence un rôle joué par l'usage des codons dans la conformation et la fonctionnalité d'ERα, qui pourrait participer à la résistance hormonale des cancers du sein ERα+.

Codon adaptation by synonymous mutations impact the functional properties of the estrogen receptor-alpha protein in breast cancer cells

Léa Clusan[1], Frederic Percevault[1], Emmanuelle Jullion[2], Pascale Le Goff[1], Christophe Tiffoche[2], Tamara Fernandez-Calero[3,4], Raphaël Métivier[2], Monica Marin[5], Farzad Pakdel[1], Denis Michel[1] and Gilles Flouriot[1*]

1 Univ Rennes, Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail) – UMR_S 1085, F-35000 Rennes, France

2 Univ Rennes, Institut de Génétique et Développement de Rennes, UMR 6290 CNRS, Rennes, France

3 Departamento de Ciencias Exactas Y Naturales, Universidad Catolica del Uruguay, 11600 Montevideo, Uruguay

4 Bioinformatics Unit, Institut Pasteur Montevideo, Mataojo 2020, 11400 Montevideo, Uruguay

5 Biochemistry-Molecular Biology, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay

*To whom correspondence should be addressed: Dr. Gilles Flouriot ; Irset, Inserm U1085, 9 rue du Prof. Léon Bernard, 35000 Rennes cedex, France ; Phone: +33-2 23 23 68 04 ; Fax: +33-2 23 23 67 94 ; E-mail: gilles.flouriot@univ-rennes1.fr

Abstract

The estrogen receptor-alpha (ERα) is intimately associated with the development of hormone-dependent breast cancers, promoting tumor growth due to the mitotic effect of its ligand, estradiol. Thus, ERα is then the main therapeutic target through the use of antiestrogens as anti-mitotic agents. Unfortunately, tumor proliferation often escapes hormonal control, making endocrine therapy inefficient. Breast cancer cells then evolve towards a less-differentiated and more proliferating phenotype that can lead to metastatic processes. Recently, two distinct translation programs with specific tRNA repertoires and codon usage frequencies were evidenced during proliferation and differentiation. Considering the phenotype switch of breast cancer cells to more proliferating and less differentiated states, we can speculate that changes in tRNA pool and codon usage that likely occur make ERα coding sequence no longer suitable, impacting translational rate, co-translational folding and finally the functional properties of the protein. To verify this hypothesis, we generated an ERα synonymous sequence whose codon usage was optimized to the frequencies observed in genes expressed specifically in proliferating cells. The functional properties of the receptor were investigated by assessing its transactivation activity, its interactions with genomic and non-genomic partners and its impact on cell fate after establishing its stable expression in MCF7 cell line. We demonstrate that codon adaptation restores functional properties of ERα observed in differentiated cells, including an improved activity of AF1 transactivation function, better repressive capability in the presence of corepressors and reduced interactions with kinases. On cell fate, it weakens estrogen-induced cell proliferation of MCF7 cells by abolishing the anti-apoptotic effect of E2. We show that functional modifications are associated with conformational changes in ERα protein. Altogether, these results shed light on a role played by codon usage in the conformation and function of ERα that could participate to endocrine resistance in ERα-positive breast cancer.

Abbreviations: AFs: Transactivation Functions; DMEM: Dulbecco's modified Eagle's medium; E2: 17β-estradiol; EGF: Epidermal Growth Factor; EMT: Epithelial-Mesenchymal Transition; ERα: Estrogen Receptor-alpha; ERE: ERα Responsive Element; FCS: Fetal Calf Serum; ICI: ICI 182-780; OHT: 4-hydroxytamoxifen; PLA: Proximity Ligation Assay; SERD: Selective Receptor Degrader; SERM: Selective Estrogen Receptor Modulators; SYN opt: Synonymous optimized; WT: wild type.

Keywords: Estrogen receptor-alpha; codon usage; co-translational folding; breast cancer; endocrine resistance

1. Introduction

As the main mediator of estrogens, the estrogen receptor-alpha (ERα) is involved in a variety of physiological processes ranging from the establishment and maintenance of the female sexual differentiation patterns to cardiovascular and neuronal systems, to liver, fat and bone metabolisms (Auchus and Fuqua, 1994; Couse and Korach, 1999). ERα influences also several pathological processes of which the most common is breast cancer (Dahlman-Wright et al., 2006; Platet et al., 2004). Approximatively 70% of diagnosed breast cancers express ERα whose activation by estrogens favors the proliferation of breast cancer cells. For these cancers, the therapy of choice is endocrine therapy, which aims to deprive the tumor of estrogens or directly block ERα activity respectively through the use of aromatase inhibitors or antiestrogens such as the tamoxifen (R. Clarke et al., 2015; Jordan and O'Malley, 2007). However, endocrine resistance occurs in 30% of cases due to a hormonal escape, with cells continuing to proliferate without estrogenic stimulation and becoming resistant to endocrine therapy. Breast tumors then switch to less-differentiated phenotypes and undergo invasive and metastatic processes associated with unfavorable vital prognosis (Jordan and O'Malley, 2007).

ERα belongs to the nuclear receptor superfamily and like all members it is primarily a transcription factor that exhibits genomic activity through the regulation of gene expression. ERα transcriptional activity is mediated by two transactivation functions: AF1, localized in the N-terminal part of the protein, and AF2 in the C-terminal ligand-binding domain. These two functional domains present interface surfaces that allow the recruitment of coactivators or corepressors in an ordered, cyclical and combinatorial manner (Beato, 1989; Métivier et al., 2003). Both AFs do not participate equally in ERα genomic activity and it has been showed that their contribution depends upon the cell differentiation stage (Mérot et al., 2004). The loss of AF1 activity in dedifferentiated cells has notably been linked to the loss of cell-cell junctions, which particularly occurs during epithelial-mesenchymal transition (EMT) (Huet et al., 2008; Mérot et al., 2004). In addition to its genomic activity, ERα is also known to modulate non-genomic membrane signaling pathways through a protein pool located at the cytoplasm and/or plasma membrane and capable of interacting with several kinases such as c-src and PI3K (Edwards, 2005). Activation of the downstream MAPK and AKT signaling pathways ultimately leads to the stimulation of cell proliferation and survival (Arnal et al., 2017).

Several mechanisms are known to promote endocrine resistance. These include ER-activating mutations, imbalance between coactivators and corepressors expression or overactivation of growth factors signaling pathways (Hanker et al., 2020a). Endocrine resistance is generally accompanied by the epithelial-mesenchymal transition (EMT) of breast cancer cells, and functional modifications of ERα (Jehanno et al., 2021; Kerdivel et al., 2014). We recently demonstrated that, during EMT, the activation of the actin/MRTFA pathway results in functional modifications of ERα due to the impairment of its localization and interactions with partners. These mechanisms ultimately lead to a decrease of its transcriptional activity and an increase of its non-genomic activity, promoting hormonal escape (Jehanno et al., 2020, 2021). Such modifications in ERα activity are likely to result from changes in environment and ERα partners as mentioned above but may also be due to protein conformational changes that have an impact on its functional activity.

Due to the degeneration of the genetic code, amino acids can be encoded by several codons. These synonymous codons are not equally used in coding sequences, leading to a codon usage bias. A correlation between preferred codons and more abundant tRNAs was observed, suggesting that tRNAs pools and codon usage are essential players in the regulation of gene expression. They may differ depending on the cell type, cell cycle phase, or differentiation status of the cell (Gingold et al., 2014; Novoa and Ribas de Pouplana, 2012). For instance, Gingold et al. demonstrated in 2014 a strong dependence between tRNA selection and abundance and the transcriptional programs involved during cell proliferation and differentiation (Gingold et al., 2014). In particular, tRNA pool was correlated to codon usage bias in genes specifically expressed in differentiated or proliferating cells improving expression of cell state specific genes. More recently, specific expression of tRNAs in metastatic breast cancer cells that may promote tumor progression has been observed (Goodarzi et al., 2016). In fact, the availability of tRNA appears to be a key parameter impacting translation speed, and translation speed modification can alter the co-translational folding of some proteins (Hanson and Coller, 2018; Marín et al., 2017; Rodnina, 2016; Yu et al., 2015). Concerning ERα, previous work notably suggested that its conformation would be sensitive to the cellular environment and codon composition (Fernández-Calero et al., 2014; Horjales et al., 2007). It should be recalled that in healthy breast tissues, ERα is expressed in differentiated cells that do not or rarely proliferate whereas in ERα-positive breast cancers, it is often expressed in proliferative cells (R. B. Clarke, 2004). A conformational change in ERα during tumor transformation or EMT could therefore be induced by a change in tRNA abundance and codon usage, altering the translation rate and final conformation of the receptor.

In the present study, we show that ERα codon usage fits the codon usage of genes specifically expressed in differentiation. We generated an estrogen receptor synonymous sequence to adapt the frequency of ERα codon usage to that observed in genes specifically expressed in proliferative cells, and then investigated the functional consequence of these synonymous mutations on ERα activity. We demonstrate that these synonymous mutations improve the AF1 transcriptional activity of ERα and alter ERα interactions with its partners by reducing its interaction with kinases and promoting the formation of complexes with

corepressors. Finally, expression of the synonymous mutant in MCF7 cell line abolishes the anti-apoptotic effect of estradiol.

2. Materials and methods

2.1. Cell culture and treatments

HepG2 (RRID:CVCL_0027), MCF7 (RRID:CVCL_0031), MDA-MB-231 (RRID:CVCL_0062) and SUM159PT (RRID:CVCL_5423) cells were routinely maintained in DMEM (Gibco) supplemented with 8% fetal bovine serum (FBS; Biowest) and antibiotics (Gibco) at 37 °C in 5% CO_2. Most cell lines are ATCC origin. Control and overexpressing GFP-ERα WT or GFP-ERα SYN-opt MCF7 clones were obtained by transfecting cells with the pcDNA6/TR plasmid and the corresponding pcDNA4/TO expression vectors (T-Rex system, Invitrogen) using JetPEI® (Polyplus transfection). GFP epitope was fused at the N-terminal domain of the proteins. The clones were selected in a medium containing 5 µg/mL blasticidin and 100 µg/mL zeocin (Invitrogen). Individual clones were isolated and grown in a medium containing selective antibiotics to maintain selection pressure. All cell lines were routinely monitored with luminal and basal-like markers and mycoplasma detection tests. When treatments with steroids were required, the cells were maintained 48 h in DMEM (Gibco) supplemented with 2.5% dextran/charcoal-stripped FBS (dsFBS; Biowest) prior to the experiments. The induction of GFP-ERα WT or GFP-ERα SYN-opt expression in MCF7 clones was performed with 1 µg/mL tetracycline for 48 h. Pharmacological treatments were performed with final concentrations of 10 nM 17β-estradiol (E2), 1000 nM 4-hydroxytamoxifen (OHT), 100 nM ICI 182-780 (ICI) or the solvent 0.1% ethanol as a control. EGF was used at 50 ng/mL.

2.2. Plasmids and transient transfections

The ERE-tk-Luc, C3-Luc, AP1-Luc and SP1-Luc reporter genes as well as the internal control CMV-βgal, the pCR-ERα, pCR-ERα Δ79, pCR-ERα Δ173, pCR-SRC1 and pCR-NCoR1 expression vectors have been previously described (Huet et al., 2008; Métivier et al., 2002; Pham et al., 2021). pCR-ERα SYN-opt was produced synthetically (GeneArt, Life technologies). pCR-ERα SYN-opt Δ79 and pCR-ERα SYN-opt Δ173 expression vectors were generated by subcloning corresponding PCR products into pCR3.1 vector (Invitrogen). To assess ERα transcriptional activity on reporter genes, 50 ng of control (pCR3.1) or ERα expression vectors were co-transfected with 200 ng of reporter genes and 200 ng CMV-βgal by well in 24-well plates. When pCR-MRTFA ΔN200, pCR-SRC1 or pCR-NCoR1 expression vectors were added to the DNA mix, 200 ng of plasmid were used. For immunofluorescence and proximity ligation assay, 250 ng of expression vectors were used by well in 24-well plates. For western blot analysis, 500 ng of expression vectors were transfected by well in 6-well plates. Transfections were carried out with JetPEI® (Polyplus transfection) according to the manufacturer's instructions. GFP-ERα WT or GFP-ERα SYN-opt expression vectors were generated from the pcDNA4/TO expression vector (T-Rex system, Invitrogen) by subcloning the corresponding PCR products.

2.3. Luciferase assay

Cells were plated in 24-well plates and incubated for 24 h before medium exchange for serum and steroid starvation. Cells were transfected as previously described and treated with ligands for 24 h before luciferase assay (Luciferase Assay System, Promega) (Huet et al., 2009). β-galactosidase activity was determined by incubating two third of each lysate with 0.7 mg/mL O-nitrophenyl β-D-galactopyranoside (ONPG) in assay buffer (0.8 mM $MgCl_2$, 35 mM β-mercaptoethanol) before measuring the absorbance at 415 nm with an iMark Microplate Absorbance Reader (Bio-Rad). Luciferase activity reflecting ERα transcriptional activity was normalized to the β-galactosidase activity.

2.4. Immunofluorescence and Proximity ligation assay (PLA)

Cells were plated on cover slides in 24-well plates and incubated for 24 h before medium exchange for serum and steroid starvation. When cell treatment was completed, phosphate buffered saline (PBS) containing 4% paraformaldehyde was used to fix the cells for 15 min. After PBS washing, cells were permeabilized in PBS containing 0.3% Triton X-100 for 15 min. After washing, an incubation with primary antibodies (1:1000) in PBS containing 3% FBS was performed overnight at 4 °C. For immunofluorescence experiments, cells were incubated the next day with secondary antibodies (1:1000) in PBS-FBS for 2 h at room temperature. For PLA, the Duolink® Proximity Ligation Assay reagents from Sigma were used to detect ERα interactions with partners, according to manufacturer's instructions, as previously described (Jehanno et al., 2021). Finally, the cover slides were mounted in Duolink® In Situ Mounting Medium with DAPI (Sigma). Images were obtained with an ApoTome Axio Z1 Imager microscope (Zeiss) and processed with AxioVision software. Fluorescent cells and PLA dots were analyzed using ImageJ software.

2.5. Western blotting

Whole-cell extracts were lysed in RIPA buffer or directly prepared in 3X Laemmli buffer from subconfluent cells in 6-well plates. Subcellar fractionation of HepG2 cells was performed as previously described (Jehanno et al., 2021). Following sonication, the protein extracts were denatured for 5 min at 95 °C, separated on 10% SDS polyacrylamide gels, and transferred to nitrocellulose membrane (Amersham Biosciences). The proteins were then probed with specific antibodies (1:1000) as previously described (Mérot et al., 2004), and detected using the Substrat HRP Immobilon Western kit from Millipore.

2.6. Proliferation assay

Cells were plated in 24-well plates and incubated for 24 h before medium exchange for serum and steroids starvation. Tetracycline induction and pharmacological treatments were applied for six days before cell trypsinization and counting.

2.7. Flow cytometry analysis

Cells were plated in 10-cm dishes and incubated for 24 h before medium exchange for serum and steroids starvation. After 48 h tetracycline induction and pharmacological treatments, cells were collected in PBS containing 30% IFA buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 4% FBS) and centrifuged for 10 min at 800 g,

4 °C. Cells were fixed for 30 min on ice with 70% ethanol, and incubated for 20 min in IFA buffer at 4 °C. A 30-min incubation with 100 µg/mL RNAse A was then performed at 37 °C, and 50 µg/mL propidium iodide (Sigma) was finally added for 10 min at 37 °C. Finally, the cell cycle was then analyzed using the FACSCalibur flow cytometer (BD Biosciences).

2.8. TUNEL staining

Apoptosis was determined by detecting DNA fragmentation using terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) staining. This was performed with an In Situ Cell Death Detection Kit, Fluorescein (Roche) according to the manufacturer's instructions, as previously described (Pham et al., 2021). Images were obtained with an ApoTome Axio Z1 Imager microscope (Zeiss) and processed with AxioVision software. TUNEL-positive cells were analyzed using ImageJ software.

2.9. Quantitative RT-PCR (RT-qPCR)

Cells were plated in 6-well plates and incubated for 24 h before medium exchange for serum and steroids starvation. After 48 h tetracycline induction and 24 h pharmacological treatments, RNA was extracted using the Nucleospin RNA Plus kit (Macherey-Nagel) according to manufacturer's instruction. Quantitative RT-PCR was performed as previously described (Pham et al., 2021), with the primer sequences indicated in Supplementary Table 1.

2.10. Limited digestion

Cells were plated in 6-well plates and incubated for 24 h before medium exchange for serum and steroids starvation. After 48 h tetracycline induction, cells were scrapped in PBS and centrifuged for 7 min at 800 g, 4 °C. Cell pellet was sonicated and the lysate recovered after a 15-min centrifugation at 12,000 g, 4 °C. After protein quantitation with the DCTM Protein Assay kit (Bio-Rad), a same amount of proteins was digested by 0.2 to 1.6 ng/µL trypsin or chymotrypsin (Sigma Aldrich) for 30 min at 30 °C. Enzymatic reaction was stopped by adding 1X Laemmli buffer and samples were denatured for 5 min at 95 °C before Western blotting.

2.11. Reagents and antibodies

17β-estradiol (E2) and 4-hydroxy-tamoxifen (OHT) were purchased from Sigma. ICI 182-780 (ICI) was obtained from TOCRIS Bioscience. The primary antibodies used for western blotting and immunofluorescence analyses were as follows: antibodies from Santa Cruz Biotechnology against C-term ER (HC-20, sc-543), p-Akt1/2/3 (Ser473) (sc-7985-R), Akt1/2/3 (H-136, sc-8312), ERK1 (K-23, sc-94) and p-ERK (E-4, sc-7383); antibodies from Abcam against H3S10p (ab5176), H3 (ab12079) and Lamin A+C (JOL2, ab40567); antibody from Thermo Fisher against N-term ER (6F11, MA5-13304). The secondary peroxidase-conjugated donkey anti-rabbit (NA934V) and sheep anti-mouse (NA931V) antibodies were purchased from GE Healthcare. Alexa Fluor® dye-conjugated secondary antibodies from Invitrogen were employed for immunofluorescence.

2.12. Statistical analysis

Statistical analyses were performed using Student's t-test. The values are provided as the mean ± standard error of the mean (SEM).

3. Results

3.1. ERα codon usage is adapted to the translation program of differentiated cells and less adapted to the translation program of proliferating cells

In order to measure the adaptability of ERα codon usage to the dual translation program in cellular proliferation and differentiation (Gingold et al., 2014), we compared the frequency of codon usage deduced from the coding sequence of the human ERα gene with those of the two functional gene sets, "M phase of mitotic cell cycle" and "pattern specification process" (which we will call hereafter "proliferation" and "differentiation") from Gingold et al. 's study (Gingold et al., 2014). As expected, codon usage frequency for ERα correlated strongly with that of genes specifically expressed in differentiation (R^2 = 0.93) and weakly with that of genes specifically expressed in proliferation (R^2 = 0.46) (Figure 1). Since ERα is increasingly expressed in proliferating cells during mammary epithelial cell carcinogenesis, we generated, based on Gingold et al. 's study (Gingold et al., 2014), an ERα cDNA in which codons were mutated to synonymous codons in order to optimize the codon usage of ERα to proliferative cells. Specifically, we chose a codon for each amino acid whose frequency of use in genes specifically expressed in proliferation approximates the frequency of the original codon observed in genes specifically expressed in differentiation. Codons with a similar frequency of use between "proliferation" and "differentiation" were maintained. This mutant, called ERα SYN-opt, has an identical amino acid sequence to ERα WT, but a different DNA sequence (Figure 1 and Supplementary Figure 1). The functional properties of this synonymous mutant were then studied.

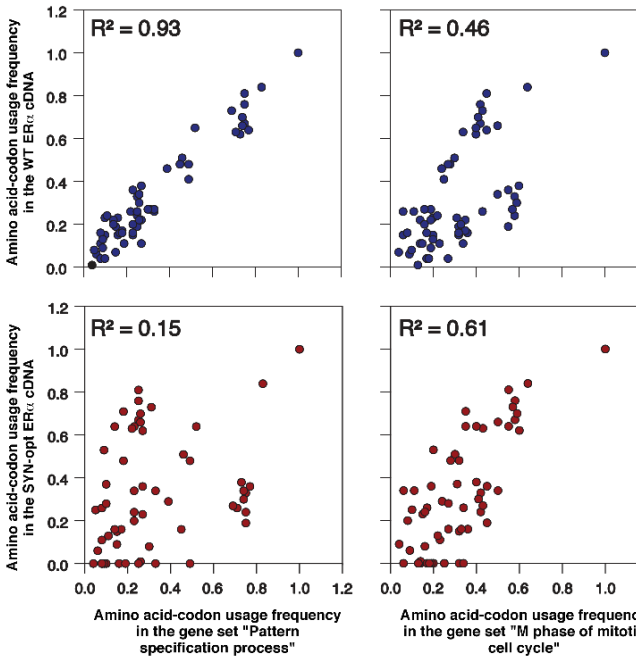


Figure 1: ERα codon usage is adapted to the translation program of differentiated cells and less adapted to the translation program of proliferating cells. Comparison of the amino acid codon usage frequency deduced from the coding sequences of WT and SYN-opt ERα with those of the gene set "M phase of mitotic cell cycle" (proliferation) and "pattern specification process" (differentiation) from the study of Gingold et al. (Gingold et al., 2014). Each point corresponds to one codon.

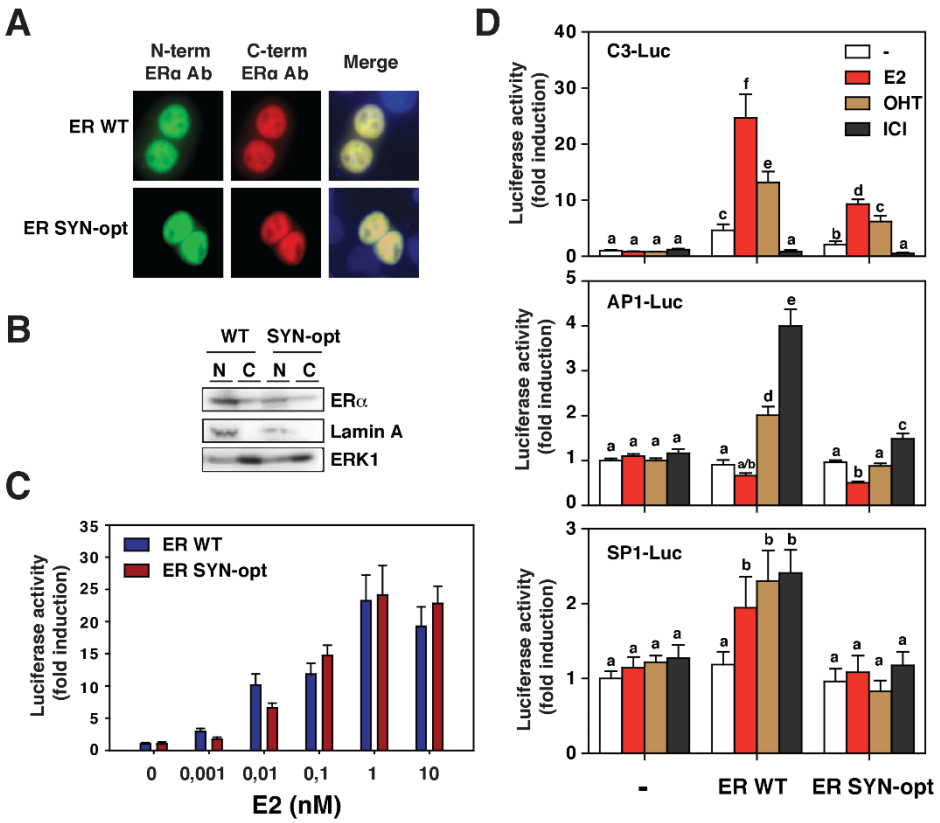


Figure 2: ERα SYN-opt exhibits different transcriptional activity than ERα WT on ERE-, AP1- and SP1-driven reporter genes. (**A**), Immunofluorescent detection of ERα WT and ERα SYN-opt proteins by a N-terminal and C-terminal ERα antibody (Ab), 48 h after transient transfection of the corresponding expression vectors in HepG2 cells. (**B**), Western blot analysis of ERα WT and ERα SYN-opt proteins in nuclear (N) and cytoplasmic (C) fractions of HepG2 cells, 48 h post-transfection of the corresponding expression vectors. Lamin A expression validates the subcellular fractionation and ERK1 the total amount of loaded proteins. (**C**), HepG2 cells were transfected with ERE-tk-Luc and CMV-βgal reporter genes together with either pCR ERα WT or pCR ERα SYN-opt expression vectors and then treated for 24 h with increasing doses of E2 (0 to 10 nM). Luciferase activity was normalized to β-galactosidase activity. Data corresponds to the mean values ± SEM from at least five separate transfection experiments and are expressed as fold change from non-treated control. (**D**), HepG2 cells were transfected with C3-Luc, AP1-luc or SP1-luc reporter genes together with the CMV-βgal internal control and either pCR3.1 (empty vector), pCR ERα WT or pCR ERα SYN-opt expression vectors and then treated for 24 h with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle. Luciferase activity was normalized to β-galactosidase activity. Data corresponds to the mean values ± SEM from at least three separate transfection experiments and are expressed as fold change from vehicle-treated pCR3.1 control. Columns with different superscript differ significantly ($p<0,05$, Student's t-test).

3.2. Codon usage optimization of ERα coding sequence by synonymous mutations (ERα SYN-opt) impacts the behaviors of ERα on ERE-, AP1- and SP1-driven reporter genes

ERα SYN-opt expression, localization and transcriptional activity were first characterized following transient transfection in HepG2 cells. This cell line was chosen for its differentiated phenotype, providing an AF1 permissive context with no endogenous ERα expression, as previously described (Huet et al., 2008; Mérot et al., 2004). As illustrated in Figures 2A and 2B, ERα SYN-opt protein is properly expressed in HepG2 and recognized by antibodies targeting both ERα N- and C-terminal domains. As ERα WT, it displays a major nuclear localization with almost similar expression level. The transactivation efficiency of ERα SYN-opt protein was first assessed on an ERE-driven luciferase reporter gene (ERE-tk-Luc) in the presence of increasing E2 concentration. Results show a dose response curve similar to that observed with ERα WT with an EC50 of approximately 0.1 nM of E2 (Figure 2C). The transactivation activity of ERα WT and SYN-opt proteins was then analyzed in the presence or absence of different ligands, including E2 (10 nM), the Selective Estrogen Modulator (SERM) 4-hydroxytamoxifen (OHT, 1 µM) and the Selective Receptor Degrader (SERD) ICI 182-780 (ICI, 100 nM). The ERE-driven luciferase reporter gene used was the C3-Luc because of its high sensitivity to the agonist activity of OHT. Interestingly, ERα SYN-opt exhibits lower ligand-independent activity than ERα WT on the C3-Luc reporter gene (Figure 2D). This activity was then induced nearly 3- or 5-fold in presence of OHT or E2 respectively, whereas it was completely inhibited by ICI. Besides binding directly to ERE-driven genes, ERα can also regulate gene expression indirectly through protein-protein interactions with other transcriptional regulators such as AP1 and SP1. OHT and ICI were previously described as potent agonists on this ERα tethering pathway (Paech et al., 1997; Schultz et al., 2005; Stender et al., 2010). We therefore compared the transactivation activity of ERα WT and SYN-opt proteins on AP1-Luc and SP1-Luc reporter genes after treatment of cells with E2, OHT or ICI (Figure 2D). As expected, ERα WT was capable to transactivate both reporter genes after an OHT or ICI treatment and E2 treatment was inefficient on the AP1-Luc. Surprisingly, ERα SYN-opt protein was unable to mediate a transcriptional activation of both reporter genes, regardless of the ligand used. Slight residual activity was observed with ICI treatment only. Thus, synonymous mutations abolish ERα transactivation on AP1 and SP1 reporter genes.

In conclusion, these results show that adapting the codon usage frequency of ERα by synonymous mutations impacts the ability of ERα to transactivate both classical (direct ERE binding) and nonclassical (tethering) reporter target genes.

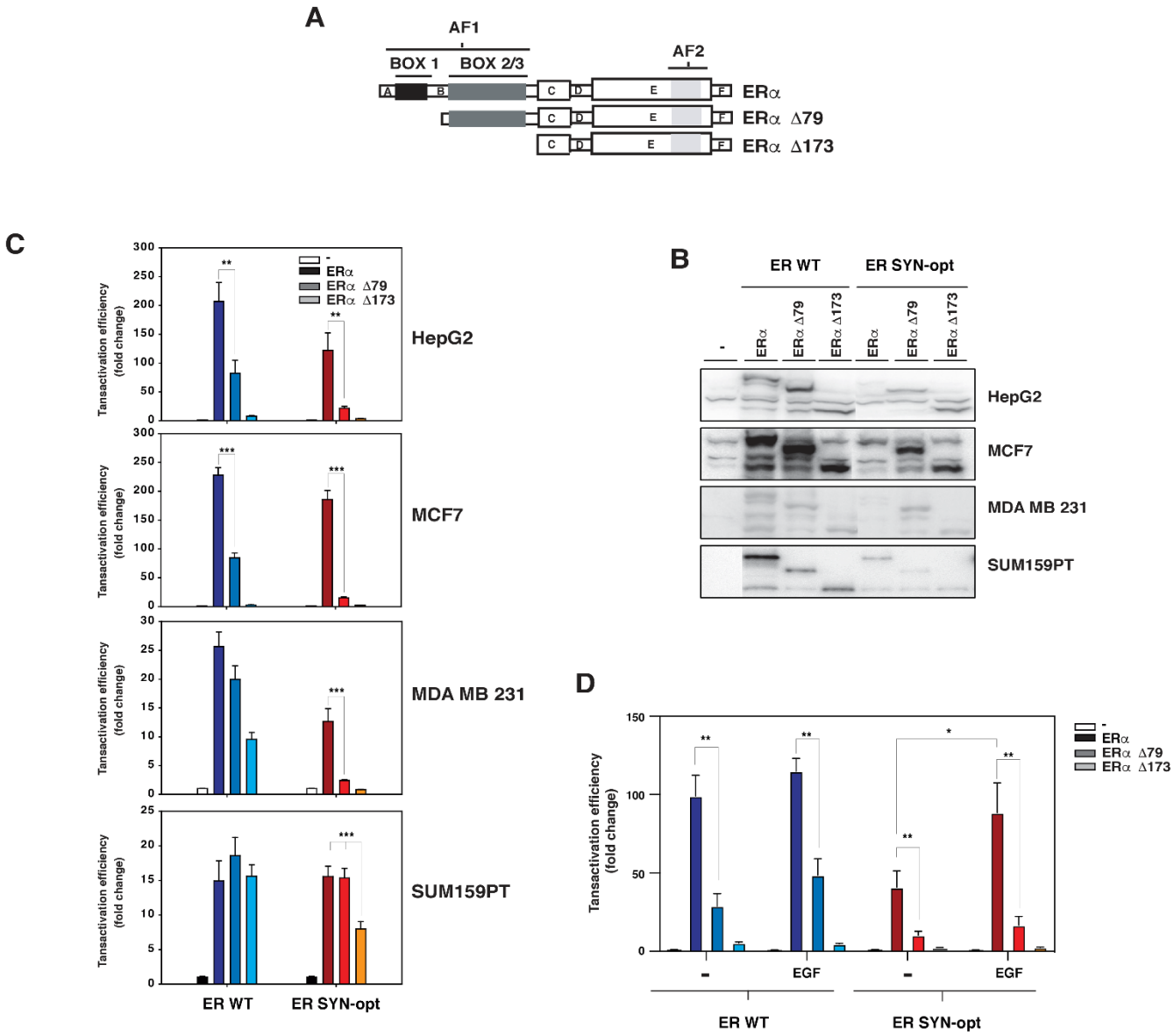


Figure 3: ERα SYN-opt protein displays an enhanced AF1 activity. (**A**), Scheme of the sequences of ERα and the two N-terminal truncated forms, ERα Δ79 and ERα Δ173, used in the study. (**B**), HepG2, MCF7, MDA-MB-231 and SUM159PT cells were transfected with pCR3.1, pCR ERα WT, pCR ERα WT Δ79, pCR ERα WT Δ173, pCR ERα SYN-opt, pCR ERα SYN-opt Δ79 or pCR ERα SYN-opt Δ173 expression vectors. 48 h post-transfection, expression of the different ERα forms was measured by western blot. (**C**), These same cell lines were transfected with the expression vectors previously described together with the C3-Luc and CMV-βgal reporter genes and then treated with 10 nM E2. (**D**), HepG2 cells were transfected with the expression vectors previously described together with the C3-Luc and CMV-βgal reporter genes, and then treated with 10 nM E2 in the presence or the absence of 50 ng/mL EGF. For all experiments, luciferase activity was normalized to β-galactosidase activity. Data correspond to the mean values ± SEM from at least three separate transfection experiments and are expressed as fold change from pCR3.1 control (*p-value <0.05, **p-value <0.01 and ***p-value <0.001, Student's t-test).

3.3. ERα SYN-opt protein enhances the transcriptional activity mediated by AF1

As previously mentioned, the transcriptional activity of ERα relies on two transactivation functions, AF1 and AF2 located in the N- and C-terminal domains of the protein, respectively. Specifically active in differentiated cells (Huet et al., 2008; Mérot et al., 2004), the AF1 has several different structural and functional units referred to as boxes 1 to 3 (Figure 3A). We therefore evaluated the respective contribution exerted by both AF2 and AF1 boxes toward the transcriptional activity of ERα SYN-opt protein. For that purpose, we generated for each ERα WT and SYN-opt proteins, two N-terminal deleted forms, ERα Δ79 (deletion of AF1 box 1) and ERα Δ173 (additional deletion of AF1 boxes 2/3). We then compared the transcriptional activity of these deleted forms with that of the full-length ERα in different cell lines with various differentiation stage and AF1 permissiveness (Huet et al., 2008, 2009; Mérot et al., 2004). A similar activity measured for full-length ERα and ERα Δ79 indicates the absence of an AF1 box 1 activity when a strict AF2-mediated transcriptional activity is revealed by the same activity for full-length ERα and ERα Δ173. The expression of the different forms was first controlled by Western blotting (Figure 3B). The full-length and deleted forms of ERα SYN-opt were generally slightly less expressed that those of ERα WT. In the strict AF1-permissive cell lines HepG2 and MCF7, ERα WT activity is mediated by both AF1 box 1 and box 2/3 as expected (Figure 3C). Interestingly, ERα SYN-opt favors the contribution of AF1 box 1 in these cells, since the difference of activity between the full-length and Δ79 form is much higher in the context of the SYN-opt receptor as compared to the WT ERα. In MDA-MB-231 cells, ERα WT activity does not require the AF1 box 1 activity while ERα SYN-opt activity remains primarily driven by this subdomain (Figure 3C). Finally, whereas the transcriptional activity of ERα WT relies strictly on the AF2 function in the less-differentiated cell line SUM159PT, ERα SYN-opt additionally exhibits an AF1 boxes 2/3 activity, but AF1 box 1 activity remains quenched in these cells.

AF1 activity is also known to be regulated by growth factors such as EGF, through the phosphorylation of the serine residue at position 118 of ERα by the mitogen-activated protein kinase MAPK (Kato et al., 1995). We therefore evaluated the impact of EGF treatment on the transcriptional activity of ERα WT and SYN-opt proteins. Results clearly show that ERα SYN-opt is more sensitive to EGF stimulation than ERα WT (Figure 3D).

Altogether, these results demonstrate that optimizing the codon usage frequency of ERα by synonymous mutations improves the transcriptional activity of AF1 function.

3.4. ERα SYN-opt protein improves tamoxifen-mediated repression in the presence of corepressor

SERMs such as OHT act as either agonists or antagonists of ERα functions through the recruitment of coactivators or corepressors, respectively, in a tissue-, cell- and promoter-specific manner (Yongfeng Shang and Brown, 2002; Smith et al., 1997a). In an attempt to study the consequences of ERα codon optimization on the SERM activity of OHT, MCF7 were transiently transfected with either ERα WT or ERα SYN-opt expression vectors in the presence or absence of expression vectors for SRC1 (Steroid Receptor coactivator-1) or NCoR1 (Nuclear Receptor Corepressor-1) and the activity of the reporter gene C3-Luc was then measured after treatment or not with E2 or OHT. The results are shown in Figure 4. Whereas overexpression

of SRC1 similarly enhances the agonist activity of OHT on ERα WT and SYN-opt proteins, overexpression of NCoR1 induces antagonist activity of OHT in the presence of ERα SYN-opt protein, only. OHT indeed represses the basal activity of the reporter gene only when ERα SYN-opt protein is co-expressed with NCoR1. It should be noted that the repression of the basal activity was also observed in absence of ligand. This result indicates that optimizing the codon usage frequency of ERα by synonymous mutations enhances the ability of ERα to repress transcription in the presence of OHT, presumably through improved corepressor recruitment.

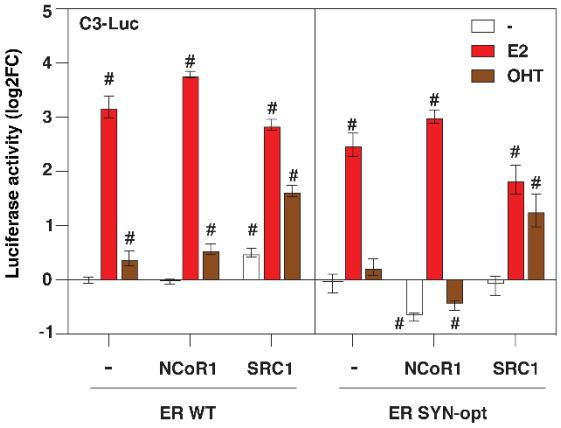


Figure 4: ERα SYN-opt protein improves tamoxifen-mediated repression in the presence of corepressor. MCF7 cells were transfected with pCR ERα WT or pCR ERα SYN-opt in the presence or the absence of pCR3.1, pCR-SRC1 or pCR-NCoR1 expression vectors together with the C3-Luc and CMV-βgal reporter genes and then treated or not for 24 h with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle. Luciferase activity was normalized to β-galactosidase activity. Data correspond to the mean values ± SEM from at least three separate transfection experiments and are expressed in log2 fold change (log2FC) from vehicle-treated pCR3.1 control. Columns with # superscript differ significantly from non-treated pCR3.1 control ($p<0.05$, Student's t-test).

3.5. E2 anti-apoptotic effect is abolished in MCF7 expressing ERα SYN-opt protein, resulting in a weaker estrogen-induced proliferation

To investigate the impact of ERα SYN-opt protein on hormone-dependent breast cancer cells, we generated MCF7 subclones using a tetracycline-inducible vector system to stably expressed ERα SYN-opt protein. We selected two clones, MCF7 ERα SYN-opt clone 1 and 2. MCF7 subclones expressing either the empty expression vector (MCF7 control) or the expression vector encoding ERα WT (MCF7 ERα WT) were also established in parallel and used as controls. ERα WT and SYN-opt proteins were fused to the GFP tag in their N-terminal part in order to discriminate artificially expressed forms of ERα from endogenous ERα in MCF7 cells. Tetracycline-induced expression of GFP-ERα WT and SYN-opt proteins was first controlled by Western blot and immunofluorescence approaches (Figure 5A and 5B). Their expression was predominantly localized in the nucleus, like endogenous ERα, with an approximately 1.5-fold higher level for GFP-ERα WT than for GFP-ERα SYN-opt protein. Interestingly, while the nuclear expression level of GFP-ERα WT varies greatly from one cell to another in accordance with what is observed with the endogenous ERα, GFP-ERα SYN-opt shows

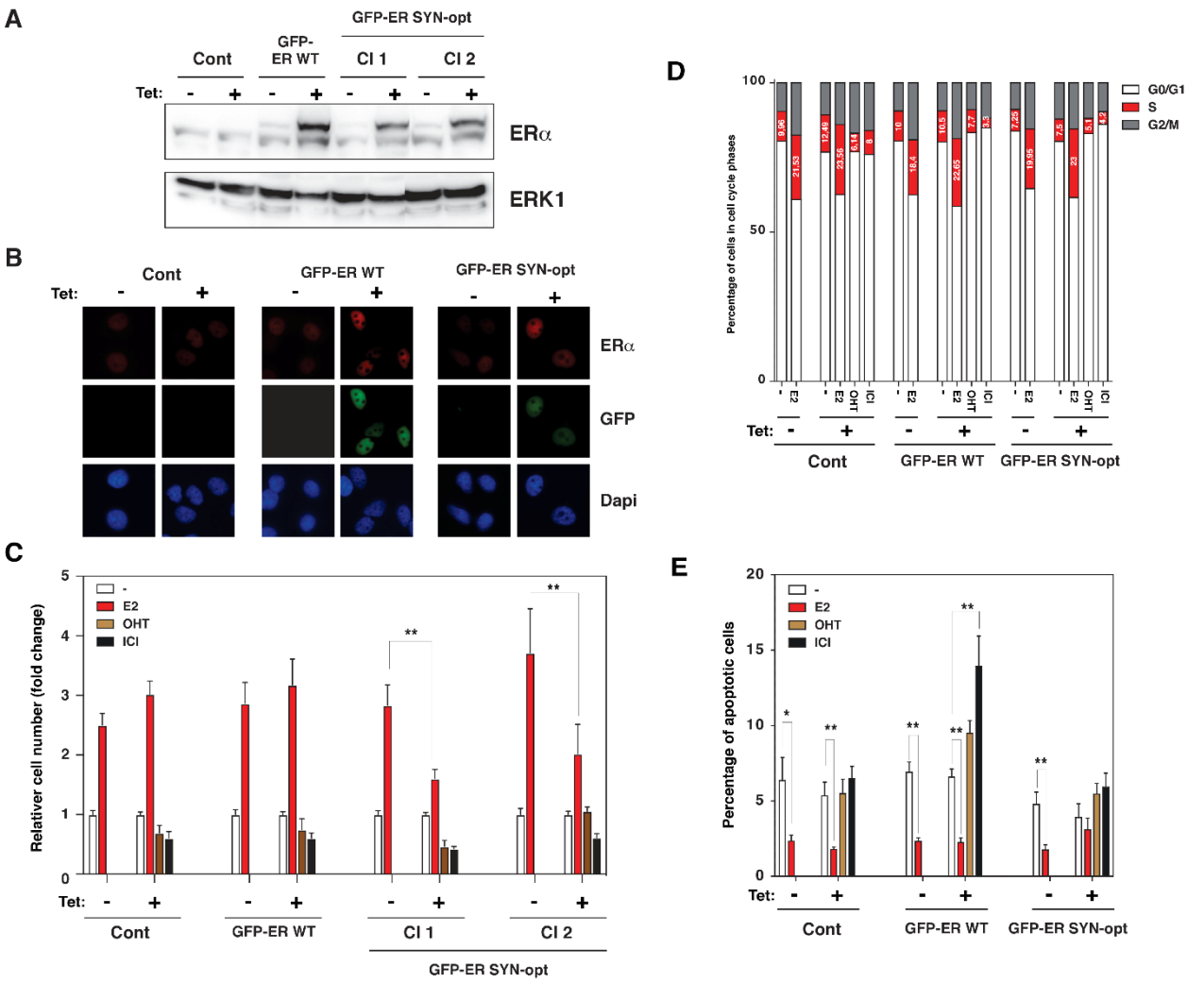


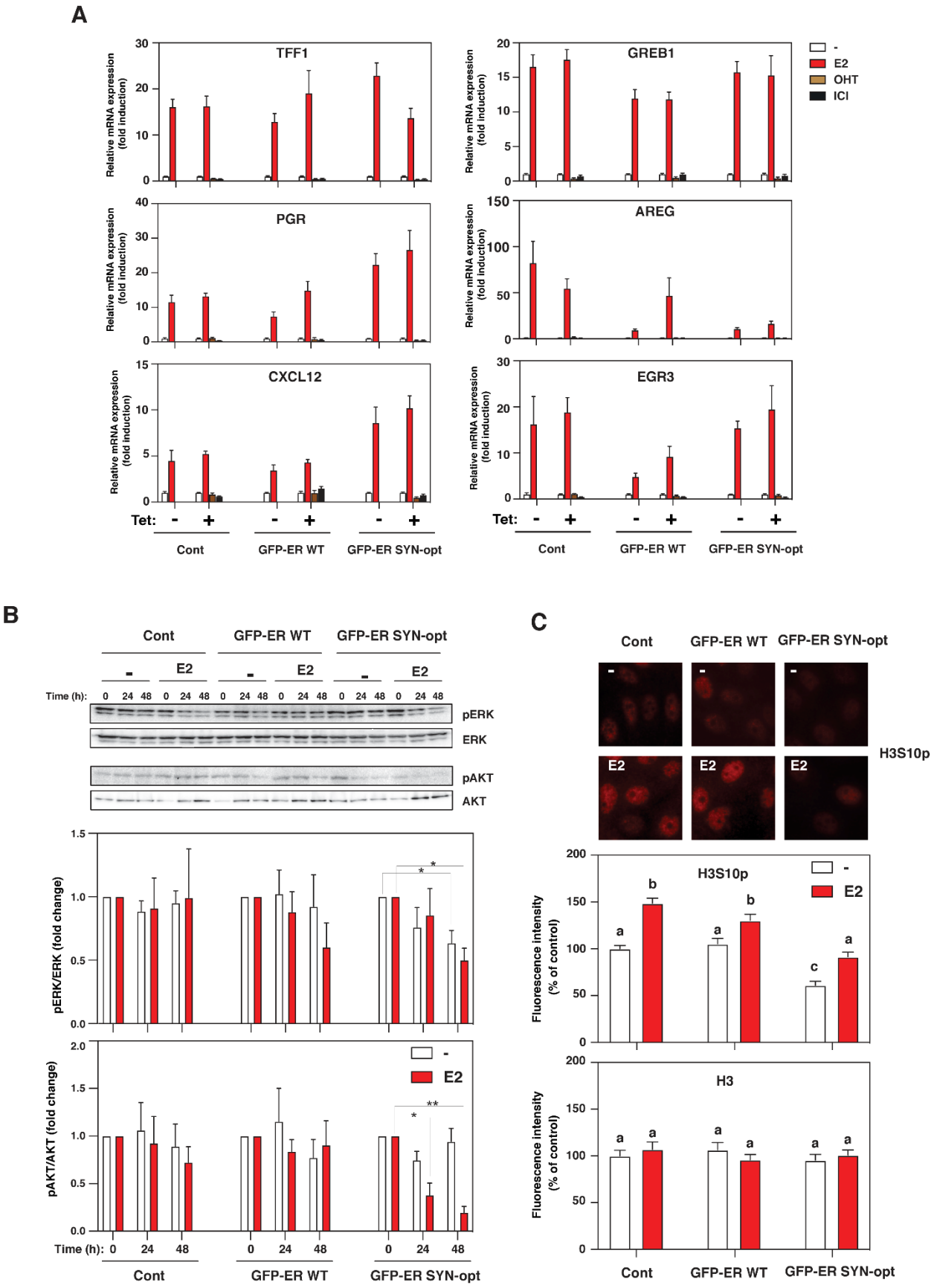
Figure 5: ERα SYN-opt impairs estrogen-induced cell proliferation of MCF7 cells by abolishing the anti-apoptotic effect of E2. (**A**), Control, GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones were treated or not with tetracycline. 48 h later, the expression of the different ERα proteins was analyzed by western blot. ERK1 was used as loading control. (**B**), GFP and immunofluorescent detection of ERα, 48 h after treatment or not with tetracycline of the control, GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones. Nuclei were stained with DAPI. (**C**), MCF7 subclones were treated or not with tetracycline together with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle for 6 days. The cells were counted manually upon trypsinization. Data correspond to the mean values ± SEM from at least three separate experiments and are expressed as fold change from cells treated with vehicle. E2 treatment values with ** superscript differ significantly between tetracycline-treated and non-treated value ($p<0.01$, Student's t-test). (**D**), MCF7 subclones were treated or not for 48 h with tetracycline together with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle during the last 24 h. Flow cytometry experiments were performed to determine the percentage of cells in each cell cycle phase. Data correspond to the mean values from at least three separate experiments. (**E**), MCF7 subclones were treated or not for 48 h with tetracycline together with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle during the last 24 h. Percentage of apoptotic cells was determined by TUNEL staining. Data correspond to the mean values from triplicate experiments ± SEM (**$p<0.01$, Student's t-test).

a very similar expression between the different cells (Supplementary Figure 2). Furthermore, nuclear expression of GFP-ERα SYN-opt remains stable in mitotic cells identified by H3S10 phosphorylation staining whereas that of GFP-ERα WT drops sharply. We then assessed the impact of GFP-tagged protein expression on cell proliferation. For that purpose, MCF7 control, GFP-ERα WT and GFP-ERα SYN-opt cells were cultivated during 6 days in the presence or absence of tetracycline and under estrogen or antiestrogen treatment, before counting the resulting cell population. As expected, E2 induced a 3-fold increase in cell numbers after 6 days of treatment when OHT and ICI slightly inhibited basal cell proliferation. For control and GFP-ERα WT MCF7 cells, tetracycline treatment had no major effect. In contrast, cell counting of GFP-ERα SYN-opt MCF7 subclones clearly shows a 2-fold reduction in cell number following tetracycline treatment indicating that GFP-ERα SYN-opt expression disrupts E2-induced cell proliferation (Figure 5C). Because similar results were obtained with GFP-ERα SYN-opt subclones 1 and 2 on cell proliferation, subsequent experiments were performed on subclone 1 only. Cell number results from the balance between cell cycle entry and apoptosis. These two parameters were therefore measured in order to determine the mechanism responsible for the limited E2-induced proliferation of breast cancer cells following GFP-ERα SYN-opt expression. First, the relative proportion of cells in each phase of the cell cycle was analyzed by flow cytometry (Figure 5D). Results show a similar dynamic in the different cell cycle phases whatever the MCF7 subclones used and the treatment or not with tetracycline. Compared to untreated cells, E2 enhances 2-fold the percentage of cells in S phase and OHT and ICI treatments slightly inhibit it, as expected. In other hands, expression of GFP-ERα WT and GFP-ERα SYN-opt proteins has no impact on the proportion of the different phases of the cell cycle in MCF7. The percentage of apoptotic cells was then determined by TUNEL staining in each condition. In control and GFP-ERα WT MCF7 cells, E2 treatment leads to a decrease in the number of apoptotic cells, whereas antiestrogen treatments have no real effect or weakly stimulate apoptosis (Figure 5E). Interestingly, when GFP-ERα SYN-opt protein was expressed in MCF7 cells, almost no reduction in the number of apoptotic cells was observed in presence of E2. Therefore, E2 loses its anti-apoptotic properties in presence of ERα SYN-opt protein.

In conclusion, these results demonstrate that the low E2-induced proliferation of MCF7 cells expressing ERα SYN-opt protein, is not due to a defect in cell cycle entry but rather to an increased apoptosis.

3.6. Expression of ERα SYN-opt protein in MCF7 cells alters MAPK and AKT signaling pathway

In MCF7 cells, ERα is known to exhibit both genomic and non-genomic effects (Arnal et al., 2017). We therefore undertook the study of these two action modes of ERα in the different MCF7 subclones. First, genomic activity of GFP-ERα forms was investigated by studying the consequences of their expression on the activity of E2-target genes. To do so, cells were treated or not for 48 h with tetracycline in the presence or absence of E2, OHT or ICI and the expression of six main E2-regulated genes, TFF1, PGR, CXCL12, GREB1, AREG and EGR3, was assessed by RT-qPCR (Figure 6A). Results show some variations in gene expression between the different subclones but which were not attributed to tetracycline-induced expression of the proteins of interest.



(legend on the next page)

Figure 6: Expression of ERα SYN-opt in MCF7 impacts MAPK and AKT signaling pathways rather than the E2-target gene expression. (**A**), Control, GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones were treated or not for 48 h with tetracycline together with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle during the last 24 h. The expression of the E2-target genes, TFF1, PGR, CXCL12, GREB1, AREG and EGR3, were quantified using RT-qPCR. Expression of E2-target genes was normalized to TBP expression. Data correspond to the mean values ± SEM from at least three separate experiments and are expressed as fold change from vehicle-treated control. (**B**), MCF7 subclones were treated with tetracycline in the presence or absence of 10 nM E2. Cells were harvested at 0, 24 and 48 h of treatment and western blots were then performed to analyze phospho-ERK (pERK), ERK, phospho-AKT (pAKT) and AKT expression. Histograms represent the mean ± SEM of pERK/ERK or pAKT/AKT ratios from four separate experiments. Results were expressed as fold change from non-treated control of each MCF7 subclones (*p-value <0.05 and **p-value <0.01, Student's t-test). (**C**), MCF7 subclones were treated for 48 h with tetracycline in the presence or absence of 10 nM E2 during the last 24 h. On the top, immunoflurescence images of H3S10p. On the bottom, densitometry quantification of the immunofluorescence images of H3S10p and total H3 (H3) expressed as percentage of the intensity measured in untreated (-E2) control cells. Values represent the mean ± SEM. Columns with different superscripts differ significantly (n = 20, p-value <0.05, Student's t-test).

While the expression of the studied genes is induced by E2 and slightly repressed in presence of OHT and ICI, this profile is indeed clearly not affected by the expression of GFP-ERα WT and GFP-ERα SYN-opt proteins. These data suggest that E2 treatment probably results in the same regulation of gene expression by ERα WT and SYN-opt proteins. Non-genomic activity of GFP-ERα forms was then investigated by examining the phosphorylation status of ERK and AKT by Western blot in control, GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones after 24 and 48 h treatment with tetracycline alone or with E2. No significant alteration of ERK nor AKT phosphorylation was observed in control and GFP-ERα WT MCF7 cells following tetracycline and E2 treatments. On the contrary, addition of tetracycline results in a decrease of ERK and AKT phosphorylation in GFP-ERα SYN-opt MCF7 subclone which was E2-independent and E2-dependent, respectively. Accordingly, in contrast to ERα WT, ERα SYN-opt represses MAPK and AKT signaling pathway (Figure 6B). Although non-genomic activity of ERα is primarily characterized by processes affecting components of signal transduction pathways, the subsequent cellular response may ultimately be a regulation of gene expression through the phosphorylation of transcription factors, coregulators or structural components of chromatin. Notably, histone H3 phosphorylation at serine 10 (H3S10p) is one of the main nucleosome response to mitogen or growth factor signaling pathways through the activation of MSK1 (Chadee et al., 1999). Since steroid hormones including progestin and estrogen have been shown to induce this signaling pathway (Reyes et al., 2014; Vicent et al., 2006), we investigated histone H3 phosphorylation status at serine 10 by immunofluorescence in the different MCF7 subclones treated or not 24 h with E2. Results show that E2 increases H3S10 phosphorylation in all three MCF7 subclones, nevertheless with a lower basal level of

H3S10p in GFP-ERα SYN-opt MCF7 cells than in control and GFP-ERα WT MCF7 cells (Figure 6C). Repression of MAPK signaling pathway in GFP-ERα SYN-opt MCF7 cells converge at the chromatin through repression of H3S10 phosphorylation.

3.7. ERα SYN-opt interacts more with corepressors and less with kinases than ERα WT in MCF7 cells

Because ERα SYN-opt protein impacts E2 signaling pathway in MCF7 cells, we extended our study to analyze the interaction of ERα with some of its main partners such as the steroid coactivator P160 family members, SRC1 and SRC3/AIB1, the corepressors NCoR1 and SMRT and the kinases c-src and PI3K. After treating GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones with the different ligands E2, OHT or ICI, protein interactions were measured by proximity ligation assay (PLA) using antibodies targeting GFP-ERα protein through the GFP tag and ERα 's partners. Confirming Western blot analyses, quantification of the GFP fluorescence intensity show that GFP-ERα SYN-opt expression is almost 1.5-fold lower than that of GFP-ERα WT and that short-term treatment with ligands has little impact on the protein expression level (Figure 7C). As expected, interaction profiles of GFP-ERα WT with the coactivators, corepressors and kinases were similar to those previously described with the endogenous receptor (Jehanno et al., 2021). Indeed, we have previously shown that whereas ERα interaction with P160 family coactivators is not further enhanced by E2, it is strongly inhibited in presence of ICI. Regarding the interaction of ERα with corepressors, treatment with E2 as well as ICI was shown to abolish much of the interaction, whereas OHT retained it. Finally, ERα/kinase complex formation was not or only partially affected by ligands (Jehanno et al., 2021). The similarity of GFP-ERα WT interaction profiles with those of the endogenous receptor indicates that fusion of GFP to the N-terminal domain of ERα does not affect interactions of the receptor with the tested partners. Interestingly, notable differences were observed in the GFP-ERα SYN-opt interaction profiles with its partners (Figure 7B). Indeed, unlike GFP-ERα WT, GFP-ERα SYN-opt interacts less with SCR1 and SRC3 coactivators in a basal state, with no further change in interactions when cells are treated with the antagonists OHT and ICI, but strongly increases its coactivator recruitment after E2 stimulation. Concerning the interactions with corepressors and kinases, GFP-ERα SYN-opt shows the same profile as GFP-ERα WT with respect to the ligand effect but respectively more ERα/corepressor complexes and less ERα/kinase complexes were detected for GFP-ERα SYN-opt protein. These differences in interaction intensity cannot fully be attributed to differences in the expression level of the corresponding proteins involved in the interaction (Figure 7C and 7D). Therefore, taken together, these results indicate that optimizing the codon usage frequency of ERα by synonymous mutations affects how ERα protein interacts with its partners, specifically favoring interaction with corepressors over interaction with kinases.

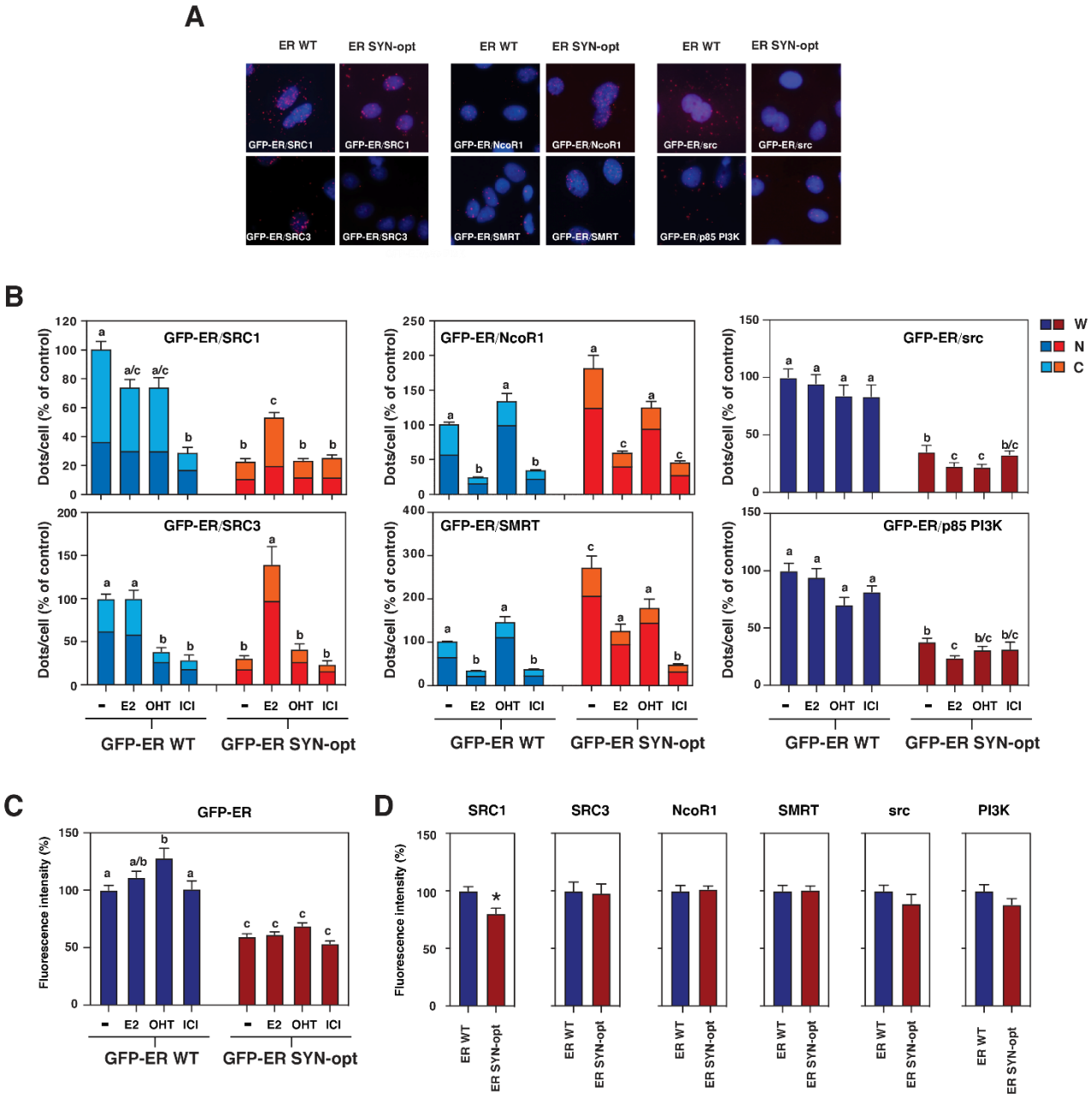


Figure 7: ERα SYN-opt interacts more with corepressors and less with kinases than ERα WT in MCF7 cells. GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones were treated 48 h with tetracycline. At the end of treatment, cells were stimulated with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle during 60 min for the study of ERα/cofactor interactions and during 10 min for the study of ERα/kinase interactions. GFP-ERα WT or GFP-ERα SYN-opt complexes with coactivators (SRC1 and SRC3), corepressors (NCoR1 and SMRT) and kinases (c-src and PI3K) were detected by PLA. (**A**), Representative pictures of the experiments are shown with DAPI-stained nuclei. (**B**), Quantification of the number of dots/cell and dots/nucleus was performed using ImageJ software and was then expressed as a percentage of the number of dots/cell measured in vehicle-treated GFP-ERα WT MCF7 subclone. The respective proportion of complexes in the nucleus (N) and in the cytoplasm (C) is shown. Values represent the mean ± SEM. Columns with different superscripts differ significantly (n = 20 to 40, p-value <0.001, Student's t-test). (**C**), Quantification of the expression level of GFP-

ERα WT and GFP-ERα SYN-opt through GFP fluorescence, 60 min after ligand treatments. Columns with different superscripts differ significantly (n = 20 to 40, p-value <0.001, Student's t-test). (**D**), Quantification of the expression level of SRC1, SRC3, NCoR1, SMRT, c-src and PI3K in GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones after immunofluorescence detection. Values represent the mean ± SEM (n =10, *p-value <0.05, Student's t-test).

3.8. Limited digestions of ERα WT and ERα SYN-opt proteins suggest conformational differences

After assessing ERα SYN-opt functional properties in MCF7 cells, a potential modification of its conformation was investigated by performing limited digestion. For that purpose, increasing amounts of chymotrypsin or trypsin were incubated with the same quantity of cell lysate from GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones. The result of these limited digestions was then analyzed by western blotting using C-terminal ERα and GFP antibodies. As shown on Figure 8, digestion of GFP-ERα WT and GFP-ERα SYN-opt proteins resulted in some differences regarding digestion kinetics and weight of digested peptides. These results indicate that limited digestions produced different fragments between GFP-ERα WT and GFP-ERα SYN-opt, revealing an altered conformation of GFP-ERα SYN-opt compared to GFP-ERα WT.

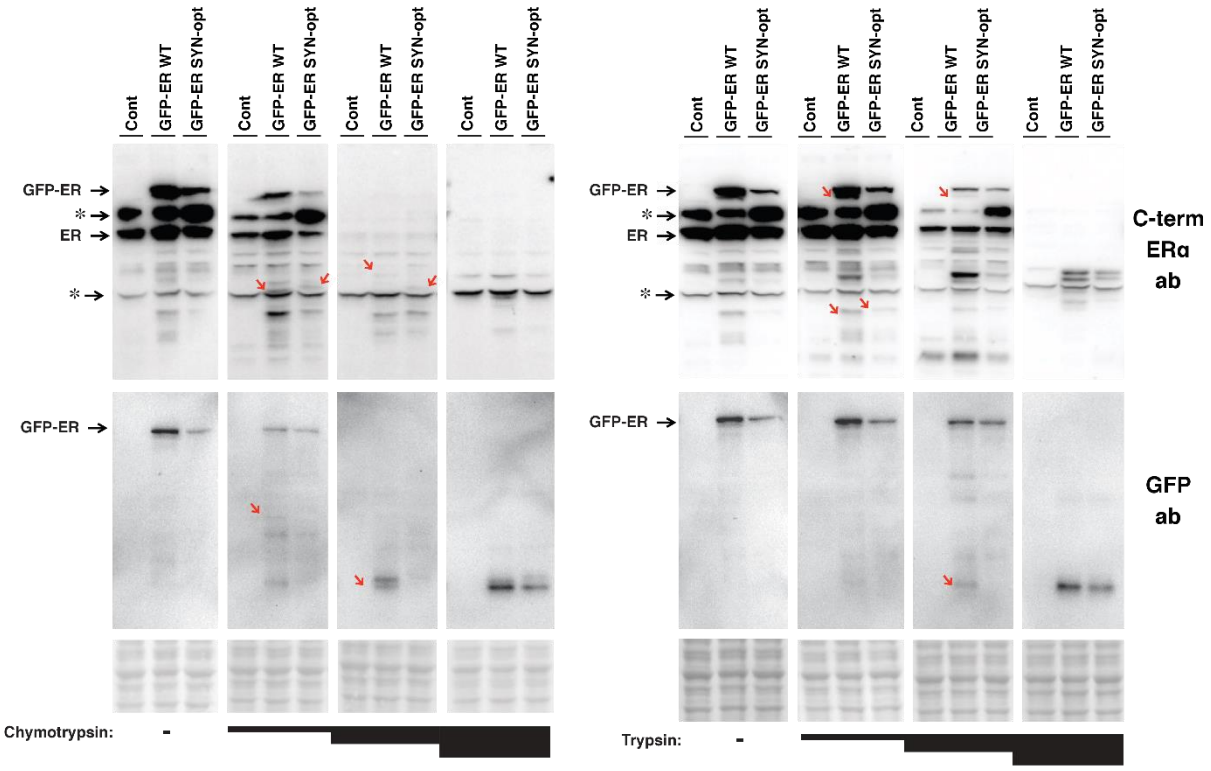


Figure 8: Limited digestions of ERα WT and ERα SYN-opt proteins suggest conformational differences. Control, GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones were treated for 48 h with tetracycline. Limited digestions of cell lysates by increasing concentrations of trypsin or chymotrypsin (0.2 to 1.6 ng/µL) were analyzed by western blot using C-terminal ERα and GFP antibodies. Ponceau red staining was used to validate the homogeneity of protein amount for each sample. * indicates non-specific band. Red arrows indicate differences observed between digested samples.

4. Discussion

The proliferation of ERα-positive breast cancer cells can be inhibited by using antiestrogens. Unfortunately, endocrine resistance often occurs, resulting in the estrogen-independent development of cancerous cells. Many works aim to understand the mechanisms leading to this hormonal escape and identified the role played notably by the imbalance of coregulators and growth factor signaling pathways. It has indeed been shown that the relative expression of coactivators and corepressors modulates the activity of the antiestrogen tamoxifen towards an ERα-agonism or antagonism, explaining why tamoxifen exerts an agonist effect in the uterus where the level of coactivators is increased compared to breast cancer cells where tamoxifen acts as an ERα-antagonist (Smith et al., 1997b). The overexpression of coactivators such as SRC1 and the repression of corepressors such as NCoR1 in breast cancer cells have then been associated with tamoxifen resistance (Altwegg and Vadlamudi, 2021). As intracellular kinases are able to activate ERα by phosphorylation, the overactivation of components of the MAPK or PI3K pathways plays a role in hormonal escape, by promoting the ligand-independent activation of ERα. This dysregulated activation results from mutations that are frequent in ERα-positive breast cancers, or the overexpression of growth factor receptors that activate such pathways, such as EGFR, which has been linked to endocrine resistance as well (Hanker et al., 2020b).

In healthy breast tissue, ERα is expressed in a small number of luminal cells, a subset of differentiated epithelial cells that are not proliferating, as assessed by a lack of co-staining with the proliferating marker Ki67 (R. B. Clarke et al., 1997; Shoker et al., 1999). In this context, ERα exerts its estrogen-induced mitogen activity by a paracrine mechanism, favoring the proliferation of surrounding ERα-negative cells only (Mallepell et al., 2006). In ERα-positive breast cancers however, an increased proportion of ERα-expressing cells appears to be proliferating, due to a shift from a paracrine to an autocrine regulation of proliferation by ERα (R. B. Clarke et al., 1997; Tan et al., 2009). This change from quiescent to proliferating ERα-expressing cells during tumor transformation results in different modifications in cell activity that can lead to changes in functional properties of ERα explaining endocrine resistance.

As previous works highlighted tRNA pool variations between differentiated and proliferating cells (Gingold et al., 2014), we hypothesized that this could occur in proliferative ERα-positive breast cancer cells compared to non-cancerous ERα-expressing cells that remain quiescent. These tRNA pool modification could then alter the translation speed of ERα, resulting in an alteration of its co-translational folding. This would ultimately lead to a modification of its conformation, explaining the dysregulation of ERα functions in breast cancer cells, participating to hormonal escape. This tRNA pool variation was correlated to a different codon usage bias in genes specifically expressed in differentiated or proliferating cells, suggesting an adequacy between codon usage and tRNA pool to enable the correct production of proteins in specific cell states (Gingold et al., 2014). In this case, the codons used to produce ERα would not anymore be adapted for its production in dedifferentiated and proliferating cells.

A synonymous version of ERα was then generated by adapting the frequency of its codon usage to that observed in genes specifically expressed in proliferative cells, to investigate if this sequence optimization affects ERα functional properties in cancerous cells. Several of our results clearly demonstrate that codon adaptation to expression in proliferating cell indeed restores functional properties of ERα normally observed in differentiated cells: first, in MCF7, ERα nuclear expression is high in quiescent cells and is down regulated during cell cycle (Penot et al., 2005), resulting in heterogeneous expression of ERα in asynchronous cells. Adaptation of ERα codons to its expression in proliferating cells leads to the homogenization of ERα expression in MCF7 cells and no down-regulation in mitotic cells. Second, in differentiated cells, ERα exerts a genomic activity mainly driven by its AF1 function, especially its box 1 structural unit (Huet et al., 2008; Mérot et al., 2004). Optimizing the codon usage of ERα for proliferating cells improved the transcriptional activity mediated by AF1 box 1 even in dedifferentiated cells. Third, hormone-dependent breast cancers are more differentiated and less invasive than the other types of breast cancers and generally respond well to endocrine therapy (Jordan and O'Malley, 2007). In particular, ERα transcriptional activity in breast cancer cells responding to endocrine therapy is repressed by the 4-hydroxytamoxifen through corepressor recruitment by ERα (Y. Shang et al., 2000; Yongfeng Shang and Brown, 2002). As mentioned previously, one mechanism of endocrine resistance is an imbalance in ERα recruitment between coactivators and corepressors in favor of coactivators. We clearly show that the ability of OHT to repress transcription was increased after ERα codon adaptation, through improved corepressor recruitment, which was further validated by PLA in MCF7 cells. Finally, non-genomic activity of ERα, which leads to the phosphorylation of intracellular kinases involved in MAPK or AKT signaling pathways, is mediated by the formation of complexes between ERα and kinases such as c-src or PI3K. These complexes are increased in aggressive breast cancer cells, playing a role in endocrine resistance (Jehanno et al., 2021; Poulard et al., 2012). We demonstrate that codon optimization of ERα reduces interaction of ERα with kinases and represses MAPK and AKT signaling pathways when the protein is stably expressed in MCF7. All these changes in ERα activity are likely related to changes directly affecting ERα protein, because only the codon sequence of ERα has been modified. Since limited digestions with trypsin or chymotrypsin show differences between WT and codon optimized forms of ERα, the changes in ERα activity are likely based on differences in ERα conformation.

The impact of ERα SYN-opt expression on MCF7 cell fate was also investigated in this study. While cells appear to enter S-phase normally after estrogenic stimulation, the cell population increases little due to the abolition of the anti-apoptotic effect of E2. There is no clear information regarding the apoptotic activity of ERα-positive epithelial cells in healthy breast tissue but it is noteworthy that ERα-positive cells are few in number and represent about 10% of the total luminal epithelial cell population in human, which may suggest some sensitivity to apoptosis.

5. Conclusion

The results obtained in this study attest that modifying the codons used to produce ERα can lead to functional modifications, resulting from alterations of its conformation. In this case, ERα SYN-opt displays a more repressive phenotype than ERα WT, meaning that adapting its codons in agreement with the frequencies observed in genes specifically expressed in proliferating cells enhanced the phenotype of ERα in differentiated epithelial cells. This suggests that the functional modifications of ERα in transformed cells are linked to a conformational modification of the receptor, which highlights a translational regulation of ERα conformation linked to its nucleotide sequence and cell state. This work sheds light on a role played by codon usage in the translational process regulation that could participate to endocrine resistance in ERα-positive breast cancer.

Acknowledgements

This work was supported by the University of Rennes 1, Inserm, CNRS and the Ligue Contre le Cancer.

Author contributions

G.F. conceived the project. L.C., F.P., E.J., P.G., C.T. and G.F. conducted the experiments. T.F.-C., R.M., M.M., F.P. and. and D.M. contributed to the data analysis. T.F.-C., D.M. and G.F. wrote the manuscript with input from all authors.

Supporting information

Table S1: List of primers used to detect target genes by RT-qPCR.

Target gene	**Forward primer**	**Reverse Primer**
Amphiregulin (AREG)	CCTGGCTATATTGTCGATTCA	GTATTTTCACTTTCCGTCTTGTTTTG
C-X-C motif chemokine ligand 12 (CXCL12)	CACCATTGAGAGGTCGGAAG	AATGAGACCCGTCTTTGCAG
Early growth response 3 (EGR3)	CCTGACAATCTGTACCCCGA	AGTTGGAAGGGGAGTCAAG
Growth regulating estrogen receptor binding 1 (GREB1)	GAGGATGTGGAGTGGAGAC	CAGTACCTCAAAGACCTG
Progesterone receptor (PGR)	GTGCCTATCCTGCCTCTCAATC	CCCGCCGTCGTAACTTTGG
TATA-box binding protein (TBP)	TGCACAGGAGCCAAGAGTGAA	CACATCACAGCTCCCCACCA
Trefoil factor 1 (TFF1)	ACCATGGAGAACAAGGTAA	CCGAGCTCTGGGACTAATCA

A

		Codon usage frequency		
Amino acid	Codon	proliferation	differentiation	ERα
ala	gca	0,32	0,16	0,15
ala	gcc	0,3	0,46	0,51
ala	gcg	0,06	0,23	0,15
ala	gct	0,32	0,15	0,19
arg	aga	0,31	0,1	0,23
arg	agg	0,2	0,16	0,23
arg	cga	0,14	0,08	0,11
arg	cgc	0,11	0,33	0,26
arg	cgg	0,15	0,27	0,11
arg	cgt	0,09	0,06	0,06
asn	aac	0,42	0,75	0,76
asn	aat	0,58	0,25	0,24
asp	gac	0,42	0,75	0,67
asp	gat	0,58	0,25	0,33
cys	tgc	0,4	0,73	0,62
cys	tgt	0,6	0,27	0,38
gln	caa	0,36	0,17	0,16
gln	cag	0,63	0,83	0,84
glu	gaa	0,55	0,25	0,19
glu	gag	0,45	0,75	0,81
gly	gga	0,35	0,14	0,22
gly	ggc	0,25	0,49	0,41
gly	ggg	0,19	0,27	0,22
gly	ggt	0,2	0,1	0,15
his	cac	0,45	0,77	0,64
his	cat	0,55	0,23	0,36
ile	ata	0,23	0,08	0,11
ile	atc	0,34	0,71	0
ile	att	0,43	0,22	0,26
leu	cta	0,1	0,05	0,08
leu	ctc	0,14	0,26	0,22
leu	ctg	0,28	0,49	0,48
leu	ctt	0,17	0,08	0,04
leu	tta	0,13	0,04	0,01
leu	ttg	0,17	0,08	0,16
lys	aaa	0,5	0,26	0,34
lys	aag	0,5	0,74	0,66
met	atg	1	1	1
phe	ttc	0,43	0,69	0,73
phe	ttt	0,57	0,31	0,27
pro	cca	0,35	0,18	0,17
pro	ccc	0,24	0,39	0,46
pro	ccg	0,06	0,25	0,26
pro	cct	0,34	0,19	0,11
ser	agc	0,19	0,33	0,27
ser	agt	0,2	0,09	0,13
ser	tca	0,19	0,09	0,09
ser	tcc	0,16	0,23	0,2
ser	tcg	0,04	0,15	0,07
ser	tct	0,22	0,11	0,24
thr	aca	0,32	0,18	0,16
thr	acc	0,27	0,45	0,48
thr	acg	0,08	0,23	0,16
thr	act	0,33	0,14	0,20
trp	tgg	1	1	1
tyr	tac	0,41	0,74	0,70
tyr	tat	0,59	0,26	0,30
val	gta	0,18	0,08	0,04
val	gtc	0,16	0,3	0,27
val	gtg	0,4	0,52	0,65
val	gtt	0,27	0,1	0,04

B

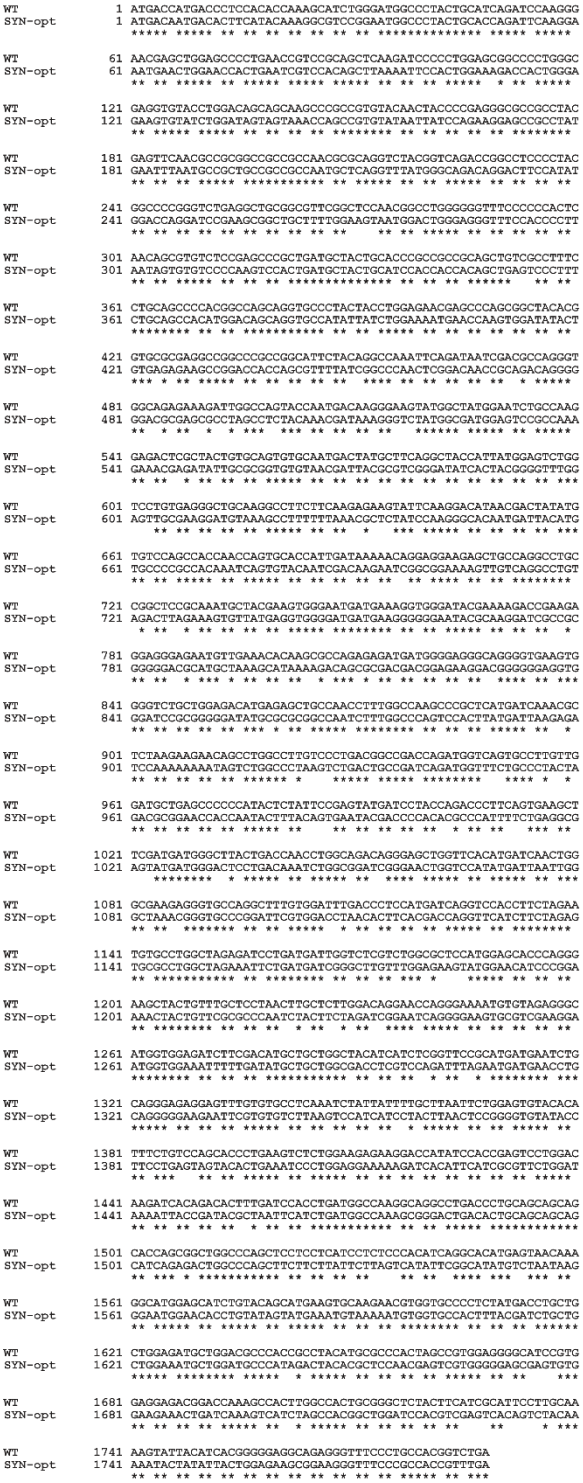


Figure S1: Codon usage optimization of ERα coding sequence by synonymous mutations. (**A**), Table of amino acid-normalized frequency of codon usage in the two functional gene sets, "M phase of mitotic cell cycle" for proliferation and "pattern specification process" for differentiation from the study of Gingold et al. (2014), with the corresponding codon usage frequency for ERα gene. (**B**), Alignment of nucleotide sequences of wild type (WT) and codon usage optimized (SYN-opt) ERα. For optimization of codon usage in ERα coding sequence, a codon was chosen for each amino acid whose frequency of use in proliferation approximates the frequency of the wild-type (WT) codon observed in differentiation.

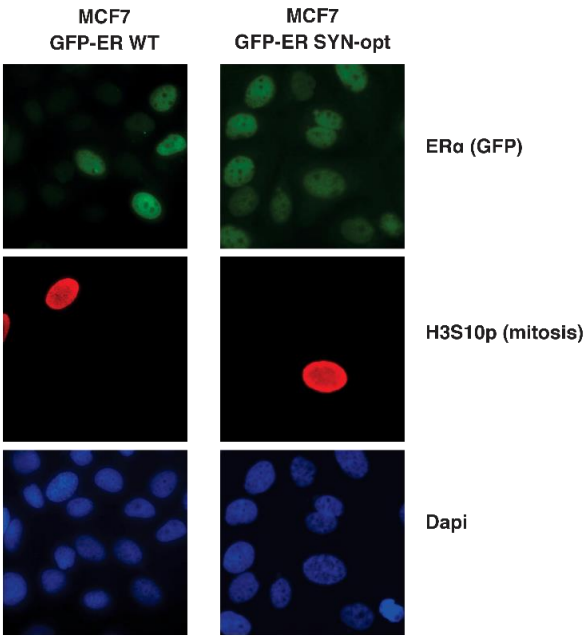


Figure S2: GFP-ERα expression in GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones. GFP-ERα WT and GFP-ERα SYN-opt MCF7 subclones were treated 48 h with tetracycline. Immunofluorescence of H3S10 phosphorylation and GFP fluorescence were measured. Nuclei were stained with DAPI. H3S10 phosphorylation identifies cells in mitosis.

References

Altwegg KA, and Vadlamudi RK (2021) Role of estrogen receptor coregulators in endocrine resistant breast cancer *Explor Target Anti-Tumor Ther* **2**, 385–400.

Arnal J-F, Lenfant F, Metivier R, Flouriot G, Henrion D, Adlanmerini M, … Katzenellenbogen J (2017) Membrane and Nuclear Estrogen Receptor Alpha Actions: From Tissue Specificity to Medical Implications *Physiol Rev* **97**, 1045–1087.

Auchus RJ, and Fuqua SA (1994) Hormone-nuclear receptor interactions in health and disease. The oestrogen receptor *Baillieres Clin Endocrinol Metab* **8**, 433–449.

Beato M (1989) Gene regulation by steroid hormones *Cell* **56**, 335–344.

Chadee DN, Hendzel MJ, Tylipski CP, Allis CD, Bazett-Jones DP, Wright JA, and Davie JR (1999) Increased Ser-10 phosphorylation of histone H3 in mitogen-stimulated and oncogene-transformed mouse fibroblasts *J Biol Chem* **274**, 24914–24920.

Clarke R, Tyson JJ, and Dixon JM (2015) Endocrine resistance in breast cancer--An overview and update *Mol Cell Endocrinol* **418 Pt 3**, 220–234.

Clarke RB (2004) Human breast cell proliferation and its relationship to steroid receptor expression *Climacteric J Int Menopause Soc* **7**, 129–137.

Clarke RB, Howell A, Potten CS, and Anderson E (1997) Dissociation between steroid receptor expression and cell proliferation in the human breast *Cancer Res* **57**, 4987–4991.

Couse JF, and Korach KS (1999) Estrogen receptor null mice: what have we learned and where will they lead us? *Endocr Rev* **20**, 358–417.

Dahlman-Wright K, Cavailles V, Fuqua SA, Jordan VC, Katzenellenbogen JA, Korach KS, … Gustafsson J-A (2006) International Union of Pharmacology. LXIV. Estrogen receptors *Pharmacol Rev* **58**, 773–781.

Edwards DP (2005) Regulation of signal transduction pathways by estrogen and progesterone *Annu Rev Physiol* **67**, 335–376.

Fernández-Calero T, Astrada S, Alberti Á, Horjales S, Arnal JF, Rovira C, … Marin M (2014) The transcriptional activities and cellular localization of the human estrogen receptor alpha are affected by the synonymous Ala87 mutation *J Steroid Biochem Mol Biol* **143**, 99–104.

Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, … Pilpel Y (2014) A dual program for translation regulation in cellular proliferation and differentiation *Cell* **158**, 1281–1292.

Goodarzi H, Nguyen HCB, Zhang S, Dill BD, Molina H, and Tavazoie SF (2016) Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression *Cell* **165**, 1416–1427.

Hanker AB, Sudhan DR, and Arteaga CL (2020a) Overcoming Endocrine Resistance in Breast Cancer *Cancer Cell* **37**, 496–513.

Hanker AB, Sudhan DR, and Arteaga CL (2020b) Overcoming Endocrine Resistance in Breast Cancer *Cancer Cell* **37**, 496–513.

Hanson G, and Coller J (2018) Codon optimality, bias and usage in translation and mRNA decay *Nat Rev Mol Cell Biol* **19**, 20–30.

Horjales S, Cota G, Señorale-Pose M, Rovira C, Román E, Artagaveytia N, … Marín M (2007) Translational machinery and protein folding: Evidence of conformational variants of the estrogen receptor alpha *Arch Biochem Biophys* **467**, 139–143.

Huet G, Mérot Y, Le Dily F, Kern L, Ferrière F, Saligaut C, … Flouriot G (2008) Loss of E-cadherin-mediated cell contacts reduces estrogen receptor alpha (ERα) transcriptional efficiency by affecting the respective contribution exerted by AF1 and AF2 transactivation functions *Biochem Biophys Res Commun* **365**, 304–309.

Huet G, Mérot Y, Percevault F, Tiffoche C, Arnal J-F, Boujrad N, … Flouriot G (2009) Repression of the estrogen receptor-alpha transcriptional activity by the Rho/megakaryoblastic leukemia 1 signaling pathway *J Biol Chem* **284**, 33729–33739.

Jehanno C, Fernandez-Calero T, Habauzit D, Avner S, Percevault F, Jullion E, … Flouriot G (2020) Nuclear accumulation of MKL1 in luminal breast cancer cells impairs genomic activity of ERα and is associated with endocrine resistance *Biochim Biophys Acta BBA - Gene Regul Mech* **1863**, 194507.

Jehanno C, Percevault F, Boujrad N, Le Goff P, Fontaine C, Arnal J-F, … Flouriot G (2021) Nuclear translocation of MRTFA in MCF7 breast cancer cells shifts ERα nuclear/genomic to extra-nuclear/non genomic actions *Mol Cell Endocrinol* **530**, 111282.

Jordan VC, and O'Malley BW (2007) Selective estrogen-receptor modulators and antihormonal resistance in breast cancer *J Clin Oncol Off J Am Soc Clin Oncol* **25**, 5815–5824.

Kato S, Endoh H, Masuhiro Y, Kitamoto T, Uchiyama S, Sasaki H, … Chambon P (1995) Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase *Science* **270**, 1491–1494.

Kerdivel G, Boudot A, Habauzit D, Percevault F, Demay F, Pakdel F, and Flouriot G (2014) Activation of the MKL1/actin signaling pathway induces hormonal escape in estrogen-responsive breast cancer cell lines *Mol Cell Endocrinol* **390**, 34–44.

Mallepell S, Krust A, Chambon P, and Brisken C (2006) Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland *Proc Natl Acad Sci U S A* **103**, 2196–2201.

Marín M, Fernández-Calero T, and Ehrlich R (2017) Protein folding and tRNA biology *Biophys Rev* **9**, 573–588.

Mérot Y, Métivier R, Penot G, Manu D, Saligaut C, Gannon F, … Flouriot G (2004) The Relative Contribution Exerted by AF-1 and AF-2 Transactivation Functions in Estrogen Receptor α Transcriptional Activity Depends upon the Differentiation Stage of the Cell *J Biol Chem* **279**, 26184–26191.

Métivier R, Penot G, Hübner MR, Reid G, Brand H, Kos M, and Gannon F (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter *Cell* **115**, 751–763.

Métivier R, Stark A, Flouriot G, Hübner MR, Brand H, Penot G, … Gannon F (2002) A dynamic structural model for estrogen receptor-alpha activation by ligands, emphasizing the role of interactions between distant A and E domains *Mol Cell* **10**, 1019–1032.

Novoa EM, and Ribas de Pouplana L (2012) Speeding with control: codon usage, tRNAs, and ribosomes *Trends Genet TIG* **28**, 574–581.

Paech K, Webb P, Kuiper GG, Nilsson S, Gustafsson J, Kushner PJ, and Scanlan TS (1997) Differential ligand activation of estrogen receptors ERalpha and ERbeta at AP1 sites *Science* **277**, 1508–1510.

Penot G, Le Péron C, Mérot Y, Grimaud-Fanouillère E, Ferrière F, Boujrad N, … Flouriot G (2005) The human estrogen receptor-alpha isoform hERalpha46 antagonizes the proliferative influence of hERalpha66 in MCF7 breast cancer cells *Endocrinology* **146**, 5474–5484.

Pham TH, Page YL, Percevault F, Ferrière F, Flouriot G, and Pakdel F (2021) Apigenin, a Partial Antagonist of the Estrogen Receptor (ER), Inhibits ER-Positive Breast Cancer Cell Proliferation through Akt/FOXM1 Signaling *Int J Mol Sci* **22**, 470.

Platet N, Cathiard AM, Gleizes M, and Garcia M (2004) Estrogens and their receptors in breast cancer progression: a dual role in cancer proliferation and invasion *Crit Rev Oncol Hematol* **51**, 55–67.

Poulard C, Treilleux I, Lavergne E, Bouchekioua-Bouzaghou K, Goddard-Léon S, Chabaud S, … Le Romancer M (2012) Activation of rapid oestrogen signalling in aggressive human breast cancers *EMBO Mol Med* **4**, 1200–1213.

Reyes D, Ballaré C, Castellano G, Soronellas D, Bagó JR, Blanco J, and Beato M (2014) Activation of mitogen- and stress-activated kinase 1 is required for proliferation of breast cancer cells in response to estrogens or progestins *Oncogene* **33**, 1570–1580.

Rodnina MV (2016) The ribosome in action: Tuning of translational efficiency and protein folding *Protein Sci Publ Protein Soc* **25**, 1390–1406.

Schultz JR, Petz LN, and Nardulli AM (2005) Cell- and ligand-specific regulation of promoters containing activator protein-1 and Sp1 sites by estrogen receptors alpha and beta *J Biol Chem* **280**, 347–354.

Shang Y., Hu X, DiRenzo J, Lazar MA, and Brown M (2000) Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription *Cell* **103**, 843–852.

Shang Yongfeng, and Brown M (2002) Molecular determinants for the tissue specificity of SERMs *Science* **295**, 2465–2468.

Shoker BS, Jarvis C, Clarke RB, Anderson E, Hewlett J, Davies MP, … Sloane JP (1999) Estrogen receptor-positive proliferating cells in the normal and precancerous breast *Am J Pathol* **155**, 1811–1815.

Smith CL, Nawaz Z, and O'Malley BW (1997a) Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen *Mol Endocrinol Baltim Md* **11**, 657–666.

Smith CL, Nawaz Z, and O'Malley BW (1997b) Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen *Mol Endocrinol Baltim Md* **11**, 657–666.

Stender JD, Kim K, Charn TH, Komm B, Chang KCN, Kraus WL, … Katzenellenbogen BS (2010) Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation *Mol Cell Biol* **30**, 3943–3955.

Tan H, Zhong Y, and Pan Z (2009) Autocrine regulation of cell proliferation by estrogen receptor-alpha in estrogen receptor-alpha-positive breast cancer cell lines *BMC Cancer* **9**, 31.

Vicent GP, Ballaré C, Nacht AS, Clausell J, Subtil-Rodríguez A, Quiles I, … Beato M (2006) Induction of progesterone target genes requires activation of Erk and Msk kinases and phosphorylation of histone H3 *Mol Cell* **24**, 367–381.

Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, and Liu Y (2015) Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding *Mol Cell* **59**, 744–754.

2. CONTRIBUTION DU CODE GENETIQUE AU CONTROLE DU NIVEAU D'EXPRESSION DES PROTEINES : L'EXEMPLE DU RECEPTEUR AUX ŒSTROGENES ALPHA

Les résultats présentés ci-après s'inscrivent dans une étude globale réalisée avec différents collaborateurs. Seule la partie à laquelle j'ai participée est alors détaillée ici, celle-ci ayant notamment mené à l'entreprise des travaux présentés au chapitre trois. Considérant leur volume important, certaines données supplémentaires (annotées Supplementary Data) ne sont pas intégrées dans ce manuscrit.

L'expression des gènes en protéines est un processus fondamental dont la dérégulation mène à la majorité des pathologies. Trois niveaux de régulation sont impliqués dans ce phénomène : la transcription, la stabilité de l'ARNm, et la traduction, dont l'équilibre est complexe et toujours le sujet de nombreuses études. Alors que la dégénérescence du code génétique est connue depuis des années, avec l'existence de plusieurs codons correspondant à un même acide aminé, seules de récentes études attribuent un rôle fonctionnel à cette caractéristique. Les codons synonymes ne sont en effet pas utilisés à la même fréquence au sein des séquences codantes, or ce biais d'usage des codons apparaît être différent selon le type de cellules et les phases du cycle cellulaire. L'usage des codons jouant un rôle dans la stabilité de l'ARNm et la traduction, cela suggère que ce paramètre pourrait être impliqué dans la régulation de la production des protéines. Le récepteur aux œstrogènes ERα est le principal acteur de la signalisation des œstrogènes dans de nombreux tissus, et joue un rôle majeur dans le développement du cancer du sein lorsque son expression ou activité sont dérégulées. L'expression d'ERα dépend en effet du tissu, type de cellule, et cycle cellulaire, mais les mécanismes la régulant ne sont pas entièrement compris. Ce récepteur nucléaire présente des activités génomique et non-génomique, qui sont prédominantes dans des cellules exprimant fortement ou faiblement ERα, respectivement. L'augmentation de son activité non-génomique étant associée à l'échappement au contrôle hormonal, et donc la résistance à l'hormonothérapie et un mauvais pronostic lors de cancers du sein, il est nécessaire d'étudier les mécanismes menant à cette modification d'activité. Nous avons précédemment démontré que l'optimisation des codons d'ERα pour son expression dans un contexte plus prolifératif menait à une modification de ses propriétés fonctionnelles, suggérant un rôle de l'usage des codons dans l'expression et l'activité de certaines protéines. L'étude suivante vise alors à mieux comprendre les paramètres dans l'usage des codons qui contrôlent le taux d'expression des protéines, en prenant ERα comme exemple. En étudiant des données omiques issues de modèles cellulaires plus ou moins différenciés, et en transfectant de nouvelles versions synonymes d'ERα, nous avons identifié un rôle joué par la troisième base des codons dans la régulation de l'expression des gènes en protéines en fonction de l'état cellulaire. Plus précisément, un enrichissement en codons terminant par G ou C (GC3) est associé à une transcription plus importante mais aussi une meilleure efficacité de traduction en cellules quiescentes ou différenciées. A l'inverse, un messager enrichi en AU3 serait plus exprimé et mieux traduit lors des phases S à M ou en cellules dédifférenciées. Cette dynamique d'expression et traduction au cours du cycle en fonction de l'enrichissement en GC3 ne régulerait pas la production de tout le protéome, mais s'appliquerait à certaines

classes de protéines clefs dans la signalisation cellulaire, dont l'activité est finement régulée au cours du cycle, telles que des facteurs de transcription ou kinases. Il existe donc une coordination entre la transcription et la traduction pour la production optimale de certaines protéines au cours du cycle en fonction de leur composition en codons.

Contribution of the genetic code in the control of the expression level of proteins: the example of the estrogen receptor-alpha

Abstract

Genes expression into proteins is a fundamental process whose dysregulation is involved in most pathological conditions, and the complex regulation still under investigation. A role played by the codon usage bias notably emerged in recent years, but what drives this bias between cell states and how it regulates genes expression into proteins remains unknown. The estrogen receptor-alpha (ERα) is a nuclear receptor that plays a major role in breast cancer development when its expression or function are altered. ERα is normally expressed in differentiated, quiescent cells, and we previously demonstrated that codon optimization of this protein for its expression in more proliferating cells resulted in functional modifications, highlighting a role played by codon usage in protein expression and activity. This study aims then to better understand the parameters in codon usage that could control the expression level of proteins, using ERα as an example. The review of published omics data revealed that GC3-enrichment of mRNAs impacts their transcription and translation efficiency depending on cell state, with GC3-rich transcripts being better expressed in quiescent cells whereas AT3-rich sequences are favored during S to M phases or in dedifferentiated cells. In addition, transfections of new synonymous versions of ERα indicated a role played by the promoter for the coordination of transcription and translation depending on codon composition. This study shed then light on a role played by the third codon base in the regulation of gene to protein expression depending on cell state.

Keywords: Codon usage; translation regulation; estrogen receptor-alpha; breast cancer; endocrine resistance

1. Introduction

Gene expression is the fundamental process in cell functioning, with the alteration leading to pathological conditions. Understanding the complex regulation of transcription, mRNA stability, and translation is still in progress. It is known that almost every amino acid can be coded by several codons, named synonymous codons. It appears however that each synonymous codon is not used at the same frequency for a given amino acid, which is called the codon usage bias. More frequently used codons would be correlated to more abundant tRNAs, suggesting a notion of optimality in codon usage for the production of proteins (Novoa and Ribas de Pouplana, 2012). It was then showed that codon usage and tRNA abundance may differ depending on the cell type, cell cycle phase, or differentiation status of the cell (Gingold et al., 2014; Novoa and Ribas de Pouplana, 2012). For instance, Gingold et al. demonstrated in 2014 a strong dependence between tRNA choice and abundance and the transcriptional programs involved during cell proliferation and differentiation.

More recently, specific expression of tRNAs in metastatic breast cancer cells that may promote tumor progression has been observed (Goodarzi et al., 2016). In addition, codon usage would impact mRNA stability, with optimal codons associated with the stabilization of mRNAs (Presnyak et al., 2015; Wu et al., 2019). By impacting mRNA stability and translation, codon usage appears to play a functional role in the production of protein depending on cell state (Hanson and Coller, 2018). More than impacting the level of mRNAs and proteins, codon usage would also regulate folding efficiency and then the activity of some proteins. It was indeed reported that codon usage might regulate translation speed, which is a key parameter in the co-translational folding of proteins (Rodnina, 2016; Tsai et al., 2008; Yu et al., 2015). Nevertheless, precisely what drives codon usage bias and how it regulates genes expression into proteins depending on cell state is not fully understood.

The estrogen receptor-alpha (ERα) is the main mediator of estrogen actions in mammals, with a higher expression in reproductive than non-reproductive tissues (Arnal et al., 2017; Droog et al., 2016). Its expression level differs also between cell types in a given tissue. The mammary epithelium contains for example basal and luminal lineages, distinguished by the absence or the presence of ERα expression, respectively. The level of ERα expression in luminal cells varies however during mammary gland development, and cells expressing no, low, and high amounts of ERα coexist at a given time. What drives these differential expression remains however misunderstood (Rusidzé et al., 2021). By mediating estrogen actions, ERα is closely associated with pathological events such as breast cancers, where its functional dysregulation is a major contributor to tumor progression (Couse and Korach, 1999; Jordan and O'Malley, 2007). In case of ERα-positive breast cancers, ERα expression is in addition heterogeneous, with a decrease associated with a more aggressive phenotype. The underlying mechanisms of such expression regulation are not completely known, but it appears notably that ERα expression falls when cells enter cell cycle (Porras et al., 2021; Zheng and Murphy, 2016). ERα is a ligand-inducible transcription factor which regulates gene expression upon direct or indirect binding to regulatory sequences (Beato, 1989; Evans, 1988). Its physiological and pathological effects rely then first on nuclear, transcriptional actions, often observed in reproductive tissues (Arnal et al., 2017; Couse and Korach, 1999). Despite its obvious importance, it is however clear that this genomic activity does not account for all of the biological functions of ERα. In fact, under certain circumstances, ERα is also extra-nuclear, in particular associated with the plasma membrane, and activates from this site many signal transduction pathways, including MAP kinase (MAPK) and phosphatidyl-inositol 3-kinase (PI3K) signaling (Edwards, 2005). This non-genomic activity of ERα has been characterized in a variety of cell types, particularly in endothelial cells where ERα expression is low (Arnal et al., 2017). These modifications in ERα functions were associated with resistance to endocrine therapy and poor prognosis in human breast cancers, which raises interest in understanding the mechanisms behind this (Poulard et al., 2012). Alterations in ERα activity are often understood in relation to a change in the environment of the receptor, but changes in the expression level of the protein are also questioned.

To investigate if codon usage plays a role in ERα expression depending on cell state, we previously generated, based on Gingold et al.'s study (Gingold et al., 2014), an ERα synonymous sequence with all codons optimized according to the codon usage observed in genes specifically expressed in proliferating cells. The results obtained when characterizing this synonymous receptor revealed modifications of the functional properties of ERα, which highlighted an effective role played by codon usage in protein expression and activity (Clusan et al., in preparation). Then in the present paper, we further explore the possible role of codon usage in the control of the expression level of proteins depending on cell state. Published omics data in several model differing by their differentiation status were investigated, and new synonymous versions of ERα were studied by transient transfection. The results obtained shed light here on a role played by the third codon base in the regulation of gene to protein expression depending on cell state.

2. Materials and methods

2.1. Analyses of published whole-genome datasets

RNA sequencing (RNA-seq) and normalized global nuclear run-on sequencing (GRO-seq) data obtained in MCF7 cells synchronized at different cell cycle phase were retrieved from the GEO as series GSE94479 (Barrett et al., 2012; Liu et al., 2017). For RNA-seq analysis we used the available lists of differentially expressed genes together with their log2FC and P-values calculated in the original paper. Differential genes expression rates were calculated from the GRO-seq data using DEseq2 R v.1.34 package following the removal of non-expressed genes and the appliance of the HTS filter R package onto the gene counts matrix to exclude low-expressed RNAs from the analysis (Love et al., 2014; Rau et al., 2013). We considered specific variations in expression those exhibiting both an absolute value of fold-change >1.5 and a Benjamini-Hochberg corrected P-value calculated by a Wald test <0.05.

Datasets from the studies of Charafe-Jauffret et al. and Smid et al. were obtained from MSigDB collections (http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp) (Charafe-Jauffret et al., 2006; Smid et al., 2008). Datasets from the studies of Stumpf et al. (Stumpf et al., 2013), Tanenbaum et al. (Tanenbaum et al., 2015), Ly et al. (Ly et al., 2014), and Fernández-Calero et al. (Fernández-Calero et al., 2020), were downloaded from the supplementary material of the corresponding papers.

2.2. Plasmids and antibodies

The C3-Luc and CMV-βgal reporter genes and the pc-ERα WT and MRTFA-ΔN200 pcDNA4/TO expression vectors have been previously described (Huet et al., 2009; Jehanno et al., 2020). pcDNA6/TR and pcDNA4/TO plasmids (T-Rex system) were purchased from Invitrogen. pc-ERα SYN 100% GC3, 50% GC3, and 0% GC3 expression vectors have been directly synthesized (GeneArt, Life technologies) (Supplementary Figure 1). CMV promoter was replaced by cyclin A2 promoter (PCR product from -215 to +245 of the human CCNA2 gene (Henglein et al., 1994)) in the pCDNA 3.1 expression vectors, between BglII and NheI restriction sites, to generate ERα expression vectors driven by the cyclin A2 promoter. Used antibodies were: anti-ERα (HC20,

Santa Cruz), anti-ERα (6F11, Abcam), anti-GR (BuGR2, Abcam), anti-ERK1/2 (137F5, Cell Signaling Technology), anti-β actin (AC-15, Santa Cruz), anti-GFP (JL-8, Becton Dickinson) and dye-conjugated secondary antibodies (Alexa Fluor, Invitrogen).

2.3. Cell culture and treatments

MCF7, ZR-75-1, T47D, MDA-MB-231, SUM159PT and HEK-293 cell lines were routinely maintained in DMEM (Invitrogen) supplemented with 10% fetal calf serum (FCS) (Biowest) and antibiotics (Invitrogen) at 37°C in 5% CO_2. MCF7 T-Rex sub-clones (T-Rex system, Invitrogen) stably transfected with either empty pcDNA4/TO expression vector (control MCF7 cells) or with MRTFA-ΔN200 pcDNA4/TO expression vector (MRTFA-ΔN200 MCF7 cells) were previously described (Jehanno et al., 2020; Kerdivel et al., 2014). Before any experiments, cells were grown in phenol red-free DMEM (Invitrogen) containing 2.5% charcoal-stripped FCS (Biowest) for at least 72 hours. To induce expression of the constitutively active mutant of MRTFA (MRTFA-ΔN200), cells were treated for 48 hours with 1 µg/mL tetracycline before harvesting. Cells were stimulated with final concentrations of 10 nM for 17β-estradiol (E2, Sigma-Aldrich), 1 µM for 4-hydroxytamoxifen (OHT, Sigma-Aldrich), and 100 nM for ICI 182-780 (ICI, TOCRIS Bioscience). In order to analyze MCF7 cells during the cell cycle, cells were maintained in steroid-free medium during 72 hours, then treated with 10 nM E2 and synchronized at the G1/S phase transition with 1 µg/mL aphidicolin (Sigma-Aldrich) for 48 hours. Release of the aphidicolin block through washings allowed the cells to progress throughout their cell cycle. The efficiency of the synchronization step was confirmed by flow cytometry analysis as previously described (Kerdivel et al., 2014), with 70% of the cells stopped in the G1/S phase transition.

2.4. Transient transfections

Transfections were carried out using jetPEI reagent (Polyplus transfection) according to manufacturer's instructions. For luciferase assays, transfections were carried out in 24-well plates with 200 ng of reporter gene, 100 ng of CMV-βGal internal control and 50 ng of expression vectors per well. Luciferase and β-galactosidase assays were performed as previously described (Huet et al., 2009). Luciferase reporter gene activity was normalized to the β-galactosidase. For western blot and RT-qPCR experiments, transfections were carried out in 6-well plates with 500 ng of ERα expression vectors and 100 ng of GFP plasmid per well.

2.5. Immunofluorescence

Cells were plated on cover slides in 24-well plates. After treatments, cells were fixed with phosphate buffered saline (PBS) containing 4% paraformaldehyde and then permeabilized in PBS containing 0.3% Triton X-100. Incubation with primary antibodies (1:1000) was performed overnight at 4°C in PBS containing 3% FCS. The next day, cells were incubated with secondary antibodies (1:1000) in PBS-FBS for 2 h at room temperature before mounting cover slides in Duolink II mounting medium with DAPI (Sigma-Aldrich). The images were obtained with an ApoTome Axio Z1 Imager microscope (Zeiss) and processed with Axio Vision Software. Fluorescent cells were analyzed using ImageJ software.

2.6. Protein extraction and western blotting

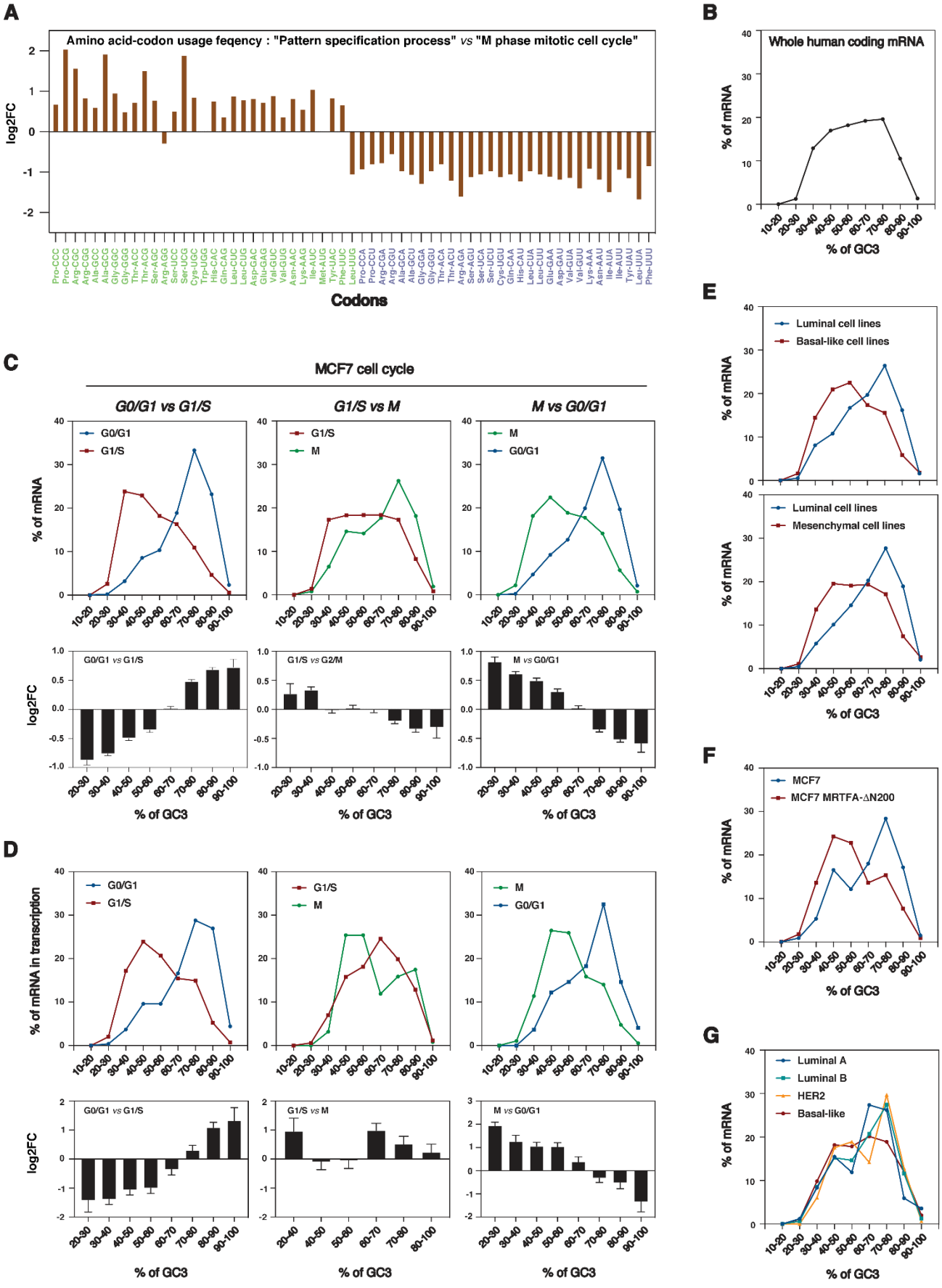
Whole-cell extracts were directly prepared in 3X Laemmli buffer. Following sonication, the proteins were denatured for 5 min at 95°C, separated on 10% SDS polyacrylamide gels, and transferred to polyvinylidene difluoride membrane (Millipore). The proteins were then probed with specific antibodies (1:1000) as previously described (Mérot et al., 2004), and detected using the Substrat HRP Immobilon Western kit from Millipore.

3. Results

3.1. Codon usage bias during the cell cycle and tumor progression is based on GC3 enrichment of mRNA coding sequences

Gingold et al. previously demonstrated using two major human GO categories, in particular the “M phase of mitotic cell cycle” and “pattern specification” gene sets, the existence of distinct codon usage between these two sets suggesting a dual program of translation regulation during cell proliferation and differentiation (Gingold et al., 2014). From the codon usage frequency they calculated which was based on a fraction per amino acid, we established a new codon usage frequency taking into account the abundance of amino acids deduced from human protein sequences (Supplementary Table 1) (Gardini et al., 2016; Rao et al., 2018). This method of calculation can be justified by the fact that the number of tRNA genes per amino acid is mostly correlated with amino acid abundance. Comparison of the codon usage frequency between “M phase of mitotic cell cycle” and “pattern specification” gene sets shows that the difference primarily relies on the last codon base, the “M phase of mitotic cell cycle” gene set being characterized by the presence of A or U at the third codon base (AU3) while the “pattern specification” gene set is enriched in G or C at this position (GC3). Only two GC3 codons do not follow this behavior, AGG and UUG codons (Figure 1A). To illustrate the phenomenon, a direct comparison of the codon usage frequency deduced from each cyclin (D1, E1, A1 and B1) known to be expressed at different phases of the cell cycle with the codon usage frequencies in “M phase of mitotic cell cycle” and “pattern specification” gene sets from Gingold et al.’s study was performed. Cyclin D1, expressed at the beginning of G1 phase, is encoded by codons whose frequency of use is similar to that of “pattern specification” gene set, with an enrichment of GC3. On the other hand, cyclins E1, A1 and B1, expressed during the S to G2/M phases, have codons whose frequency of use is similar to that of “M phase of mitotic cell cycle” gene set, with an enrichment of AU3 (Supplementary Figure 2).

We first investigated the consequence that the choice of the last base of the codon might have on the expression level of the transcripts, focusing mainly on breast cancers and their status towards ERα. For this purpose, we calculated the overall percentage of GC3 for each coding sequence of all protein-coding genes of the human genome and then reanalyzed transcriptomic data according to this percentage of GC3 (Supplementary Data). Transcripts were ranked according to their percentage of GC3, grouped in window every 10 units and then data were expressed as a percentage of the total transcripts examined.



(legend on the next page)

Figure 1: mRNA expression during MCF7 cell cycle and in breast cancer cell subtypes is correlated with codon enrichment in GC3. (**A**) Codon usage frequency differentially expressed in log2FC between gene sets of the human GO categories "pattern specification" and "M phase of the mitotic cell cycle". Codons ending with G or C (GC3) are in green and those ending with A or U (AU3) are in purple. (**B-G**) Transcripts were ranked according to their percentage in GC3, grouped in windows of 10 units from 10 to 100. The number of transcripts per window is then expressed as a percentage of the total transcripts examined. (**B**) Analysis carried out on the whole human coding mRNA (n=20,729). (**C**) Analysis performed on differentially expressed mRNAs between G0/G1 phase, G1/S phase transition and M phase of synchronized MCF7 cells. The log2FC of differentially expressed mRNAs per window is also shown (p-value<0.05; log2FC>0.585; n=4,905 for G0/G1 vs. G1/S; n=2,412 for G1/S vs. M; n=4,323 for M vs. G0/G1) (Liu et al., 2017). (**D**) Analysis performed on differentially transcribed mRNAs between G0/G1 phase, G1/S phase transition and M phase of synchronized MCF7 cells. The log2FC of differentially transcribed mRNAs per window is also shown (p-value<0.05; log2FC>1.5; n=673 for G0/G1 vs. G1/S; n=362 for G1/S vs. M; n=625 for M vs. G0/G1) (Liu et al., 2017). (**E**) Analysis performed on differentially expressed mRNAs between luminal, basal-likes and mesenchymal cell lines (n=815 for luminal vs. basal-like; n=888 for luminal vs. mesenchymal) (Charafe-Jauffret et al., 2006). (**F**) Analysis carried out on differentially expressed mRNAs between MCF7 and MCF7 MRTFA-ΔN200 (p-value<0.05; log2FC>0.585; n=676) (Fernández-Calero et al., 2020). (**G**) Analysis performed on up-regulated mRNAs constituting the genetic signature of the luminal A (n=84), luminal B (n=164), HER2 (n=148) and basal-like (n=650) breast cancer subtypes (Smid et al., 2008).

Of course, GC3-rich transcripts are AU3-poor and vice versa. The whole human coding mRNA shows a large majority of transcripts (86.7%) with a percentage of GC3 of their coding sequence between 50 and 80%. 1.3% of the transcripts were below 30% GC3 and nearly 12% above 80% (Figure 1B). In agreement with Gingold et al.'s study, ontology and pathway enrichment analysis reveals an overexpression of genes involved in cell cycle, mitosis, meiosis or DNA repair among genes with GC3 less than 45% and an enrichment in genes involved in tissue development, regulation of cell differentiation or cell-cell signaling among genes with GC3 greater than 65% (Supplementary Table 2) (Gingold et al., 2014).

To evaluate the impact of GC3% of the coding sequence on mRNA expression level during the cell cycle, we re-exploited RNA sequencing (RNA-seq) and global nuclear run-on sequencing (GRO-seq) data of the study of Liu et al. performed on the luminal (ERα positive) breast cancer cell line MCF7 (Liu et al., 2017). In this study, cells were synchronized to G0/G1 with hormone starvation, to G1/S with double thymidine treatment and to M phase with thymidine-nocodazole treatment. Genes were considered differentially expressed when P-value was <0.05 and the log2 FC >0.585. Among the differentially expressed genes between G0/G1 and G1/S phases, genes which are up-regulated in S phase are GC3 poor while genes enriched in G0/G1 show a high level of GC3 (Figure 1C). Differentially expressed genes between G1/S and M phases showed a relatively neutral profile regarding GC3 enrichment, with a minor decrease for G1/S phase genes and a slight increase

for M phase genes. Finally, comparison of differentially expressed genes between M and G0/G1 phases shows a clear return to GC3-enriched genes in quiescent cells. The same analysis was conducted for GC1 and GC2 enrichments and no differences were highlighted between cell cycle phases, revealing that the dynamics identified is specific to the third codon base (Supplementary Figure 3). Analysis of GRO-seq data shows identical dynamics, indicating that GC3 enrichment of gene coding sequences are closely associated with gene promoter activities (Figure 1D).

We then extended our analysis to the differentially expressed genes between the main breast cancer tissues or cell line subtypes: luminal, HER2-enriched and basal-like (Perou et al., 2000). We first processed the data from the study of Charafe-Jauffret et al. that clustered a total of 31 breast cell lines into luminal, basal-like and mesenchymal cell lines (Charafe-Jauffret et al., 2006). A set of 1,233 genes differentially expressed between luminal and basal-like cell lines was selected and 1,309 genes discriminating luminal and mesenchymal cell lines were analyzed. Classification of characterized genes according to their percentage of GC3 shows a clear GC3 enrichment in genes specifically expressed in luminal cell lines and on the contrary a depletion in GC3 in genes up-regulated in basal-like cell lines (Figure 1E). In comparison to luminal cell lines, the genetic signature of mesenchymal cell lines was more neutral and less oriented with respect to GC3 enrichment. We previously showed that activation and nuclear accumulation of the myocardin-related transcription factor A (MRTFA), a major actor in the epithelial to mesenchymal transition (EMT), mediates endocrine resistance of ERα-positive breast cancers by initiating a partial transition from luminal to basal-like phenotype and impairing ERα cistrome and transcriptome (Jehanno et al., 2020, 2021). We identified through RNA-seq differentially expressed genes between the native MCF7 cell line and a MCF7 cell line expressing a mutated form of MRTFA (MRTFA-ΔN200) devoid of its N-terminal actin binding sites, which allows a permanent translocation and a constitutive activity of MRTFA into the nucleus (Fernández-Calero et al., 2020). As expected, the analysis of GC3 enrichment of differentially expressed genes between the two cell lines shows a prevalence of GC3-rich genes in natives MCF7 and AU3-rich genes in MCF7 MRTFA-ΔN200 (Figure 1F). Finally, analysis of GC3 abundance in up-regulated genes constituting the genetic signature of different breast cancer subtypes deduced from tissues shows an enrichment of GC3-rich genes in luminal A, luminal B and HER2 breast cancers in comparison with basal-like breast cancers (Figure 1G) (Smid et al., 2008).

3.2. GC3 enrichment of mRNAs impacts translational efficiency and protein level during cell cycle

In light of the above results, we then investigated whether changes in GC3 enrichment of the coding sequence of transcripts could impact their translational efficiency (TE). For this end, we reprocessed, taking into account the GC3 enrichment of the coding region of the transcripts, data from the two studies that screened transcripts under translational control at different phases of the cell cycle in synchronous mammalian cell lines (Stumpf et al., 2013; Tanenbaum et al., 2015). In Stumpf et al.'s study, HeLa cells were synchronized in G1/S and S phases by release from thymidine block and in M phase by nocodazole treatment (Stumpf et al., 2013). A CDK1 inhibitor (RO-3306), on the other hand, was used in Tanenbaum et al.'s study

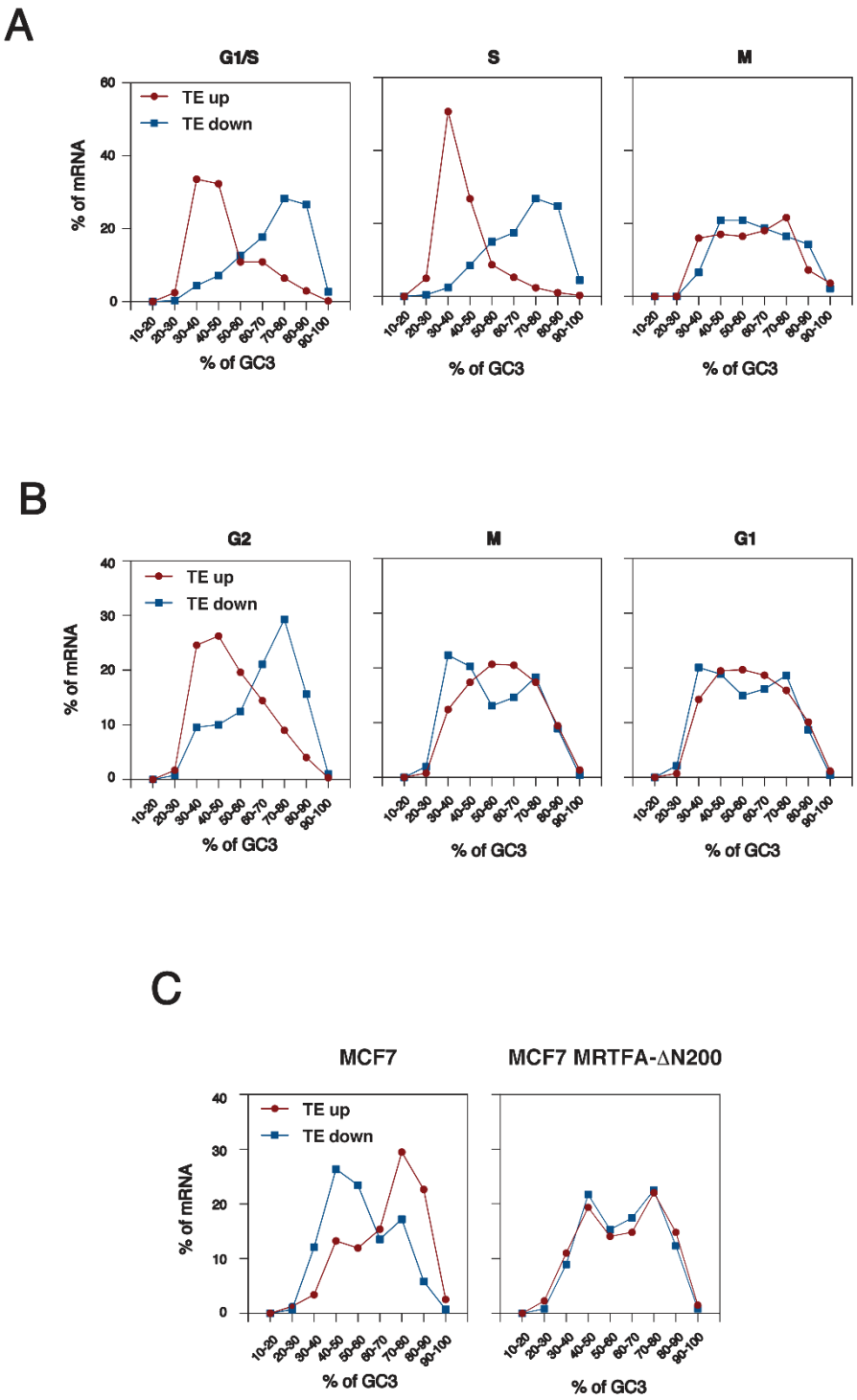
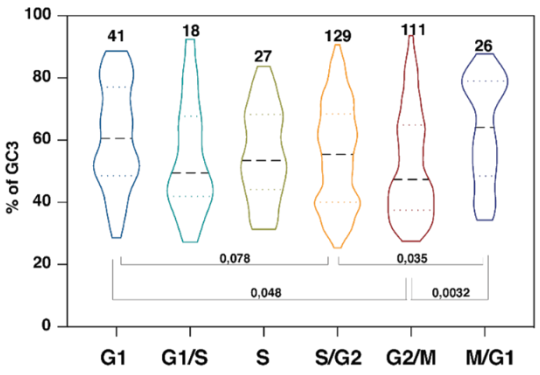


Figure 2: mRNA transitional efficiency during HeLa and RPE-1 cell cycle and in breast cancer cells MCF7 and MCF7 MRTFA-ΔN200 is correlated with codon enrichment in GC3. (**A-C**) mRNA translational efficiency was collected from several studies. Transcripts were ranked according to their percentage in GC3, grouped in windows of 10 units from 10 to 100. The number of transcripts per window is then expressed as a percentage of the total transcripts examined. (**A**) Differentially translated mRNA in G1/S transition (n=695), S phase (n=627) and M phase (n=285) of synchronized HeLa cells (Stumpf et al., 2013). (**B**) Differentially translated mRNA in G2 phase (n=3,994), M phase (n=3,994) and G1 phase (n=3,365) of synchronized RPE-1 cells (log2FC>0.5) (Tanenbaum et al., 2015). (**C**) Differentially translated mRNA in MCF7 (n=288) and MCF7 MRTFA-ΔN200 (n=498) (p-value<0.05; log2FC>1.5) (Fernández-Calero et al., 2020).

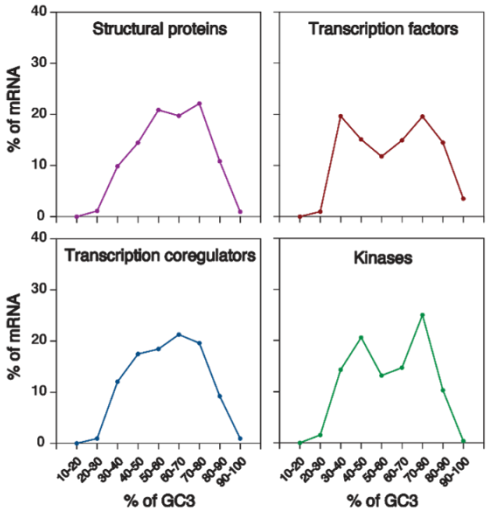
to synchronize RPE-1 cells from G2 phase to M and G1 phases upon removal of the inhibition (Tanenbaum et al., 2015). Results clearly show that the differentially translated mRNAs in cell cycle phase identified in both studies can be partly correlated to GC3 enrichment, especially in G1/S phase transition and in S and G2 phases (Figure 2A and 2B). In these phases, transcripts with reduced translational efficiency appear GC3-rich while those with increased translational efficiency are enriched in AU3. In the others analyzed phases (M and G1), differentially translated transcripts cannot be discriminated though their percentage in GC3, the GC3 enrichment remaining relatively neutral and undirected. Since these cells express little or no E-cadherin, a main epithelial marker involved in cell-cell contact, they might have difficulty to rest in quiescence and enter G0 phase. Therefore, we completed the study by processing the differentially translated transcripts in regard to their GC3 enrichment, in starved and rested MCF7 in G0/G1 phase as well as in MCF7 cells expressing MRTFA-ΔN200 (Supplementary Data). Interestingly, translationally up-regulated and down-regulated mRNA in starved MCF7 are GC3- and AU3-rich, respectively (Figure 2C). In contrast, no difference and a relatively neutral pattern in GC3 abundancy was observed between up- and down-regulated transcripts in MCF7 MRTFA-ΔN200. Together, these data suggest that GC3 enrichment of the coding sequence of transcripts regulates their translational efficiency during the different phases of cell cycle, likely adapting protein production.

To test whether GC3-mediated regulation has an impact at the protein level, we exploited data from Ly et al.'s study, in which a proteomic analysis of the cell cycle progression of the human myeloid leukemia NB4 cell line was performed through flow cytometric elutriation of cells in suspension (Ly et al., 2014). 358 proteins whose abundance was cell cycle regulated were identified in the study. GC3 enrichment of their coding sequence shows indeed a significant GC3 depletion for proteins specifically expressed in S/G2 and G2/M sets compared to G1 and M/G1 sets (Figure 3A). The expression level of a given protein can thus be partly controlled by the abundance of GC3 codons in its coding sequence to be optimal either in quiescent cells in G0/G1 phase or in dividing cells. To identify the protein families that mostly exploited this bistable expression system controlled by GC3 enrichment, we looked at the distribution of proteins according to GC3 abundance in four major protein families: structural proteins (GO:0005198), transcription coregulators [GO:0003713 (coactivators); GO:0003714 (corepressors)], transcription factors (Lambert et al., 2018), and kinases (List of Human Kinases – KinHub). While structural proteins and transcription coregulators show a distribution of GC3 enrichment in their coding sequence closed to that observed for the entire human coding genome, codon usage in transcriptions factors and kinases appears to be more tailored and specialized to both bistable expression systems with GC3-rich and AU3-rich populations (Figure 3B). This is obviously related to the fact that transcriptions factors and kinases are key modulators of signaling pathways whose activities are tightly regulated during cell cycle. It is interesting to note that many kinases which exist as multiple isoforms with high homology in their amino acid sequences show opposite GC3 enrichment in their codons between the different isoforms (Figure 3C).

A



B



C

xName	HGNC Name	Family	% GC3
Erk1	MAPK3	MAPK	81,8
Erk2	MAPK1	MAPK	57,3
Erk3	MAPK6	MAPK	36,5
Erk4	MAPK4	MAPK	84,1
AKT1	AKT1	Akt	84,8
AKT2	AKT2	Akt	80,2
AKT3	AKT3	Akt	36,6
GSK3A	GSK3A	GSK	71,6
GSK3B	GSK3B	GSK	42,8
MSK1	RPS6KA5	RSK	42,9
MSK2	RPS6KA4	RSK	85,8
RSK1	RPS6KA1	RSK	73,5
RSK2	RPS6KA3	RSK	31,9
RSK3	RPS6KA2	RSK	72,3

Figure 3: Protein expression is correlated with codon enrichment in GC3. (**A**) Violin plots representing the distribution of the differentially expressed proteins in different phases of the cell cycle of the human myeloid leukemia NB4 cell line according to the GC3 enrichment of their coding sequence. Protein number is indicated on the top of each plot. P-values are specified (Ly et al., 2014). (**B**) Protein distribution according to GC3 enrichment of their coding sequence in four protein families: structural proteins (GO:0005198), transcription coregulators [GO:0003713 (coactivators); GO:0003714 (corepressors)], transcription factors (Lambert et al., 2018), and kinases (List of Human Kinases – KinHub). Coding sequences were ranked according to their percentage in GC3, grouped in windows of 10 units from 10 to 100. The number of proteins per window is expressed as a percentage of the total proteins examined. (**C**) Table listing some kinases existing under several isoforms with the percentage of GC3 of their coding sequence.

To exemplify GC3-mediated regulation, we then probed by western blot the expression level of ERα and GR (glucocorticoid receptor) transcription factors and ERK1 and ERK2 kinase isoforms due to their opposite GC3 enrichment (Figure 4A), in breast cancer cell lines with differential proliferation and EMT status and during the different cell cycle phases of the luminal cell line MCF7. The protocol of cell cycle synchronization of MCF7 cells is depicted on Supplementary Figure 4, with its validation by flow cytometry. ERα coding sequence is 72% GC3-rich and its codon usage frequency correlates strongly with that of "pattern specification" gene set (R^2 = 0.88) and weakly with that of "M phase of mitotic cell cycle" gene sets (R^2 = 0.24) from Gingold et al.'s study (Figure 4B, first line). As expected, ERα is expressed in the luminal breast cancers cell lines MCF7, ZR-75-1 and T47D and is down-regulated in highly proliferative basal-like breast cancer cell lines MDA-MB-231 and SUM 159PT as well as in MCF7 expressing the EMT-induced form MRTFA-ΔN200. During the cell cycle of MCF7, its expression is also strongly decreased (Figure 4C, first line). Unlike ERα, GR has a coding sequence poor in GC3 (42%) which is mainly correlated to the codon usage frequency of "M phase of mitotic cell cycle" gene sets (R^2 = 0.75) from Gingold et al.'s study (Figure 4B, second line). Its expression remains similar whatever the analyzed cell lines and increases in MCF7 cells during the cell cycle (Figure 4C, second line). A similar dynamic is observed between both kinases (Figure 4B and 4C, third lines): ERK1 whose coding sequence is 81.8% GC3-rich and adapted to quiescent cells shows progressively reduced expression from luminal to basal-like cell lines as well as during MCF7 cell cycle. With a less specialized coding sequence in term of codon usage frequency, ERK2, on the other hand, shows almost no variation. It is interesting to note that β actin, used as a control in the study, shows almost unchanged expression profile whatever analyzed samples, while its coding sequence is 84.3% GC3-rich (Figure 4). This indicates that codon usage can partially but not systematically control the expression level of proteins.

3.3. GC3 enrichment impacts ERα protein level and activity in a promoter-specific manner

In order to ascertain definitively the key role of GC3 codons in the protein expression level, we generated new ERα cDNAs in which all codons were mutated into synonymous codons in order to obtain 100%, 50% and 0% GC3 codons. These mutants, called ERα SYN 100, 50 or 0% GC3, have an amino acid sequence identical to that of wild-type (WT) ERα protein, but a different cDNA sequence. For the SYN 50% GC3 mutant, an alternation of codons ending by GC3 or AU3 was realized by respecting the order of the purine and pyrimidine bases. Methionine and tryptophan codons were of course conserved in ERα SYN 0% GC3. The expression of these SYN mutants was compared to that of the WT in HEK-293 cell line in transient transfection experiments. Since GC3 enrichment of gene coding sequences was shown to be closely associated with gene promoter activities (Figure 1D), we selected two different promoters to control the expression of proteins of interest: the classical CMV promoter and the cyclin A2 promoter, whose activity is maximal in S phase (Henglein et al., 1994). Analyzed by both immunofluorescence and western blot experiments, the expression level of ERα proteins was obviously regulated in a promoter- and codon-specific manner (Figure 5A-C). Under the control of the CMV promoter, high expression level of ERα WT and SYN 100% GC3 was observed while ERα SYN 50 and 0% GC3 were very weakly expressed. It should be noted that in the few cells expressing ERα SYN 50 and

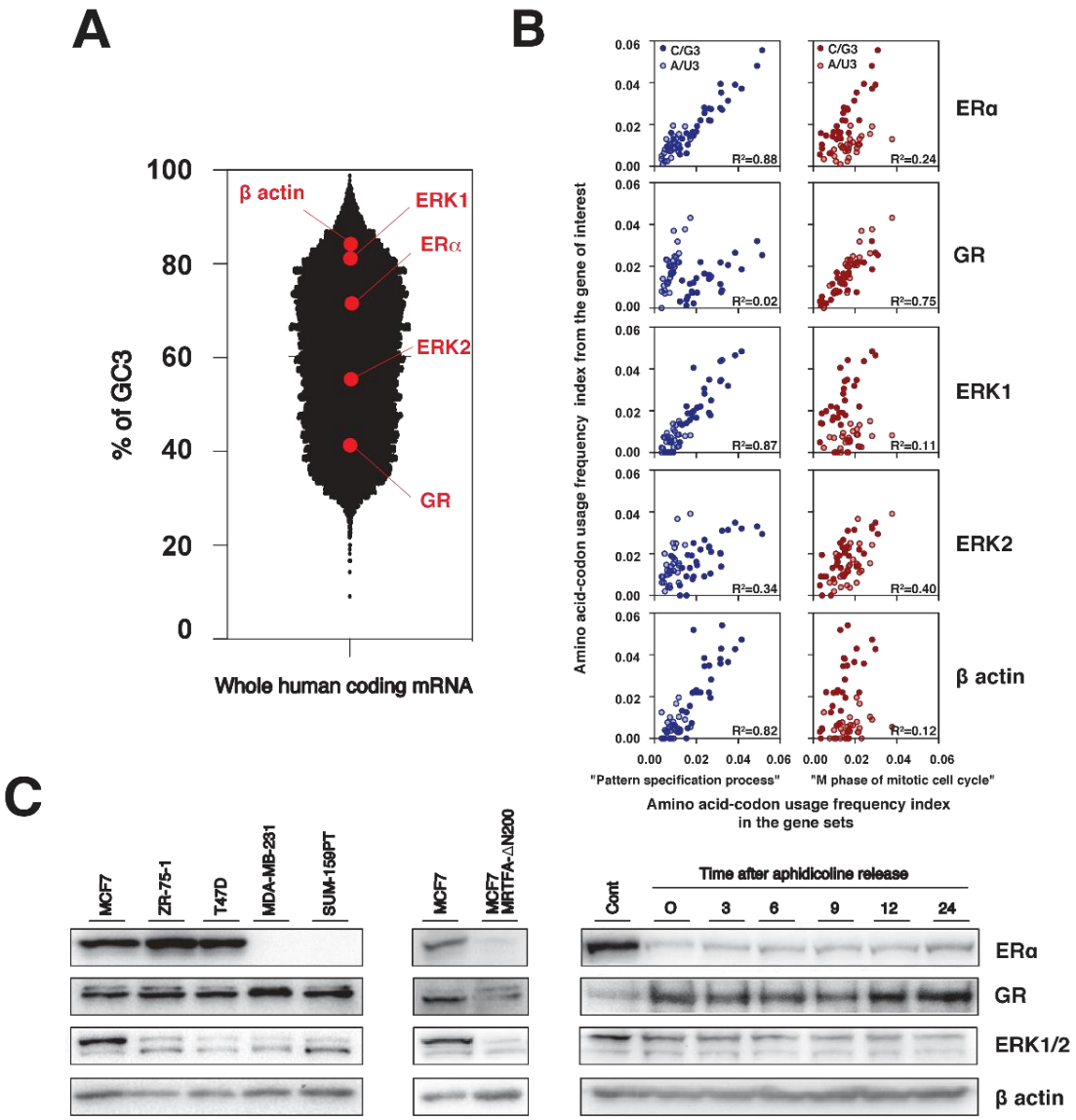


Figure 4: ERα, GR, ERK1, ERK2 and β actin expression in different breast cancer cell lines and during the MCF7 cell cycle. (**A**) Representation of the percentage in GC3 of the coding sequence of ERα, GR, ERK1, ERK2 and β actin in the whole human coding genome. (**B**) Comparison of the codon usage frequency deduced from ERα, GR, ERK1, ERK2 and β actin coding sequences with the codon usage frequency in "M phase of mitotic cell cycle" and "pattern specification" gene sets from Gingold et al.'s study (Gingold et al., 2014). GC3 and AU3 codons are represented by different symbols. The correlation coefficient is indicated. (**C**) Western blot analysis of the expression of the proteins of interest in whole cell extracts prepared from asynchronous MCF7, ZR-75-1, T47D, MDA-MB-231, SUM-159PT and MCF7 MRTFA-ΔN200 cell lines as well as from synchronized MCF7 cells following aphidicolin block.

0% GC3 forms, the proteins were correctly produced as evidenced by antibodies targeting both ERα N- and C-terminal domains (Supplementary Figure 5). In comparison to the expressions under CMV control, cyclin A2 promoter in contrast significantly enhances the expression of ERα SYN 50 and 0% GC3 and reduces those of WT and SYN 100% GC3. Used as internal control, the expression of a GFP protein under the control of a CMV promoter remained unchanged. While the difference in strength between both promoters might explain in part the decrease of ERα WT and SYN 100% GC3 expression between CMV and cyclin A2 promoters, the increased expression of ERα SYN 50 and 0% GC3 under the control of the cyclin A2 promoter provides strong support for the link between GC3 enrichment and promoter activity during cell cycle.

We then assessed the functional properties of the produced ERα proteins. As ERα is primarily a ligand-inducible transcription factor, the transactivation efficiency of ERα WT and SYN forms were measured on a ERE-driven reporter gene in the presence or absence of different ligands, including estradiol (E2, 10 nM), the Selective Estrogen Modulator (SERM) 4-hydroxytamoxifen (OHT, 1 µM) and the Selective Receptor Degrader (SERD) ICI 182-780 (ICI, 100 nM). As illustrated in Figure 5E, ERα WT and SYN 100% GC3 show clear transcriptional activity in presence of E2, which is stronger when ERα are produced through CMV than cyclin A2 promoter, obviously because of the differences in protein expression. In contrast, no transcriptional activation was detected for ERα SYN 50 and 0% GC3 forms, no matter which promoter was used to produce them. The lack of reporter gene induction could be related either to a too low level of ERα AU3-rich SYN forms even under the control of cyclin A2 promoter, or to the fact that these ERα proteins are non-functional as transcription factors. A detailed analysis of the induced reporter gene activity in the presence of decreasing amounts of ERα WT protein shows that for a similar protein amount to that observed with ERα SYN 0% GC3 form, ERα WT is still able to induce a reporter gene activity (Figure 5D and 5E, right panel). Therefore, ERα SYN 0% GC3 form and probably ERα SYN 50% GC3 are non-functional regarding to their transcriptional activity. The combination of GC3 enrichment and promoter activity impacts then protein production and activity.

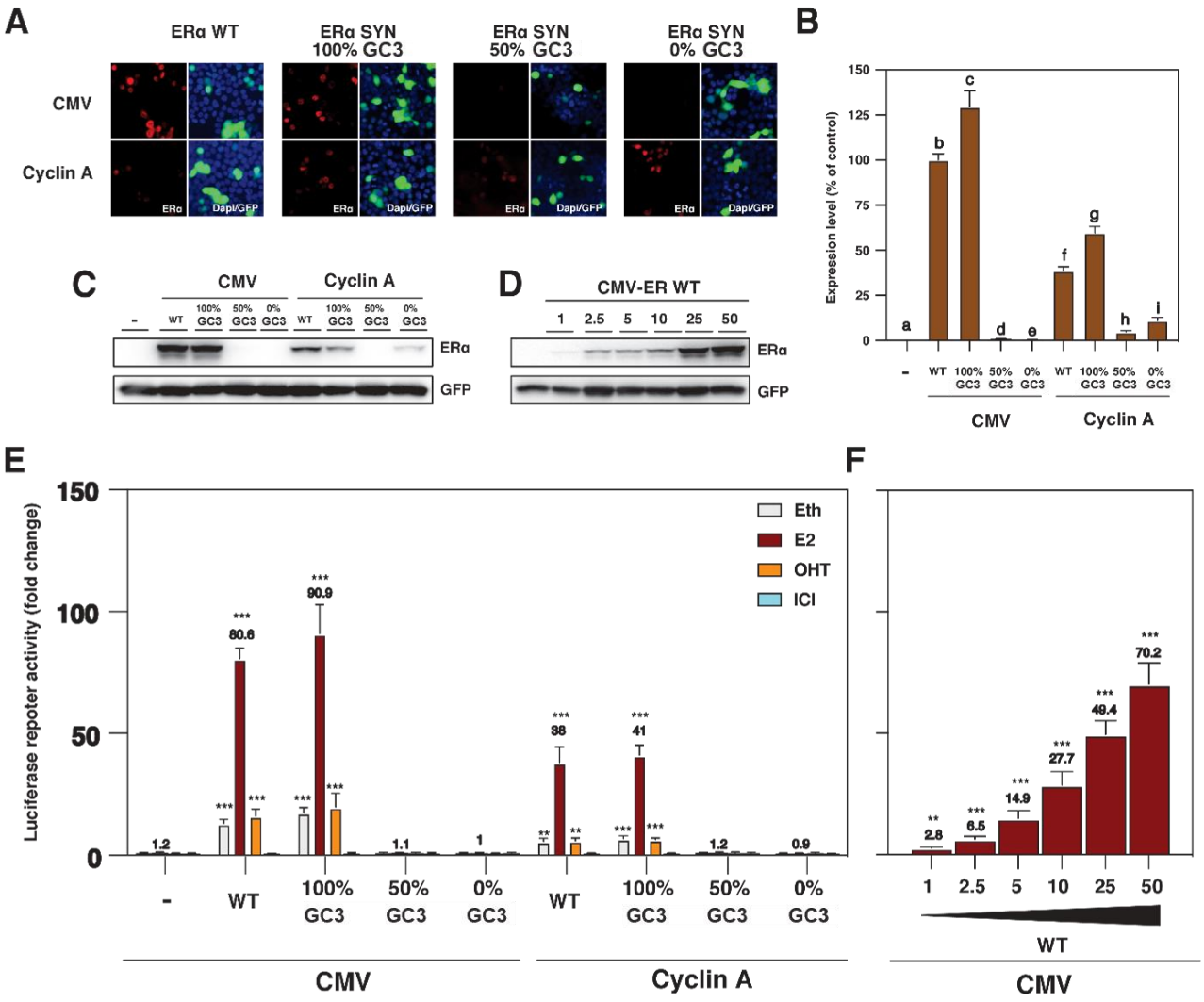


Figure 5: Change in GC3 enrichment of ERα coding sequence impacts the expression level and activity of the protein in a promoter-specific manner. (**A-C**) HEK-293 cells were transfected with the expression vectors pcDNA ERα WT, SYN 100, 50 or 0% GC3 under the control of CMV or cyclin A2 promoters together with the internal control CMV-GFP plasmid. (**A**) Immunofluorescent detection of ERα WT and SYN proteins by a C-terminal ERα antibody associated with GFP florescence. Representative images are shown with exposure times different between each ERα form but identical between each promoter of a same ERα form. (**B**) Densitometry quantification of the immunofluorescence images, with identical exposure times, expressed as a percentage of the intensity measured for ERα WT form. Results are the means ± SEM. Columns with different superscripts differ significantly (n=30-40; p-value <0.05). (**C**) Western blots analyzing the expression level of ERα WT and SYN forms as well as GFP. (**D**) HEK-293 cells were transfected with increasing amounts (0 to 50 ng) of ERα WT expression vector compensated by an empty plasmid together with 50 ng of CMV-GFP plasmid. Western blots were performed to measure ERα and GFP expressions. (**E**) HEK-293 cells were transfected with C3-Luc and CMV-βgal reporter genes together with pcDNA ERα WT, SYN 100, 50 or 0% GC3 under the control of CMV or cyclin A2 promoters. Cells were then treated for 24 h with E2 (10 nM), OHT (1 µM), ICI (100 nM) or vehicle (Eth). Data correspond to the mean values ± SEM and are expressed as fold change from vehicle-treated control (n=9-15; **p-value <0.01 and ***p-value <0.001, Student's t-test). E2-

induced fold change of the reporter gene is indicated for each ERα form. (**F**) HEK-293 cells were transfected with C3-Luc and CMV-βgal reporter genes together with increasing amounts (0 to 50 ng) of ERα WT expression vector compensated with an empty plasmid. Cells were then treated for 24 h with E2 (10 nM). Data correspond to the mean values ± SEM and are expressed as fold change from empty expression vector (0 ng) (n=9; **p-value <0.01 and ***p-value <0.001, Student's t-test).

4. Discussion

Translation of mRNAs into proteins is a key event in the regulation of gene expression. Gingold et al. demonstrated the existence of two distinct translation programs that operate during proliferation and differentiation, by examining codon usage in two major GO categories: "M phase of mitotic cell cycle" and "pattern specification" (Gingold et al., 2014). These two programs seem to differ from each other partly on the enrichment in GC of the last base of the codons. Our detailed analysis of differentially expressed genes during MCF7 cell cycle shows a progressive switch from GC3-rich genes in G0/G1 to AT3-rich genes in S phase and then a gradual return of GC3-rich genes after mitosis. A similar dynamics was observed when comparing the different subtypes of breast cancers. Notably, differentially expressed genes between the differentiated luminal and poorly differentiated basal-like breast cancers showed clear enrichment of GC3 in genes specifically expressed in luminal cancers, disappearing in dedifferentiated cancers in favor of AT3 codon bias. This difference is obviously related to distinct mitotic index between these subtypes of breast cancers. In the course of our work, several studies have come to reinforce these observations. Translation program of proliferation-related mRNAs was shown to be biased by rare codons that were poorly adapted to tRNA pools but whose translation was boosted in rapid dividing cells compared to mRNAs encoded with common codons (Guimaraes et al., 2020). Using ribosome and transcriptome profiling, Bornelov et al. further revealed distinct codon signatures during human embryonic stem cell differentiation, including a strong increase in GC content during differentiation independent of the nature of its stimulus (Bornelöv et al., 2019). Finally, from a systematic meta-analysis of transcriptomes representing 40 pathologies, change in the use of synonymous codon was identified for each disease analyzed, with notably an increase in the use of A- or T-ending codons in all cancers compared to healthy control tissues (Fornasiero and Rizzoli, 2019). Our study reveals that GC3 bias in the differential abundance of mRNAs during the cell cycle is closely associated with gene transcriptional activity, with genes active in G0/G1 being preferentially GC3-rich and those induced in S phase preferentially AT3-rich. This transcriptional regulation could also be complemented by changes in mRNA stability. Indeed, many studies have demonstrated a main role of codons in the control of mRNA decay in yeast and mammalian cells (Hia et al., 2019; Kudla et al., 2006; Presnyak et al., 2015). Stable mRNAs are generally enriched in optimal codons whereas unstable mRNAs contain predominately non-optimal codons. Particularly, stabilized mRNAs were GC3-rich while destabilized mRNAs were composed by AU3 codons in human cells (Courel et al., 2019; Hia et al., 2019). Corroborating this observation, several proteins were identified as being involved in the recognition of GC3-rich or AU3-rich mRNAs and control their degradation

(Courel et al., 2019; Hia et al., 2019; Radhakrishnan et al., 2016; Shu et al., 2020). Although this is highly likely, further experiments should obviously be conducted to determine whether codon-mediated changes in mRNA stability also occur in the differentially expressed mRNAs during cell cycle and pathological processes. As last step in protein synthesis, the translational level shows a strong regulation during the cell cycle with a decrease in the translational efficiency of GC3- and AU3-rich mRNAs in S/G2 and in G0/G1 phases, respectively. All these data suggest a tight control of protein production through GC3 codon bias in addition to gene transcriptional regulation. Partly supported by proteomic data analysis, the hypothesis was definitively demonstrated by monitoring the expression of re-coding ERα cDNAs with synonymous codons to obtain different GC3 codon enrichments, from 0 to 100%. As reported in previous studies, optimal codons increased protein production (Guimaraes et al., 2020; Kudla et al., 2006). More importantly, we show that the association between GC3 enrichment of mRNA and promoter activity is critical, with the efficient translation of AU3-rich mRNAs requiring promoters with high mid-cell cycle activity. However, when ERα SYN 0% GC3 was produced under the control of cyclin A2 promoter, this synonymous version of ERα lost its transcriptional activity. Even if the recognition of this protein by antibodies targeting both N- and C-terminal domains suggests that its conformation is not dramatically modified, the results obtained regarding its transcriptional activity suggest that the protein structure differs from the wild-type. More experiments would be needed to investigate other properties of this synonymous receptor and confirm this hypothesis, but these results highlight a role played by the third codon base in protein activity, probably by altering the co-translational folding of the protein. Previous work in fact suggested that ERα conformation would be sensitive to the cellular environment and codon composition (Fernández-Calero et al., 2014; Horjales et al., 2007). These results enable then to go further in the previously suggested hypothesis regarding the role of codon usage in protein production and activity, by shedding light here on the role of the third codon base in this process (Clusan et al., in preparation). It is interesting to note that GC3-poor ERα cDNAs were faintly translated under the control of CMV promoter, possibly suggesting cell cycle- or cell state-sensitive activity of this promoter. In comparison to other promoters, CMV was shown indeed to exhibit low activity in embryonic stem cells which then increases following differentiation (Chung et al., 2002; Wang et al., 2008). Despite a better characterization thanks to all these data, the precise nature of GC3-mediated translation regulation remains to be defined. Variations in individual tRNA abundance have been proposed (Gingold et al., 2014; Goodarzi et al., 2016), but other studies show instead the absence of major changes in the individual expression of tRNAs in a cell cycle-, cell state- or cell-type specific manner (Guimaraes et al., 2020; Pouyet et al., 2017; Rudolph et al., 2016). Our own preliminary results show limited variation in the tRNA expression profile between native luminal MCF7 cells and transformed MRTFA-ΔN200 expressing MCF7 cells (Marin M., unpublished data). GC3-mediated translation regulation could also be based on mRNA and tRNA modification mechanisms. Especially, acetylation of cytidine in mRNA was shown to promote translational efficiency (Arango et al., 2018), and self-renewing embryonic stem cells optimize translation of codons that depend on inosine-modified tRNAs in the anticodon wobble position (Bornelöv et al., 2019). Finally, GC3-mediated translation regulation might depend on the use of specialized ribosomes following rRNA editing (e.g., rRNA

2’-O-methylation and rRNA pseudouridylation) (Erales et al., 2017; King et al., 2003), post-translational modifications of ribosomal proteins (Bohlen et al., 2021), or changes in their stoichiometry (Shi et al., 2017; Slavov et al., 2015). Of particular interest, phosphorylation of serine 38 in RPL12, a known mitotic CDK1 substrate, which occurs mainly in monosomes, was shown to regulate the translation of specific subsets of mRNAs during mitosis (Imami et al., 2018). In conclusion, the bias in mRNA enrichment in GC3 codons appears to be at the center of different mechanisms of gene expression regulation, linking in an intricate, complex and state-specific manner, gene transcriptional activity, mRNA stability and translation, and protein properties.

Supporting information

ERα SYN 100% GC3

atgaccatgaccctccacaccaaggcgtccgggatggccctgctgcaccagatccaggggaacgagctggagccccctgaaccgcccgcagctcaagatccccctggagcggcccctg
ggcgaggtgtacctggacagcagcaagcccgccgtgtacaactacccccgagggcgccgcctacgagttcaacgccgcggccgccgccaacgcgcaggtctacggccagaccggcct
cccctacggccccgggtccgaggccgcggcgttcggctccaacggcctgggggggcttcccccccgctcaacagcgtgtccccgagcccgctgatgctgctgcacccgccgccgcagctg
tcgcccttcctgcagccccacggccagcaggtcccctactacctggagaacgagcccagcggctacacggtgcgcgaggccggcccgccggcgttctacaggccgaactcggacaac
cggcgccagggcggcagggagaggttggccagcaccaacgacaagggggagcatggccatggagtccgccaaggagacccgctactgcgcggtgtgcaacgactacgcctcgggct
accactacggggtctggtcctgcgagggctgcaaggccttcttcaagaggagcatccaggggcacaacgactacatgtgcccggccaccaaccagtgcaccatcgacaagaacaggag
gaagagctgccaggcctgccggctccgcaagtgctacgaggtggggatgatgaagggcgggatccggaaggaccggagggggggggaggatgttgaagcacaagcgccagaggga
cgacggggagggcaggggcgaggtggggtccgccggggacatgagggccgccaacctctggccgagcccgctcatgatcaagcgctccaagaagaagaacagcctggccttgtccctga
cggccgaccagatggtcagcgccttgttggacgccgagccccccatcctctactccgagtacgaccccaccaggcccttcagcgaggcctcgatgatgggcttgctgaccaacctggcgg
acagggagctggtccacatgatcaactgggcgaagagggtgccgggcttcgtggacttgaccctccacgaccaggtccacctcctggagtgcgcctggctggagatcctgatgatcggcc
tcgtctggcgctccatggagcacccggggaagctgctgttcgcccccaacttgctcttggacaggaaccaggggaagtgcgtggagggcatggtggagatcttcgacatgctgctggcca
cgtcgtcccggttccgcatgatgaacctgcaggggaggagttcgtgtgcctcaagtccatcatcttgctcaactccggggtgtacacgttcctgtccagcaccctgaagtccctggaggag
aaggaccacatccaccgggtcctggacaagatcacggacaccttgatccacctgatggccaaggcgggcctgaccctgcagcagcagcaccagcggctggcccagctcctcctcatcct
ctcccacatcaggcacatgagcaacaagggcatggagcacctgtacagcatgaagtgcaagaacgtggtgcccctctacgacctgctgctggagatgctggacgcccaccgcctgcacg
cgcccaccagccgcgggggggcgtccgtggaggagacggaccagagccacttggccaccgcgggctccacctcgtcgcactccttgcagaagtactacatcacgggggaggcggag
ggcttccccgccacggtctga

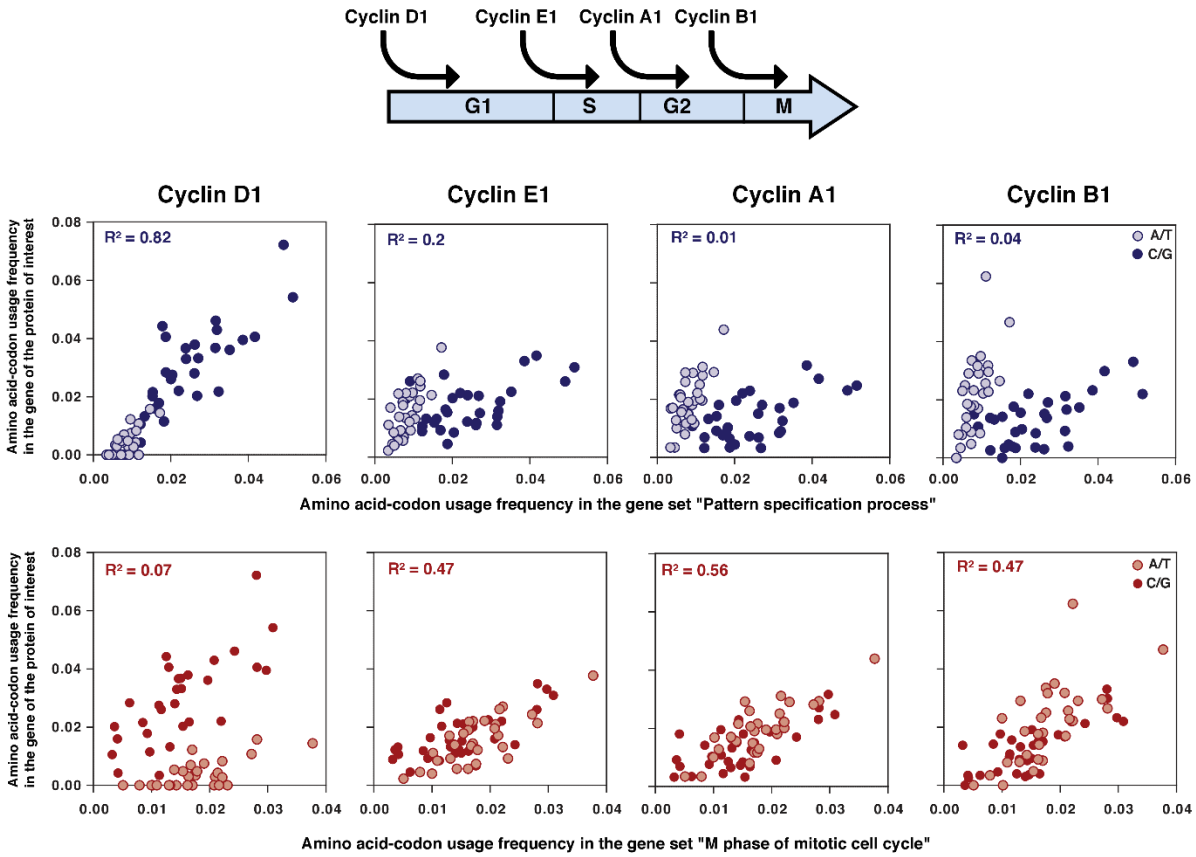
ERα SYN 50% GC3

atgactatgactctccataccaaagcgtctgggatggctctgctacaccaaatccaagggaatgagctagagcctctgaatcgcccacagcttaagattcccctagagcgacccctaggcga
agtgtatctggatagcagtaagcctgccgtatacaattaccctgagggtgccgcttacgaattcaatgccgcagccgctgccaatgcgcaagtctatggccaaaccggtctcccttacggtcc
cggatccgaagccgcagcgtttggctctaacggtctgggaggctttcccccactcaatagcgtatccccaagcccactgatgctactgcatccgccaccgcaactgtcacccttttctgcaacc
ccatggccaacaggtaccctattacctagagaatgagcctagcggttacacagtgcgtgaggctggcccaccggcattctataggccaaactcagacaatcggcgtcagggtggcagaga
gagattggctagcactaacgataagggaagcatggctatggaatccgctaaggaaacccgttactgtgcggtatgcaatgactatgcctcaggctatcactatgggtttggtcttgcgaagg
ctgtaaggctttctttaagagaagcattcagggacacaatgactatatgtgtccggctaccaatcagtgtaccattgacaaaaacagaaggaaaagctgtcaggcttgccgactccgtaagtgt
tacgaagtgggaatgatgaaaggcggaatccgaaaggatcggagaggggaaggatgttaaagcataagcgtcagagagacgatggggaaggcagaggcgaagtgggatccgctgg
ggatatgagagccgctaacctttggccaagcccactcatgattaagcgttccaaaaagaatagcctagccttatccctaacggctgaccaaatggttagcgctttgttagacgctgagcctccc
atactctattccgaatacgatcccactaggcctttcagtgaggcttcgatgatgggtttgctaaccaatctggcagacagagagctagtccatatgattaactgggcgaaaagggtaccgggttt
cgtagacttaacccttcacgatcaggttcaccttctggaatgcgcttggctagagattctgatgattggccttgtctggcgctctatggaacacccagggaaactgctattcgctcccaatttgctt
ttggataggaatcagggaaagtgtgtggaaggcatggtagagattttcgatatgctactggctacgtcatcccgattccgtatgatgaatctgcaagggggaagagtttgtgtgtctcaaatccatt
atcttactcaattccggagtgtatacgtttctgtctagcactctgaaatccctagaggaaaaggatcacattcaccgagtcctagacaaaatcacagacactttgattcacctaatggctaaggca
ggcctaaccctacagcaacagcatcagcgactggctcagcttctccttatcctttcccatatcagacacatgagtaacaaaggcatggaacacctatacagtatgaaatgcaaaaacgtagtg
cctctctatgacctactgctagagatgctagacgctcaccgtctgcatgcgcctaccagtcgcggaggggcatccgtagaggaaacggatcagagtcacttagccactgcgggttccacttc
gtcacactctttgcaaaagtattacattacgggagaggcagagggtttccctgccacagtctga

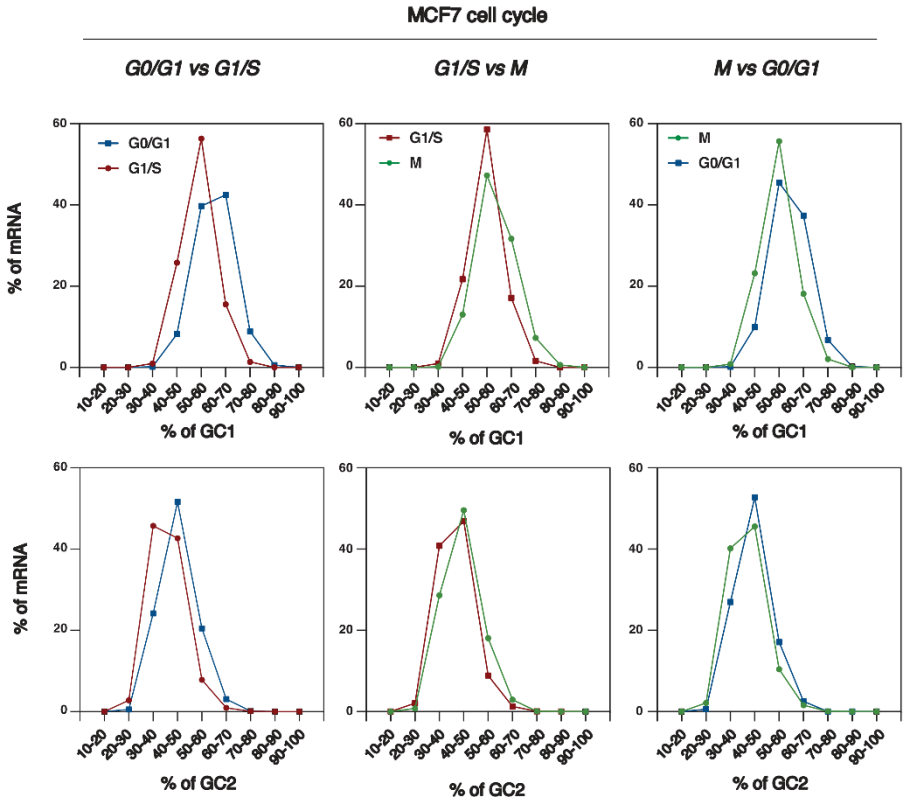
ERα SYN 0% GC3

atgactatgactcttcatactaaagcatctggaatggctctactacatcaaattcaaggaaatgaactagaacctctaaatcgtccacaacttaaaattcctctagaacgacctctaggtgaagtat
atctagatagtagtaaacctgctgtatataattatcctgaaggtgctgcttatgaatttaatgctgcagctgctgctaatgcacaagtttatggtcaaactggtcttccttatggtcctggatctgaag
ctgcagcatttggttctaatggtctaggaggttttcctccacttaatagtgtatctccaagtccactaatgctactacatccaccaccacaactatcacctttttctacaacctcatggtcaacaagtac
cttattatctagaaaatgaacctagtggttatacagtacgtgaagctggtccaccagcatttttatagaccaaattcagataatcgacgtcaaggtggtagagaaagattagctagtactaatgata
aaggaagtatggctatggaatctgctaaagaaactcgttattgtgcagtatgtaatgattatgcttcaggttatcattatggagtttggtcttgtgaaggttgtaaagctttttttaaaagaagtattca
aggacataatgattatatgtgtccagctactaatcaatgtactattgataaaaatagaagaaaaagttgtcaagcttgtcgacttcgtaaatgttatgaagtaggaatgatgaaaggtggaatacg
aaaagatcgaagaggaggaagaatgttaaaacataaacgtcaaagagatgatggagaaggtagaggtgaagtaggatctgctggagatatgagagctgctaatctttggccaagtccactt
atgattaaacgttctaaaaaaaaatagtctagctttatctctaacagctgatcaaatggttagtgctttattagatgctgaacctcctatactttattctgaatatgatcctactagacctttttagtgaagct
tcaatgatgggtttactaactaatctagcagatagagaactagttcatatgattaattgggcaaaaagagtaccaggttttgtagatttaactcttcatgatcaagttcatcttctagaatgtgcttggc
tagaaattctaatgattggtcttgtttggcgttctatggaacatccaggaaaactactatttgctcctaatttacttttagatagaaatcaaggaaaatgtgtagaaggtatggtagaaatttttgatatg
ctactagctacatcatctcgatttcgtatgatgaatctacaaggagaagaatttgtatgtcttaaatctattatttttacttaattctggagtatatacatttctatctagtactctaaaatctctagaagaaa
aagatcatattcatcgagttctagataaaattacagatactttaattcatctaatggctaaagcaggtctaactctacaacaacaacatcaacgactagctcaacttcttcttattctttctcatattaga
catatgagtaataaaggtatggaacatctatatagtatgaaatgtaaaaatgtagtacctctttatgatctactactagaaatgctagatgctcatcgtctacatgcacctactagtcgtggaggag
catctgtagaagaaacagatcaaagtcatttagctactgcaggttctacttcatcacattctttacaaaaatattatatacaggagaagcagaaggttttcctgctacagtttga

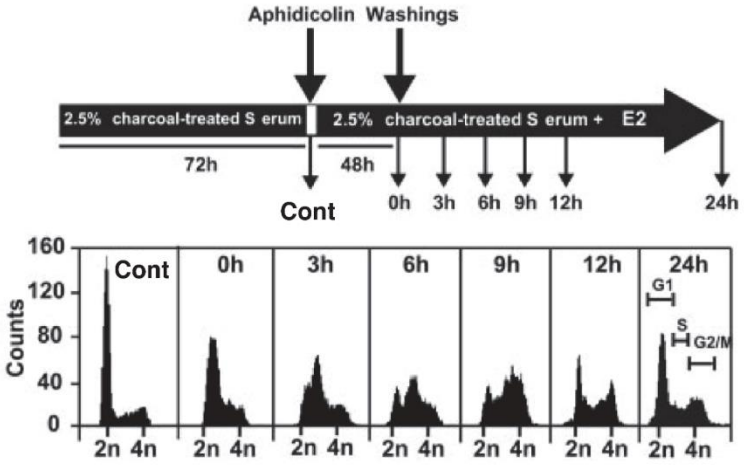
Supplementary Figure 1: Coding sequence of ERα SYN 100% GC3, ERα SYN 50% GC3 and ERα SYN 0% GC3 cDNAs.



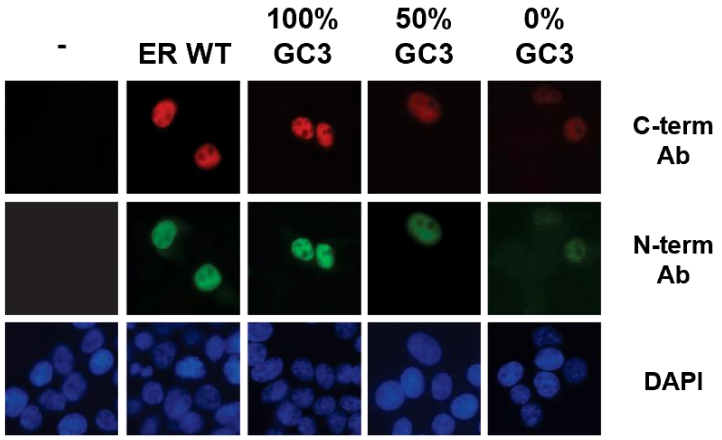
Supplementary Figure 2: Comparison of the codon usage frequency deduced from human cyclin D1, cyclin E1, cyclin A1 and cyclin B1 coding sequences with the codon usage frequency in “M phase of mitotic cell cycle” and “pattern specification” gene sets from Gingold et al.’s study (Gingold et al., 2014). GC3 and AU3 codons are represented by different symbols. The correlation coefficient is indicated.



Supplementary Figure 3: mRNA distribution depending on their composition in G or C at the first (GC1) or second (GC2) base of codons. Transcripts were ranked according to their percentage in GC1 (first line) or GC2 (second line), grouped in windows of 10 units from 10 to 100. The number of transcripts per window is then expressed as a percentage of the total transcripts examined. Analysis performed on differentially expressed mRNAs between G0/G1 phase, G1/S phase transition and M phase of synchronized MCF7 cells (Liu et al., 2017).



Supplementary Figure 4: Scheme of the cell cycle synchronization protocol with aphidicolin and validation of the synchronization of MCF7 cells by flow cytometry.



Supplementary Figure 5: Immunofluorescent detection of ERα WT and SYN proteins by a N-terminal and C-terminal ERα antibody (Ab), 48 h after transient transfection of the corresponding expression vectors in HEK-293 cells. Exposure time is different between each ERα form. The nuclei are stained with DAPI.

Supplementary Table 1: Codon usage frequencies calculated from the abundance of amino acids deduced from human protein sequences specific of proliferating or differentiating states.

Amino acid	Codon	Frequency	
		Proliferation	Differentiation
ala	gca	0.022	0.011
ala	gcc	0.021	0.032
ala	gcg	0.004	0.016
ala	gct	0.022	0.010
arg	aga	0.018	0.006
arg	agg	0.011	0.009
arg	cga	0.008	0.005
arg	cgc	0.006	0.019
arg	cgg	0.009	0.015
arg	cgt	0.005	0.003
asn	aac	0.015	0.027
asn	aat	0.021	0.009
asp	gac	0.020	0.035
asp	gat	0.027	0.012
cys	tgc	0.009	0.017
cys	tgt	0.014	0.006
gln	caa	0.017	0.008
gln	cag	0.030	0.039
glu	gaa	0.038	0.017
glu	gag	0.031	0.051
gly	gga	0.023	0.009
gly	ggc	0.017	0.032
gly	ggg	0.013	0.018
gly	ggt	0.013	0.007
his	cac	0.012	0.020
his	cat	0.014	0.006
ile	ata	0.010	0.004
ile	atc	0.015	0.031
ile	att	0.019	0.010
leu	cta	0.010	0.005
leu	ctc	0.014	0.026
leu	ctg	0.028	0.049
leu	ctt	0.017	0.008
leu	tta	0.013	0.004
leu	ttg	0.017	0.008
lys	aaa	0.028	0.015
lys	aag	0.028	0.042
met	atg	0.022	0.022
phe	ttc	0.016	0.026
phe	ttt	0.022	0.012
pro	cca	0.021	0.011
pro	ccc	0.015	0.024
pro	ccg	0.004	0.015
pro	cct	0.021	0.012
ser	agc	0.015	0.027
ser	agt	0.016	0.007
ser	tca	0.015	0.007
ser	tcc	0.013	0.019
ser	tcg	0.003	0.012
ser	tct	0.018	0.009
thr	aca	0.017	0.010
thr	acc	0.014	0.024
thr	acg	0.004	0.012
thr	act	0.018	0.007
trp	tgg	0.013	0.013
tyr	tac	0.011	0.020
tyr	tat	0.016	0.007
val	gta	0.011	0.005
val	gtc	0.010	0.018
val	gtg	0.024	0.032
val	gtt	0.016	0.006

Supplementary Table 2: Ontology (GO – biological process) and pathway (KEGG) enrichment analysis of genes depending on their GC3%, grouped in three windows: 0 to 45%, 45 to 65%, and 65 to 100% GC3.

GO - biological process

From 0 to 45% GC3						
Gene Set	**Description**	**Size**	**Expect**	**Ratio**	**P Value**	**FDR**
GO:0007049	cell cycle	1739	367.08	1.5010	0	0
GO:0051276	chromosome organization	1143	241.27	1.5501	0	0
GO:0006259	DNA metabolic process	970	204.76	1.6556	0	0
GO:0000278	mitotic cell cycle	927	195.68	1.5638	0	0
GO:0006974	cellular response to DNA damage stimulus	806	170.14	1.6810	0	0
GO:0006281	DNA repair	511	107.87	1.8263	0	0
GO:0007059	chromosome segregation	312	65.860	2.1257	0	0
GO:0033554	cellular response to stress	1867	394.10	1.3550	2.2204e-16	1.5528e-13
GO:0006397	mRNA processing	487	102.80	1.7510	2.2204e-16	1.5528e-13
GO:0000280	nuclear division	420	88.657	1.8160	2.2204e-16	1.5528e-13
GO:0048285	organelle fission	459	96.890	1.7752	3.3307e-16	2.1628e-13
GO:0016071	mRNA metabolic process	765	161.48	1.5729	1.1102e-15	5.9371e-13
GO:0051321	meiotic cell cycle	248	52.350	2.0439	3.1086e-15	1.5700e-12
GO:0007017	microtubule-based process	729	153.88	1.5726	5.9952e-15	2.8685e-12
GO:0051726	regulation of cell cycle	1106	233.46	1.4520	7.9936e-15	3.6335e-12
GO:0008380	RNA splicing	417	88.024	1.7722	8.9928e-15	3.8930e-12
GO:0070647	protein modification by small protein conjugation or removal	1029	217.21	1.4686	1.0214e-14	4.2207e-12
GO:0006260	DNA replication	268	56.572	1.9621	2.8977e-14	1.1453e-11
GO:0010564	regulation of cell cycle process	689	145.44	1.5677	5.5622e-14	2.1069e-11
GO:0006403	RNA localization	228	48.128	2.0362	6.0729e-14	2.2084e-11
GO:0044772	mitotic cell cycle phase transition	487	102.80	1.6829	7.3941e-14	2.4512e-11
GO:0051169	nuclear transport	356	75.147	1.8098	7.5495e-14	2.4512e-11
GO:0044770	cell cycle phase transition	526	111.03	1.6302	5.2791e-13	1.5481e-10
GO:0051052	regulation of DNA metabolic process	405	85.491	1.7195	8.6064e-13	2.3012e-10

From 45 to 65% GC3						
Gene Set	**Description**	**Size**	**Expect**	**Ratio**	**P Value**	**FDR**
GO:0051606	detection of stimulus	632	208.45	1.6598	0	0
GO:0050906	detection of stimulus involved in sensory perception	478	157.66	1.8711	0	0
GO:0007606	sensory perception of chemical stimulus	473	156.01	1.8332	0	0
GO:0007608	sensory perception of smell	404	133.25	1.9662	0	0
GO:0007186	G protein-coupled receptor signaling pathway	1290	425.48	1.2903	4.7518e-14	4.7998e-11
GO:0050877	nervous system process	1348	444.62	1.2078	2.1488e-8	0.000019535
GO:0006613	cotranslational protein targeting to membrane	99	32.653	1.7762	1.4077e-7	0.00011634
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	95	31.334	1.7872	1.7235e-7	0.00013057
GO:0072599	establishment of protein localization to endoplasmic reticulum	111	36.611	1.7208	2.0817e-7	0.00014557
GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	120	39.580	1.6422	0.0000013159	0.00074766
GO:0033141	positive regulation of peptidyl-serine phosphorylation of STAT protein	20	6.5967	2.5771	0.0000023791	0.0012723
GO:0006412	translation	613	202.19	1.2414	0.000015698	0.0079282
GO:0043043	peptide biosynthetic process	636	209.77	1.2347	0.000018393	0.0088003
GO:0043604	amide biosynthetic process	766	252.65	1.2112	0.000020569	0.0093496
GO:0033139	regulation of peptidyl-serine phosphorylation of STAT protein	22	7.2563	2.3428	0.000026217	0.011349
GO:0000956	nuclear-transcribed mRNA catabolic process	206	67.946	1.4129	0.000030374	0.012551
GO:0052696	flavonoid glucuronidation	9	2.9685	3.0318	0.000045998	0.018181
GO:0022613	ribonucleoprotein complex biogenesis	440	145.13	1.2679	0.000052969	0.020064
GO:0051607	defense response to virus	235	77.511	1.3676	0.000065288	0.023741
GO:0006518	peptide metabolic process	780	257.27	1.1894	0.00010123	0.035397
GO:0043603	cellular amide metabolic process	1033	340.72	1.1593	0.00013797	0.045094
GO:0006538	glutamate catabolic process	8	2.6387	3.0318	0.00013960	0.045094
GO:0002250	adaptive immune response	382	126.00	1.2699	0.00014385	0.045094

From 65 to 100% GC3						
Gene Set	**Description**	**Size**	**Expect**	**Ratio**	**P Value**	**FDR**
GO:0009888	tissue development	1925	773.42	1.3201	0	0
GO:2000026	regulation of multicellular organismal development	1908	766.59	1.2523	0	0
GO:0051049	regulation of transport	1765	709.13	1.2565	0	0
GO:0045595	regulation of cell differentiation	1699	682.62	1.2745	0	0
GO:0006811	ion transport	1608	646.06	1.3157	0	0
GO:0007267	cell-cell signaling	1576	633.20	1.3771	0	0
GO:1901700	response to oxygen-containing compound	1556	625.16	1.2557	0	0
GO:0022008	neurogenesis	1551	623.15	1.2966	0	0
GO:0055085	transmembrane transport	1491	599.05	1.2720	0	0
GO:0030182	neuron differentiation	1313	527.53	1.3023	0	0
GO:0060429	epithelium development	1213	487.35	1.3276	0	0
GO:0051241	negative regulation of multicellular organismal process	1143	459.23	1.3240	0	0
GO:0048878	chemical homeostasis	1119	449.59	1.3323	0	0
GO:0009790	embryo development	980	393.74	1.3384	0	0
GO:0009887	animal organ morphogenesis	974	391.33	1.4055	0	0
GO:0051093	negative regulation of developmental process	905	363.61	1.3449	0	0
GO:0030855	epithelial cell differentiation	747	300.13	1.3961	0	0
GO:0099536	synaptic signaling	698	280.44	1.4156	0	0
GO:0045596	negative regulation of cell differentiation	678	272.40	1.3876	0	0
GO:0007389	pattern specification process	433	173.97	1.5233	0	0
GO:0042592	homeostatic process	1817	730.03	1.2219	2.2204e-16	5.1759e-14
GO:0010647	positive regulation of cell communication	1733	696.28	1.2280	2.2204e-16	5.1759e-14
GO:0051240	positive regulation of multicellular organismal process	1661	667.35	1.2347	2.2204e-16	5.1759e-14
GO:0045944	positive regulation of transcription by RNA polymerase II	1168	469.27	1.2828	2.2204e-16	5.1759e-14
GO:0048646	anatomical structure formation involved in morphogenesis	1067	428.70	1.2970	2.2204e-16	5.1759e-14
GO:0023056	positive regulation of signaling	1739	698.69	1.2252	4.4409e-16	9.8469e-14

KEGG

From 0 to 45% GC3						
Gene Set	**Description**	**Size**	**Expect**	**Ratio**	**P Value**	**FDR**
hsa03013	RNA transport	171	33.976	2.2663	4.1633e-14	1.3572e-11
hsa04120	Ubiquitin mediated proteolysis	137	27.220	2.2777	1.0127e-11	1.6508e-9
hsa00640	Propanoate metabolism	32	6.3580	2.9884	0.0000010048	0.00010919
hsa03022	Basal transcription factors	45	8.9410	2.5724	0.0000026929	0.00021947
hsa03018	RNA degradation	79	15.696	2.1024	0.0000063845	0.00038142
hsa03440	Homologous recombination	41	8.1462	2.5779	0.0000070199	0.00038142
hsa04110	Cell cycle	124	24.637	1.8265	0.000013290	0.00061894
hsa03015	mRNA surveillance pathway	91	18.081	1.9358	0.000029264	0.0011925
hsa04114	Oocyte meiosis	124	24.637	1.7453	0.000071749	0.0025989
hsa04141	Protein processing in endoplasmic reticulum	165	32.784	1.6167	0.00011536	0.0034861
hsa03040	Spliceosome	134	26.624	1.6902	0.00011763	0.0034861
hsa03420	Nucleotide excision repair	47	9.3383	2.0346	0.00094452	0.023789
hsa03460	Fanconi anemia pathway	54	10.729	1.9573	0.00094863	0.023789
hsa00280	Valine, leucine and isoleucine degradation	48	9.5370	1.9922	0.0012766	0.029727
hsa00500	Starch and sucrose metabolism	36	7.1528	2.0971	0.0022363	0.048602
From 45 to 65% GC3						
Gene Set	**Description**	**Size**	**Expect**	**Ratio**	**P Value**	**FDR**
hsa04740	Olfactory transduction	448	158.59	1.8097	0	0
hsa04650	Natural killer cell mediated cytotoxicity	131	46.374	1.4448	0.00014026	0.022863
hsa03010	Ribosome	153	54.161	1.3847	0.00032564	0.035386

From 65 to 100% GC3						
Gene Set	**Description**	**Size**	**Expect**	**Ratio**	**P Value**	**FDR**
hsa04080	Neuroactive ligand-receptor interaction	277	113.56	1.5234	1.8674e-13	6.0877e-11
hsa04020	Calcium signaling pathway	183	75.023	1.5862	2.7896e-11	4.5471e-9
hsa05322	Systemic lupus erythematosus	133	54.525	1.6690	1.0732e-10	8.7466e-9
hsa05217	Basal cell carcinoma	63	25.828	1.8198	5.6860e-8	0.0000037072
hsa05224	Breast cancer	147	60.264	1.4934	4.7564e-7	0.000025843
hsa04934	Cushing syndrome	154	63.134	1.4731	7.4654e-7	0.000034767
hsa04010	MAPK signaling pathway	295	120.94	1.3313	0.0000011015	0.000044885
hsa04925	Aldosterone synthesis and secretion	96	39.356	1.5754	0.0000023108	0.000083702
hsa04024	cAMP signaling pathway	199	81.582	1.3728	0.0000075260	0.00019727
hsa04390	Hippo signaling pathway	154	63.134	1.4255	0.0000077667	0.00019727
hsa04921	Oxytocin signaling pathway	152	62.314	1.4282	0.0000078373	0.00019727
hsa04911	Insulin secretion	85	34.847	1.5783	0.0000078664	0.00019727
hsa04015	Rap1 signaling pathway	206	84.452	1.3617	0.0000095645	0.00022272
hsa04261	Adrenergic signaling in cardiomyocytes	144	59.034	1.4229	0.000017082	0.00034805
hsa05226	Gastric cancer	149	61.084	1.4079	0.000023803	0.00043462
hsa04750	Inflammatory mediator regulation of TRP channels	99	40.586	1.5030	0.000025000	0.00043462
hsa04550	Signaling pathways regulating pluripotency of stem cells	139	56.985	1.4214	0.000025330	0.00043462
hsa04270	Vascular smooth muscle contraction	121	49.605	1.4515	0.000026842	0.00043753
hsa05200	Pathways in cancer	526	215.64	1.2011	0.000045221	0.00067010
hsa04916	Melanogenesis	101	41.406	1.4732	0.000058983	0.00080904
hsa05166	Human T-cell leukemia virus 1 infection	255	104.54	1.2914	0.000059562	0.00080904
hsa05225	Hepatocellular carcinoma	168	68.873	1.3503	0.00010230	0.0012828
hsa04062	Chemokine signaling pathway	189	77.483	1.3293	0.00010231	0.0012828
hsa04927	Cortisol synthesis and secretion	64	26.238	1.5626	0.00015541	0.0018765
hsa05202	Transcriptional misregulation in cancer	186	76.253	1.3114	0.00025198	0.0029064
hsa04810	Regulation of actin cytoskeleton	213	87.322	1.2826	0.00035093	0.0038134
hsa04913	Ovarian steroidogenesis	49	20.088	1.5930	0.00049098	0.0045731

References

Arango D, Sturgill D, Alhusaini N, Dillman AA, Sweet TJ, Hanson G, … Oberdoerffer S (2018) Acetylation of Cytidine in mRNA Promotes Translation Efficiency *Cell* **175**, 1872-1886.e24.

Arnal J-F, Lenfant F, Metivier R, Flouriot G, Henrion D, Adlanmerini M, … Katzenellenbogen J (2017) Membrane and Nuclear Estrogen Receptor Alpha Actions: From Tissue Specificity to Medical Implications *Physiol Rev* **97**, 1045–1087.

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, … Soboleva A (2012) NCBI GEO: archive for functional genomics data sets--update *Nucleic Acids Res* **41**, D991–D995.

Beato M (1989) Gene regulation by steroid hormones *Cell* **56**, 335–344.

Bohlen J, Roiuk M, and Teleman AA (2021) Phosphorylation of ribosomal protein S6 differentially affects mRNA translation based on ORF length *Nucleic Acids Res* **49**, 13062–13074.

Bornelöv S, Selmi T, Flad S, Dietmann S, and Frye M (2019) Codon usage optimization in pluripotent embryonic stem cells *Genome Biol* **20**, 119.

Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adélaïde J, Cervera N, … Bertucci F (2006) Gene expression profiling of breast cell lines identifies potential new basal markers *Oncogene* **25**, 2273–2284.

Chung S, Andersson T, Sonntag K-C, Björklund L, Isacson O, and Kim K-S (2002) Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines *Stem Cells Dayt Ohio* **20**, 139–145.

Courel M, Clément Y, Bossevain C, Foretek D, Vidal Cruchez O, Yi Z, … Weil D (2019) GC content shapes mRNA storage and decay in human cells *ELife* **8**, e49708.

Couse JF, and Korach KS (1999) Estrogen receptor null mice: what have we learned and where will they lead us? *Endocr Rev* **20**, 358–417.

Droog M, Mensink M, and Zwart W (2016) The Estrogen Receptor α-Cistrome Beyond Breast Cancer *Mol Endocrinol* **30**, 1046–1058.

Edwards DP (2005) Regulation of signal transduction pathways by estrogen and progesterone *Annu Rev Physiol* **67**, 335–376.

Erales J, Marchand V, Panthu B, Gillot S, Belin S, Ghayad SE, … Diaz J-J (2017) Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes *Proc Natl Acad Sci U S A* **114**, 12934–12939.

Evans RM (1988) The steroid and thyroid hormone receptor superfamily *Science* **240**, 889–895.

Fernández-Calero T, Astrada S, Alberti Á, Horjales S, Arnal JF, Rovira C, … Marin M (2014) The transcriptional activities and cellular localization of the human estrogen receptor alpha are affected by the synonymous Ala87 mutation *J Steroid Biochem Mol Biol* **143**, 99–104.

Fernández-Calero T, Davyt M, Perelmuter K, Chalar C, Bampi G, Persson H, … Marín M (2020) Fine-tuning the metabolic rewiring and adaptation of translational machinery during an epithelial-mesenchymal transition in breast cancer cells *Cancer Metab* **8**, 8.

Fornasiero EF, and Rizzoli SO (2019) Pathological changes are associated with shifts in the employment of synonymous codons at the transcriptome level *BMC Genomics* **20**, 566.

Gardini S, Cheli S, Baroni S, Di Lascio G, Mangiavacchi G, Micheletti N, … Niccolai N (2016) On Nature's Strategy for Assigning Genetic Code Multiplicity *PloS One* **11**, e0148174.

Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, … Pilpel Y (2014) A dual program for translation regulation in cellular proliferation and differentiation *Cell* **158**, 1281–1292.

Goodarzi H, Nguyen HCB, Zhang S, Dill BD, Molina H, and Tavazoie SF (2016) Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression *Cell* **165**, 1416–1427.

Guimaraes JC, Mittal N, Gnann A, Jedlinski D, Riba A, Buczak K, … Zavolan M (2020) A rare codon-based translational program of cell proliferation *Genome Biol* **21**, 44.

Hanson G, and Coller J (2018) Codon optimality, bias and usage in translation and mRNA decay *Nat Rev Mol Cell Biol* **19**, 20–30.

Henglein B, Chenivesse X, Wang J, Eick D, and Bréchot C (1994) Structure and cell cycle-regulated transcription of the human cyclin A gene *Proc Natl Acad Sci U S A* **91**, 5490–5494.

Hia F, Yang SF, Shichino Y, Yoshinaga M, Murakawa Y, Vandenbon A, … Takeuchi O (2019) Codon bias confers stability to human mRNAs *EMBO Rep* **20**, e48220.

Horjales S, Cota G, Señorale-Pose M, Rovira C, Román E, Artagaveytia N, … Marín M (2007) Translational machinery and protein folding: Evidence of conformational variants of the estrogen receptor alpha *Arch Biochem Biophys* **467**, 139–143.

Huet G, Mérot Y, Percevault F, Tiffoche C, Arnal J-F, Boujrad N, … Flouriot G (2009) Repression of the estrogen receptor-alpha transcriptional activity by the Rho/megakaryoblastic leukemia 1 signaling pathway *J Biol Chem* **284**, 33729–33739.

Imami K, Milek M, Bogdanow B, Yasuda T, Kastelic N, Zauber H, … Selbach M (2018) Phosphorylation of the Ribosomal Protein RPL12/uL11 Affects Translation during Mitosis *Mol Cell* **72**, 84-98.e9.

Jehanno C, Fernandez-Calero T, Habauzit D, Avner S, Percevault F, Jullion E, … Flouriot G (2020) Nuclear accumulation of MKL1 in luminal breast cancer cells impairs genomic activity of ERα and is associated with endocrine resistance *Biochim Biophys Acta Gene Regul Mech* 194507.

Jehanno C, Percevault F, Boujrad N, Le Goff P, Fontaine C, Arnal J-F, … Flouriot G (2021) Nuclear translocation of MRTFA in MCF7 breast cancer cells shifts ERα nuclear/genomic to extra-nuclear/non genomic actions *Mol Cell Endocrinol* **530**, 111282.

Jordan VC, and O'Malley BW (2007) Selective estrogen-receptor modulators and antihormonal resistance in breast cancer *J Clin Oncol Off J Am Soc Clin Oncol* **25**, 5815–5824.

Kerdivel G, Boudot A, Habauzit D, Percevault F, Demay F, Pakdel F, and Flouriot G (2014) Activation of the MKL1/actin signaling pathway induces hormonal escape in estrogen-responsive breast cancer cell lines *Mol Cell Endocrinol* **390**, 34–44.

King TH, Liu B, McCully RR, and Fournier MJ (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center *Mol Cell* **11**, 425–435.

Kudla G, Lipinski L, Caffin F, Helwak A, and Zylicz M (2006) High guanine and cytosine content increases mRNA levels in mammalian cells *PLoS Biol* **4**, e180.

Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, … Weirauch MT (2018) The Human Transcription Factors *Cell* **172**, 650–665.

Liu Y, Chen S, Wang S, Soares F, Fischer M, Meng F, … He HH (2017) Transcriptional landscape of the human cell cycle *Proc Natl Acad Sci U S A* **114**, 3473–3478.

Love MI, Huber W, and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 *Genome Biol* **15**, 550.

Ly T, Ahmad Y, Shlien A, Soroka D, Mills A, Emanuele MJ, … Lamond AI (2014) A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells *ELife* **3**, e01630.

Mérot Y, Métivier R, Penot G, Manu D, Saligaut C, Gannon F, … Flouriot G (2004) The Relative Contribution Exerted by AF-1 and AF-2 Transactivation Functions in Estrogen Receptor α Transcriptional Activity Depends upon the Differentiation Stage of the Cell *J Biol Chem* **279**, 26184–26191.

Novoa EM, and Ribas de Pouplana L (2012) Speeding with control: codon usage, tRNAs, and ribosomes *Trends Genet TIG* **28**, 574–581.

Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, … Botstein D (2000) Molecular portraits of human breast tumours *Nature* **406**, 747–752.

Porras L, Ismail H, and Mader S (2021) Positive Regulation of Estrogen Receptor Alpha in Breast Tumorigenesis *Cells* **10**, 2966.

Poulard C, Treilleux I, Lavergne E, Bouchekioua-Bouzaghou K, Goddard-Léon S, Chabaud S, … Le Romancer M (2012) Activation of rapid oestrogen signalling in aggressive human breast cancers *EMBO Mol Med* **4**, 1200–1213.

Pouyet F, Mouchiroud D, Duret L, and Sémon M (2017) Recombination, meiotic expression and human codon usage *ELife* **6**, e27344.

Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, … Coller J (2015) Codon optimality is a major determinant of mRNA stability *Cell* **160**, 1111–1124.

Radhakrishnan A, Chen Y-H, Martin S, Alhusaini N, Green R, and Coller J (2016) The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality *Cell* **167**, 122-132.e9.

Rao Y, Wang Z, Luo W, Sheng W, Zhang R, and Chai X (2018) Base composition is the primary factor responsible for the variation of amino acid usage in zebra finch (Taeniopygia guttata) *PLOS ONE* **13**, e0204796.

Rau A, Gallopin M, Celeux G, and Jaffrézic F (2013) Data-based filtering for replicated high-throughput transcriptome sequencing experiments *Bioinformatics* **29**, 2146–2152.

Rodnina MV (2016) The ribosome in action: Tuning of translational efficiency and protein folding *Protein Sci Publ Protein Soc* **25**, 1390–1406.

Rudolph KLM, Schmitt BM, Villar D, White RJ, Marioni JC, Kutter C, and Odom DT (2016) Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States *PLOS Genet* **12**, e1006024.

Rusidzé M, Adlanmérini M, Chantalat E, Raymond-Letron I, Cayre S, Arnal J-F, … Lenfant F (2021) Estrogen receptor-α signaling in post-natal mammary development and breast cancers *Cell Mol Life Sci* **78**, 5681–5705.

Shi Z, Fujii K, Kovary KM, Genuth NR, Röst HL, Teruel MN, and Barna M (2017) Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide *Mol Cell* **67**, 71-83.e7.

Shu H, Donnard E, Liu B, Jung S, Wang R, and Richter JD (2020) FMRP links optimal codons to mRNA stability in neurons *Proc Natl Acad Sci U S A* **117**, 30400–30411.

Slavov N, Semrau S, Airoldi E, Budnik B, and van Oudenaarden A (2015) Differential Stoichiometry among Core Ribosomal Proteins *Cell Rep* **13**, 865–873.

Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JGM, … Martens JWM (2008) Subtypes of breast cancer show preferential site of relapse *Cancer Res* **68**, 3108–3114.

Stumpf CR, Moreno MV, Olshen AB, Taylor BS, and Ruggero D (2013) The translational landscape of the mammalian cell cycle *Mol Cell* **52**, 574–582.

Tanenbaum ME, Stern-Ginossar N, Weissman JS, and Vale RD (2015) Regulation of mRNA translation during mitosis *ELife* **4**.

Tsai C-J, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, and Nussinov R (2008) Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima *J Mol Biol* **383**, 281–291.

Wang R, Liang J, Jiang H, Qin L-J, and Yang H-T (2008) Promoter-dependent EGFP expression during embryonic stem cell propagation and differentiation *Stem Cells Dev* **17**, 279–289.

Wu Q, Medina SG, Kushawah G, DeVore ML, Castellano LA, Hand JM, … Bazzini AA (2019) Translation affects mRNA stability in a codon-dependent manner in human cells *ELife* **8**, e45396.

Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, and Liu Y (2015) Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding *Mol Cell* **59**, 744–754.

Zheng Y, and Murphy LC (2016) Regulation of Steroid Hormone Receptors and Coregulators during the Cell Cycle Highlights Potential Novel Function in Addition to Roles as Transcription Factors *Nucl Recept Signal* **14**, nrs.14001.

3. LES RIBOSOMES SPECIALISES SONT-ILS IMPLIQUES DANS LA REGULATION DE LA TRADUCTION EN FONCTION DE L'USAGE DES CODONS ?

La transcription et la traduction sont des processus coordonnés permettant la production des protéines nécessaires au fonctionnement cellulaire, dont la régulation précise n'est toujours pas entièrement comprise. Il semblerait notamment que le biais d'usage des codons aurait un rôle fonctionnel dans cette coordination : un enrichissement en GC serait par exemple lié à la stabilisation de l'ARNm et au repliement co-traductionnel de certaines protéines. La nature de la troisième base des codons (GC3 ou AT3) semblerait elle jouer un rôle particulier dans l'expression des gènes en protéines en fonction de l'état cellulaire. De quelle manière le biais d'usage des codons participe à la régulation de la traduction reste cependant inconnu. Des ribosomes spécialisés pourraient être impliqués dans la traduction d'ARNm spécifiques en fonction de leur composition en codon, mais leur découverte reste controversée. Un tel mécanisme de régulation de la traduction pourrait affecter l'expression et l'activité de nombreuses protéines, dont le récepteur aux œstrogènes ERα. Ce récepteur nucléaire est le principal médiateur de la signalisation des œstrogènes et joue un rôle majeur dans le développement du cancer du sein. Dans ce cas, l'hormonothérapie est utilisée mais les cellules cancéreuses mammaires échappent au contrôle hormonal pour 30% des patientes. Cette résistance hormonale est liée à des modifications d'expression et d'activité d'ERα, dont tous les mécanismes moléculaires sous-jacents n'ont pas été élucidés. Dans de précédentes études, nous avons montré que l'usage des codons contrôle l'expression et les propriétés fonctionnelles d'ERα. Nous avons ensuite identifié un programme traductionnel basé sur la troisième base des codons, favorisant l'expression de protéines riches en GC3 en cellules quiescentes, et au contraire celle de protéines riches en AT3 lors des phases S à M du cycle cellulaire. Afin de mieux comprendre cette régulation de la traduction en fonction de l'usage des codons et de l'état cellulaire, des expériences de polysome profiling suivi de séquençage de l'ARN et de spectrométrie de masse ont été réalisées dans des cellules de cancer du sein ERα+ quiescentes ou transformées. Les résultats obtenus mettent en évidence une dynamique globale de la traduction au cours du cycle cellulaire, favorisant une traduction de messagers riches en AU3 médiée par les polysomes en phase S. Des modifications de la composition de la machinerie traductionnelle ont été identifiées entre les cellules quiescentes et transformées, qui pourraient expliquer le résultat précédent. La protéine ribosomale eS26 apparait en effet enrichie dans les ribosomes issus de cellules transformées, et pourrait jouer un rôle dans la traduction d'ARNm spécifiques en fonction de leur composition en codons, suggérant l'existence de ribosomes spécialisés dans un contexte dédifférencié. Les protéines liant l'ARN en fonction de son enrichissement en GC3, DDX6 et RBM47, sont quant à elles prépondérantes dans les cellules quiescentes, et pourraient ainsi être impliquées dans la répression de la traduction de messagers riches en AU3 en cellules différenciées. Aussi, des acteurs du cytosquelette sont enrichis dans notre modèle de cellules de cancer du sein ERα+ transformées, or des interactions avec la machinerie traductionnelle semblent avoir lieu et participer à la régulation de la traduction. Ces résultats ouvrent alors la voie à de nouvelles études pour

comprendre les mécanismes à l'œuvre dans la régulation de la traduction en fonction de l'enrichissement en GC3 de la séquence codante des ARNm et de l'état cellulaire.

Are specialized ribosomes involved in the regulation of translation depending on codon usage?

Abstract

Cell functions require the production of proteins and its modulation, which is achieved through the coordinated processes of transcription and translation. Their precise regulation depending on cell state still involves mysteries, but novel insights into these processes may have been shed through the recent discovery that a codon usage bias could be involved. Hence, such a mechanism of translation regulation could affect the expression and activity of many proteins, including the estrogen receptor ERα, a major player in breast cancer development and endocrine resistance. In previous studies, we demonstrated that codon usage controls ERα expression and function. In addition, we showed that a GC3-based translational program favors the expression of proteins with GC3-rich coding sequences in quiescent cells whilst proteins with AT3-rich coding sequences are preferentially found in S to M phases of the cell cycle. In this study, to better understand the regulation of translation depending on codon usage and cell state, we used ERα-positive breast cancer cells which were either synchronized in different phases of their cell cycle or forced to undergo an epithelial-mesenchymal transition. We then performed polysome profiling experiments and RNA sequencing to characterize the mRNAs translated by monosomes and polysomes depending on cell state. In addition, mass spectrometry was used on purified ribosomes to determine the composition of the translational machinery in quiescent versus transformed breast cancer cells. Through these combined approaches we identified dynamics in polysome-mediated translation during cell cycle, favoring the translation of AU3-rich mRNAs in S phase. Modifications of the translational machinery composition could explain these results, as, importantly, we detected specialized ribosomes in transformed cells. This study paves thus the way for better understanding translation regulation mechanisms that could be implicated in pathological conditions such as endocrine resistance in breast cancer.

Keywords: Codon usage; polysomes; translation regulation; specialized ribosomes; riboproteome

1. Introduction

Transcription and translation are the fundamental processes enabling genes and proteins expression modulations required for cell homeostasis and adaptation to physiological environment. Despite their study for decades, mysteries persist regarding their precise coordinated regulation. In recent years, it was reported that codon usage is biased, with some codons being more used in coding sequences than other for the same amino acid. Strikingly, it was shown that this codon usage bias differs between cell states and pathological conditions, particularly between differentiated and proliferating cells. These observations rise questions

regarding the functional role of such bias in gene expression and protein production (Fornasiero and Rizzoli, 2019; Gingold et al., 2014; Meyer et al., 2021). GC-enrichment was notably associated with mRNA stabilization (Courel et al., 2019; Hia et al., 2019; Kudla et al., 2006), and codon usage appears to be involved in the regulation of the co-translational folding of proteins by impacting their translation speed (Stein and Frydman, 2019). More precisely, our last study highlighted the role played by the third codon base in the regulation of gene expression and translation depending on cell state. A GC3-based translational program seems indeed to exist to ensure the production of proteins throughout the cell cycle, with GC3-rich sequences associated with proteins expressed in G0/G1 phase and GC3-poor (AT3-rich) sequences corresponding to proteins specifically expressed during S to M phases (chapter 2). Similarly, a switch from proteins with a GC3-rich coding region to proteins with a GC3-poor coding sequence has been observed in the progression of several diseases, including cancer (Fornasiero and Rizzoli, 2019) (chapter 2). How exactly this codon usage parameter regulates protein production is however not known. A coordination with tRNA abundance could be involved (Dittmar et al., 2005; Goodarzi et al., 2016), but some studies reported an absence of major changes in tRNA pool composition between cell states (Guimaraes et al., 2020; Pouyet et al., 2017; Rudolph et al., 2016). In agreement with these last studies, our own models of native versus transformed breast cancer cells do not show major variation in their tRNA pool composition (Marin M., unpublished data). The regulation of translation by codon usage seems then to rely on other parameters than its adequacy with the available tRNAs in a given cell state. The concept of specialized ribosomes emerged notably in recent years, with evidences that ribosomes differing in their core composition, post-translational modifications, or rRNA modifications coexist in cells, enabling the translation of specific mRNAs. This discovery however remains controversial, with conflicting results obtained, notably depending on cell type (Dalla Venezia et al., 2019; Papagiannopoulos et al., 2022). Codon usage and ribosomes appear then to play a role in the regulation of translation, impacting cell proteome depending on cell state.

Such a phenomenon can affect the expression and activity of many proteins, including the estrogen receptor alpha (ERα). ERα is a nuclear receptor mediating estrogen signaling. Its activation results in genes regulations promoting cell growth and proliferation. Accordingly, alterations in ERα expression or activity have major impact on breast cancer development (Arnal et al., 2017; Manavathi et al., 2013). In case of ERα-positive breast cancer, endocrine therapy exists to deprive the tumor in estrogen or directly block ERα. A resistance is however observed for 30% of the patients, where ERα escapes hormonal control and enables breast cancer development in an estrogen-independent manner (Hanker et al., 2020). ERα action relies primarily on its activity at the genome level, but a non-genomic activity exists as well (Rusidzé et al., 2021). Whereas its genomic activity is generally predominant in cells expressing high amounts of nuclear ERα, a low expression of the receptor appears to be linked to an increased non-genomic signaling. This non-genomic activity being notably associated with resistance to endocrine therapy (Poulard et al., 2012), it is of great interest to understand how ERα expression and properties are regulated. In line with the hypothesis that codon usage plays a role in protein production, we showed that optimizing ERα codons to express it in a more proliferating

context restored activities normally observed in differentiated cells (Clusan et al., in preparation). In this study, codons optimization was based on their frequency of use in divergent biological conditions, but we next further precisely identified the third base of codons as playing this role in regulating translation and ERα activities (chapter 2). Considering these results, we aimed here to characterize the mechanisms involved in putting the regulation of translation as a function of codon usage and cell state, using the luminal (ERα-positive) breast cancer cell line, MCF7, as a model. To do so, polysome profiling experiments were performed to characterize the translation depending on MCF7 cell state. High throughput RNA sequencing was used to identify the mRNAs translated by monosomes and polysomes, and mass spectrometry determined the composition of the translational machinery in quiescent versus proliferating cells as well as after inducing MCF7 cells through an epithelial-mesenchymal transition (EMT). According to our results, a translational dynamics exists during cell cycle, favoring a polysome-mediated translation of AU3-rich mRNAs in S phase. This dynamics seems to be coordinated with gene expression for the production of key factors, and could be linked to modifications of the translational machinery composition.

2. Materials and methods

2.1. Cell culture and cell cycle synchronization

MCF7 cells were routinely maintained in DMEM (Gibco) supplemented with 8% fetal calf serum (FCS; Biowest) and antibiotics (Gibco) at 37 °C in 5% CO_2. Control and overexpressing MRTFA-ΔN200 MCF7 clones were previously described and maintained in DMEM supplemented with 8% FCS, antibiotics, and the selective antibiotics blasticidin (5 µg/mL) and zeocin (100 µg/mL) from Invitrogen (Flouriot et al., 2014). Prior to the experiments, MRTFA-ΔN200 expression was induced by a 48 h treatment of the MCF7 clones with 1 µg/mL tetracycline in DMEM supplemented with 2.5% dextran/charcoal-stripped FCS (dsFCS; Biowest), at 37 °C in 5% CO_2. To synchronize the cell cycle of MCF7 cells, cells were plated and incubated for 24 h before medium exchange for serum and steroid starvation. Samples in the G0/G1 phase were obtained after 72 h maintenance in DMEM supplemented with 2.5% dsFCS before lysis. Once in G0/G1 phase, cells were treated 48 h with 5 µg/mL aphidicolin (Sigma) and 10 nM 17β-estradiol (E2; Sigma) to provoke cell cycle entry and blockade at the G1/S transition. Cells were then washed in phosphate-buffered saline (PBS) before the addition of 10 nM E2 for 6 h (S phase) or 9 h (S/G2 transition).

2.2. Plasmids and transient transfections

The pCR-ERα-cmv_WT and MRTFA-ΔN200 pcDNA4/TO expression vectors have been previously described (Huet et al., 2009; Jehanno et al., 2020), as well as those expressing pCR-ERα-cmv_SYN 100% GC3, pCR-ERα-cmv_SYN 0% GC3, pCR-ERα-ccna2_WT, pCR-ERα-ccna2_SYN 100% GC3 and pCR-ERα-ccna2_SYN 0% GC3 (chapter 2). To assess the mRNA distribution of ERα in ribosomes depending on its GC% and promoter, 5 µg of ERα-expressing vectors were transfected in each 10-cm dish. Transfections were carried out with JetPEI® (Polyplus transfection) according to the manufacturer's instructions.

2.3. Flow cytometry analysis

Cells were plated in 10-cm dishes and incubated for 24 h before medium exchange for serum and steroid starvation. Cell cycle synchronization was performed as described above to obtain cell lysates in G0/G1, S and S/G2 phases. After treatment, cells were trypsinized, collected in PBS containing 30% IFA buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 4% FCS, 0.1% NaN_3), pelleted at 1,000 rpm for 10 min, and fixed for 30 min on ice with 70% ethanol. Fixed cells were incubated in IFA buffer containing 100 µg/mL RNase A for 15 min at 37 °C and 25 µg/mL propidium iodide was added before analysis with a FACScan equipment (Beckton-Dickinson).

2.4. Polysome profiling

Cells were plated in 10-cm dishes and incubated for 24 h before medium exchange for serum and steroid starvation. Cell cycle synchronization was performed as described above to obtain cell lysates in G0/G1, S and S/G2 phases. Cell lysis and polysome fractionation were performed as described by Guzzi et al. with some modifications (Guzzi et al., 2018). Prior to their lysis, cells were incubated 10 min at 37 °C with 50 µg/mL cycloheximide (CHX) from Sigma. Then, cells were immediately placed on ice, washed with ice-cold PBS containing 50 µg/mL CHX, and lysed in 500 µL passive lysis buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.25% NP-40, 0.1% Triton X-100, 320 U/mL RNasin® Ribonuclease Inhibitor from Promega, 150 µg/mL CHX, 20 mM DTT). Cells were incubated on ice for 40 min before being scrapped, and the resulting lysate was centrifuged for 10 min at 15,000 g, 4 °C, to recover the supernatant. The lysates from 4 dishes were pooled and concentrated by ultrafiltration with a 50 kDa molecular weight cutoff, using Amicon® Ultra centrifugal filter units from Millipore. Each lysate was loaded onto a 10-50% (w/v) sucrose gradient (25 mM Tris-HCl pH 7.4, 25 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT) and centrifuged in a SW 55 Ti rotor (Beckman) for 75 min at 45,000 rpm, 4 °C. Fractions of 100 µL were manually collected and their absorbance was measured at 260 nm to generate polysome profiles. To analyze the mRNA distribution in monosomes and polysomes, RNA was extracted from 7 samples per gradient using Direct-Zol RNA Miniprep (Zymo Research). Each sample was composed of 2 concomitant fractions selected along the polysome profile, as indicated in Supplementary Table 1. 260 µL of TRIzol was added to each 200 µL sample before following manufacturer's instructions, including a DNase I treatment when cells were transfected by ERα-coding plasmids. Finally, RNA was resuspended in 25 µL RNAse-free water.

2.5. Quantitative RT-PCR (RT-qPCR)

After polysome fractionation and RNA extraction from ribosome-containing fractions, 200 ng of RNA were reverse-transcribed in cDNA and quantitative PCR (qPCR) was performed as previously described (T. H. Pham et al., 2021), with the primer sequences indicated in Supplementary Table 2. It is to note that the same primer pair was used to detect all transfected versions of ERα (WT, SYN 100% GC3 as well as SYN 0% GC3). The sequence amplified using these primers indeed lies just upstream the coding sequence, which is a common region between each transfected plasmid and which allows discrimination from intrinsic ERα produced by MCF7 cells. mRNA detection along gradients was analyzed according to the method described by Panda et

al. (Panda et al., 2017). Briefly, the 2ΔCt method was used, where the Ct measured in the negative control sample (on top of the gradient, where no ribosome stands) was subtracted to the Ct measured in each monosome or polysome sample. This specific detection value was multiplied by a corrective factor to obtain the final mRNA value detected in the sample. This corrective factor applies because the same amount of RNA was reverse-transcribed instead of an equal volume, whereas RNA concentration varies greatly between samples. The corrective factor corresponds then to the dilution factor of the initial concentration of RNA in the RT-PCR reaction (200 ng of RNA in a final amount of 20 µL, corresponding to a final RNA concentration of 10 ng/µL). The mRNA detection values from all the samples composing a gradient were added, to obtain the total amount of mRNA detected in the gradient, from which the percentage of mRNA detected in each sample was determined. For the detection of ERα mRNAs encoded by transfected plasmids, as the primers used do not span an exon-exon junction, qPCRs were performed using no-reverse-transcribed RNA as a control for each sample. The low level of signal obtained from these control qPCRs validated the specific detection of cDNA for each transfected version of ERα.

2.6. RNA sequencing

After polysome fractionation and RNA extraction from ribosome-containing fractions, samples corresponding to the monosome peak or polysomes were pooled. As spiked-in reference, 0.01 ng of luciferase control RNA (Promega) was added to 1 µg RNA, and RNA quality was validated using an RNA Nano Chip (RNA 6000 Nano Kit, Agilent) and a 2100 Bioanalyzer (Agilent). RNAs libraries were prepared, sequenced and demultiplexed at the GeT-PlaGe core facility, INRAE US 1426 Toulouse, on a NovaSeq6000 Illumina apparatus. Reads were then filtered by quality and trimmed using Trim_Galore v.0.6.7 (https://github.com/FelixKrueger/TrimGalore). Alignment onto the human genome (hg19) was performed using bowtie2 v2.4.3 with --fast-local --fr parameters creating .sam files which were converted to .bam format using samtools v1.15.1 (Langmead and Salzberg, 2012; H. Li et al., 2009). Reads pilups were assigned to Ensembl annotation of Human genes (hg19.gtf) by using the featurecounts function integrated into the R subread package v2.10.1 (Y. Liao et al., 2014; Yang Liao et al., 2019). Spike-in normalization was done by counting sequenced reads that aligned on the luciferase cDNA using as above the bowtie2/samtools/featurecounts pipeline. These counts were processed through the BRGenomics package (https://mdeber.github.io) to calculate spike-in normalized reads per million coefficient (SRPMC) that was next used to normalize all sample counts altogether. Low-expressed mRNAs were excluded by applying the HTS filter R package onto the gene count matrix (Rau et al., 2013). To define monosome/polysome specific mRNAs, the DESeq2 R v.1.34 package that identifies differentially expressed populations of mRNAs was used (Love et al., 2014). Variations were considered as specific when exhibiting both an absolute value of fold-change (FC) >1.5 and a Benjamini-Hochberg corrected P-value calculated by a Wald test <0.05. The mRNAs that did not reach these values of significance were declared as having no preference for monosome or polysome-mediated translation.

Normalized data depicting the transcriptome of MCF7 cells at different cell cycle phases (G0/G1, S and M) were obtained from the published GSE94479 GEO dataset (Liu et al., 2017). Identical thresholds for significance (absolute FC >1.5 and adjusted P-value <0.05) were applied on these data to identify mRNAs exhibiting specific variations in expression at different cell cycle phase. Genes were defined as being "specific" for a given phase when their expression was statistically higher in this phase than in the two others.

2.7. Mass spectrometry

Cells were plated, synchronized and lysed as described for polysome profiling, without the 10-min CHX treatment and in absence of CHX in PBS and passive lysis buffer. Concentrated lysates were loaded on 10-50% (w/v) sucrose gradients and centrifuged in a SW 55 Ti rotor (Beckman) for 75 min at 45,000 rpm, 4 °C. Fractions of 100 µL were manually collected and their absorbance was measured at 260 nm to generate polysome profiles. Fractions composing the monosome peak were pooled and concentrated by ultracentrifugation. To do so, an equal volume of sucrose-free buffer (25 mM Tris-HCl pH 7.4, 25 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT) was added to the sample before a 45-min centrifugation at 75,000 rpm, 4 °C, in a TLA-120.2 rotor (Beckman). The supernatant was removed and the ribosome pellet resuspended in 20 µL of sucrose-free buffer.

Sample preparation (denaturation, reduction, alkylation) and a 3-hour tryptic digestion were performed using the iSC-BCT kit (Preomics) following manufacturer's instructions. After sample purification, 200 ng of resulting peptides were separated on a 75 µm x 250 mm IonOpticks Aurora 2 C18 column (Ion Opticks Pty Ltd, Australia). A gradient of basic reversed-phase buffers (Buffer A: 0.1% formic acid, 98% H_2O MilliQ, 2% acetonitrile; Buffer B: 0.1% formic acid, 100% acetonitrile) was run on a NanoElute HPLC System (Bruker Daltonik GmbH, Bremen, Germany) at a flow rate of 200 nL/min at 50 °C. The liquid chromatography (LC) run lasted for 40 min (2% to 13% of buffer B during 19 min; up to 19% at 26 min; up to 22% at 30 min; up to 95% at 33 min and finally 95% for 7 min to wash the column). The column was coupled online to a TimsTOF Pro (Bruker Daltonik GmbH, Bremen, Germany) with a CaptiveSpray ion source (Bruker Daltonik). The temperature of the ion transfer capillary was set at 180 °C. Ions were accumulated for 114 ms, and mobility separation was achieved by ramping the entrance potential from −160 V to −20 V within 114 ms. The acquisition of the MS and MS/MS mass spectra was done with average resolutions of 60,000 and 50,000 full width at half maximum (mass range 100-1,700 m/z), respectively. To enable the PASEF method, precursor m/z and mobility information was first derived from full scan TIMS-MS experiments (with a mass range of m/z 100-1,700). The quadrupole isolation width was set to 2 and 3 Th and, for fragmentation, the collision energies varied between 31 and 52 eV depending on precursor mass and charge. TIMS, MS operation and PASEF were controlled and synchronized using the control instrument software OtofControl 5.1 (Bruker Daltonik). LC-MS/MS data were acquired using the PASEF method with a total cycle time of 1.31 s, including 1 TIMS MS scan and 10 PASEF scans (100 ms each) which contained each an average of 12 MS/MS scans. Ion mobility resolved mass spectra, nested ion mobility vs. m/z distributions, as well as summed fragment ion

intensities were extracted from the raw data file with DataAnalysis 5.1 (Bruker Daltonik GmbH, Bremen, Germany).

Peptide and protein identification were performed using the Mascot (Mascot server v2.5.01; http://www.matrixscience.com) database search engine. MS/MS spectra were queried against the UniProtKB Homo Sapiens proteome database UP000005640 UniProtKB 2022_03 restricted to one protein sequence per gene (20,577 sequences) and a common proteomic contaminant database from the Max Planck Institute of Biochemistry, Martinsried (247 sequences). Mass tolerance for MS and MS/MS was set at 15 ppm and 0.05 Da. The enzyme selectivity was set to full trypsin with one miscleavage allowed. Protein modifications were set to fixed carbamidomethylation of cysteines, and to oxidation of methionine, acetylation of lysine, acetylation of N-terminal proteins and deamidation of glutamine and asparagine as variable modifications. Identification results from Mascot (.dat files) were imported into the Proline Studio software (Bouyssié et al., 2020). This software was then used to validate protein identification with a peptide rank = 1 and a 0.1% FDR based on the Benjamini-Hochberg procedure at the peptide spectrum matches level (Couté et al., 2020). Proteins identified with exactly the same set of peptides or with a subset of the same peptides were grouped in a Protein Set which is represented by the best-identified protein (best score) or in case of same set proteins, the one with a SwissProt accession if possible. When proteins with shared peptides were identified with other peptides not belonging to the Protein Set, different Protein Sets were created, even if there are no specific peptides (i.e., if theses peptides were also shared by other Protein Sets). Proline Studio software was also used to the spectral count comparison of the identified proteins in each sample as previously described (Simoes Eugénio et al., 2021). For each protein, a weighted spectral count is calculated, as suggested in Abacus (Fermin et al., 2011), where shared peptides are combined and weighted according to the specific spectral counts of the different Protein Sets sharing the same peptide(s). To detect significant differences between samples, a beta-binomial test was performed on these weighed spectral counts and a P-value was calculated for each Protein Set using the R package BetaBinomial 1.2 implemented in Proline Studio (T. V. Pham et al., 2010).

3. Results

3.1. Polysome enrichment occurs during the S phase of MCF7 cells

In order to study translation through polysome profiling during the cell cycle progression, the ERα-positive breast cancer cells MCF7 were starved during 72 hours and then synchronized with aphidicolin in order to produce samples enriched in cells in G0/G1 phase, in the middle of S phase or S/G2 phase transition (Figure 1A). As confirmed in Figure 1B, 85% of cells are in G0/G1 phase after 72 hours starvation. Following E2-induced cell cycle entry, cells then accumulate in G1/S transition due to the aphidicolin treatment. Six hours after aphidicolin removal, most of the cells (60%) are in S phase and then enter G2 phase three hours later. Polysome profiling allowing dissociation of monosomes (one ribosome per mRNA) from polysomes (several

ribosomes per mRNA) through a sucrose gradient, was then achieved to visualize the overall translation in the selected phases of cell cycle. The main difference in the polysomal profile between the three phases of the cell cycle analyzed is observed in the fractions of polysomes. Indeed, MCF7 cells display a low level of polysomes (fractions 32 to 50) in G0/G1 phase, while the latter significantly increases in S phase. In G2 entry, light polysomes (fractions 32 to 40) exhibit a level similar to the one observed in S phase, whilst heavy polysomes (fractions 40 to 50) decrease to levels detected in quiescent cells (Figure 1C). These results suggest changes in the translation activity during cell cycle, with an increase in polysomes-mediated translation in S phase.

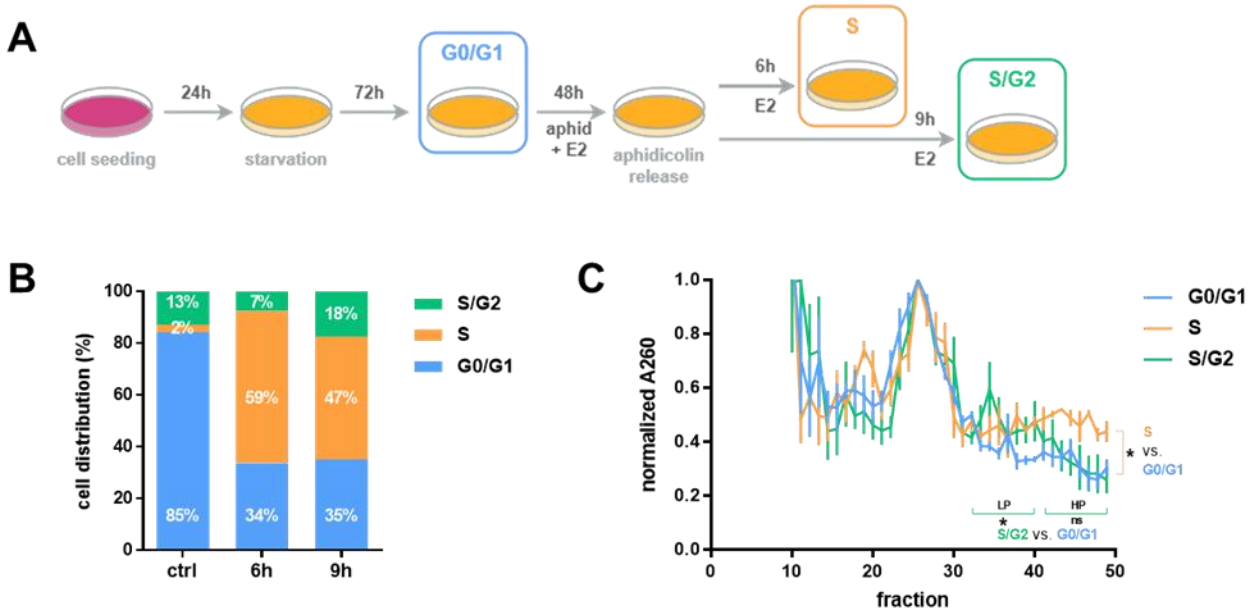
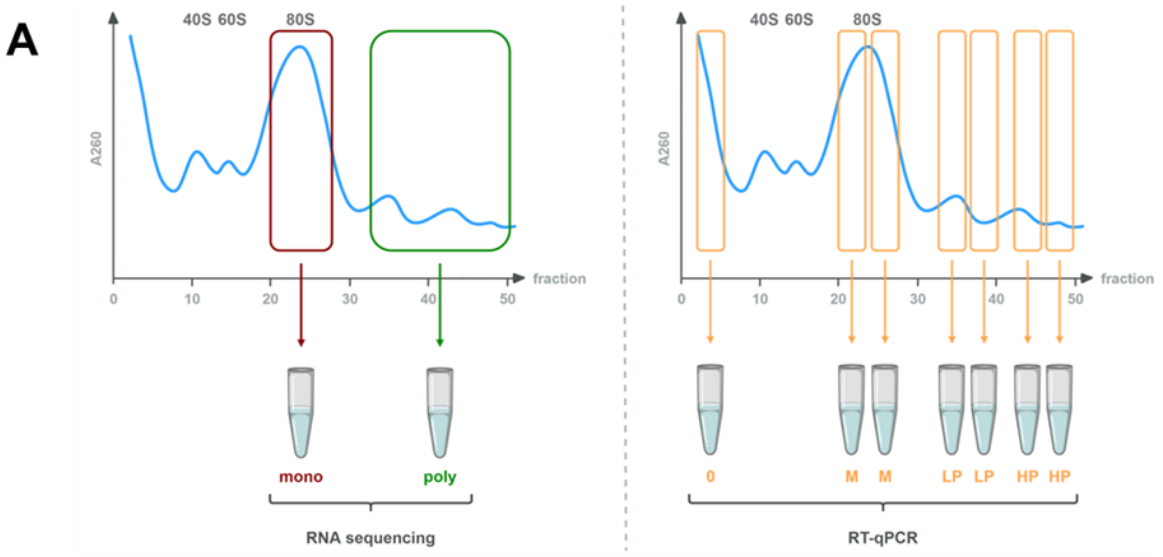
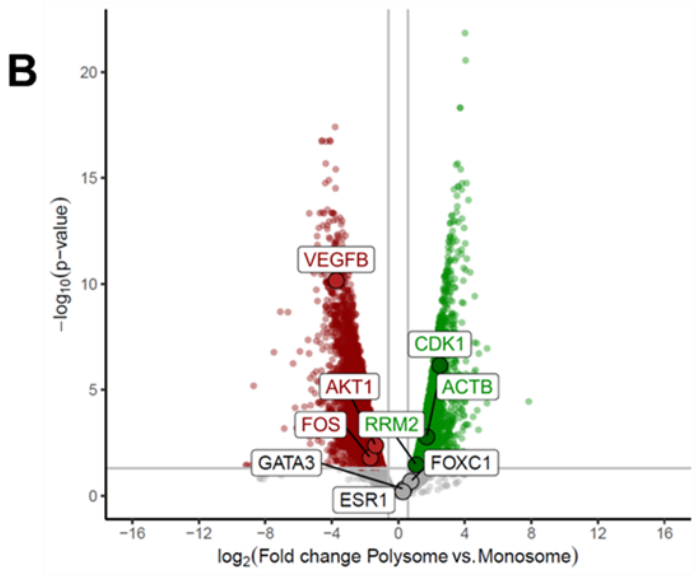
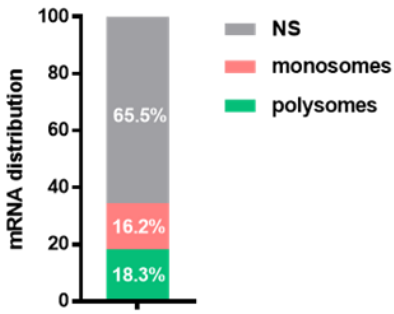


Figure 1: MCF7 cells display more polysomes in S phase than in quiescent G0/G1 phases. (**A**) Schematic overview of MCF7 cells synchronization protocol. Aphidicolin and 17β-estradiol (E2) were used at a final concentration of 5 µg/mL and 10 nM, respectively. (**B**) Cell distribution in each cell cycle phase determined by flow cytometry during the synchronization protocol. ctrl: cell sample following 72h starvation (G0/G1); 6h and 9h: time following aphidicolin release. Times 6 and 9h correspond to cells in S phase and in S/G2 transition, respectively. (**C**) Polysome profiling experiments following MCF7 cells synchronization in G0/G1, S or S/G2 phases. Data shown correspond to the mean values from triplicate experiments ± SEM, after normalizing the absorbance at 260 nm to the monosome peak. Significant differences between polysomes levels are assessed by a Student's t-test with P-value <0.05 (*).


A
40S 60S
80S
A260
fraction
0
10
20
30
40
50
mono
poly
RNA sequencing
0
M
M
LP
LP
HP
HP
RT-qPCR


B
−log10(p−value)
log2(Fold change Polysome vs. Monosome)
VEGFB
AKT1
FOS
GATA3
ESR1
CDK1
ACTB
RRM2
FOXC1


C
mRNA distribution
65.5%
16.2%
18.3%
NS
monosomes
polysomes

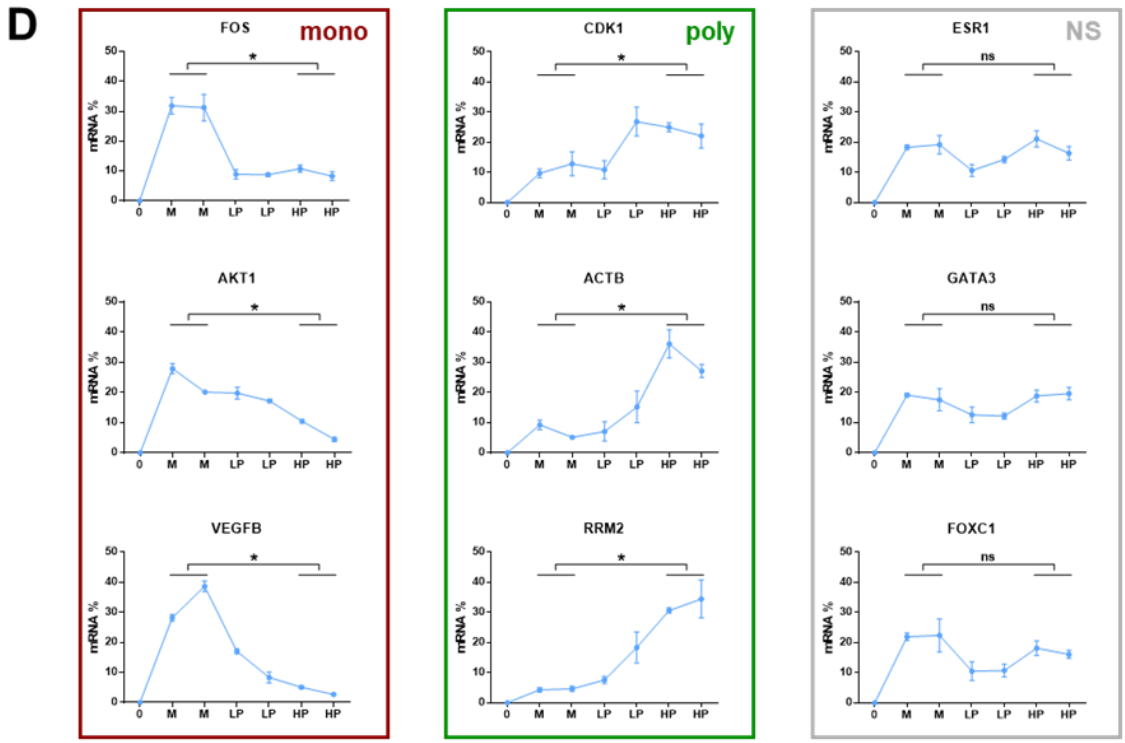

D
FOS
mono
AKT1
VEGFB
CDK1
poly
ACTB
RRM2
ESR1
NS
GATA3
FOXC1
mRNA %
0
M
M
LP
LP
HP
HP
*
ns

Figure 2: Monosomes and polysomes partly translate different sets of mRNAs. (**A**) Schematic overview of the preparation of the samples subjected to RNA sequencing or RT-qPCR following polysome profiling. (**B**) Volcano plot of the mRNAs detected by RNA sequencing following polysome profiling in quiescent G0/G1 MCF7 cells. Reads from duplicate experiments were normalized and analyzed as described in the materials and methods section. mRNAs whose the detection was higher in monosomes are colored in red, in polysomes in green, or without differential detection in gray. (**C**) Proportion of mRNAs that are more detected in monosomes, in polysomes or without differential detection (NS) (n=11,887; P-value <0.05; fold-change >1.5). (**D**) Relative expression level of different mRNAs detected by RT-qPCR along the sucrose gradient following polysome profiling from quiescent G0/G1 MCF7 cells. In red, green and grey squares are examples of mRNAs which were more detected in monosomes, in polysomes or without differential detection (NS) by the high throughput RNA sequencing approach, respectively. Data correspond to the mean values from triplicate experiments ± SEM. Significant differences are assessed by a Student's t-test with P-value <0.05. 0: first sample containing no ribosome; M: monosomes; LP: light polysomes; HP: heavy polysomes.

3.2. Monosomes and polysomes partly translate different sets of mRNAs

Considering the dynamics of polysomes during cell cycle, the existence of mRNAs specifically translated by either monosomes or polysomes were investigated to identify a potential divergence between ribosomal populations. To do so, polysome profiling was performed and fractions composing the monosome peak were pooled whereas those constituting polysomes were combined before subjecting both samples to RNA sequencing (Figure 2A, left panel). This analysis was first performed in quiescent MCF7 cells only, before comparing with other cell cycle phases. mRNAs were considered differentially translated between monosomes and polysomes when P-value was <0.05 and fold-change >1.5. As shown in Figures 2B and 2C, near 12,000 ribosome-bound mRNAs were detected, among which 16% were enriched in monosomes and 18% in the polysomes-containing sample. Therefore, 35% of total MCF7 cells mRNAs were preferentially translated by either monosomes or polysomes. The remainder (65.5%) were indifferently present in both monosome and polysome fractions. RT-qPCRs were achieved to validate the enrichment or not of some mRNAs in one ribosomal population. In this case, two fractions were successively pooled along the sucrose gradient following polysome profiling, in order to visualize the translational profile of these genes more accurately (Figure 2A, right panel). The results obtained in these experiments validated the specific mRNA enrichments in monosomes (FOS, AKT1, VEGFB) or polysomes (CDK1, ACTB, RRM2). Similarly, when no particular enrichment was determined by RNA sequencing, the translational profiles obtained by RT-qPCR evidenced an equivalent translation by monosomes and polysomes (ESR1, GATA3, FOXC1) (Figure 2D). To investigate whether preferential mRNA translations between monosomes and polysomes are altered during cell cycle, the same set of mRNAs was analyzed by RT-qPCR following polysome profiling from cells synchronized in G0/G1, S phase and S/G2 phase transition. Surprisingly, the distribution between monosomes and polysomes for the same mRNA remained similar regardless the cell cycle phase analyzed

(Supplementary Figure S1). These results could suggest that the parameters driving mRNA translation towards monosomes or polysomes were robust enough to result in the same translational profiles. On the other hand, the obtained mRNA expression profile could also be the result of the remaining unsynchronized cell population (30-40%), since cell synchronization by aphidicolin is not complete. In light of these results, high throughput RNA sequencing was not performed in the other phases of the cell cycle.

3.3. Polysomes translate more specifically mRNAs expressed during cell division

As previously mentioned, 35% of mRNAs are preferentially translated by either monosomes or polysomes. Functional annotation of these mRNAs specifically enriched in one ribosomal population by Metascape (https://metascape.org) evinced divergent functional enrichments (Figure 3). Notably, mRNAs translated in monosomes were found to encode ribosomal proteins, translation elongation factors, as well as proteins involved in cellular respiration (COX, NADH:ubiquinone oxidoreductase, or ATP synthase subunits). In contrast, mRNAs preferentially translated by polysomes encode proteins preferentially linked to RNA metabolism (ADAR, DEAD box proteins, heterogeneous nuclear ribonucleoproteins…) as well as cell cycle and mitosis (cyclins, CDK1, PCNA…). This observation suggests that a preferential translation by monosomes or polysomes convey a functional meaning with, in particular, an enrichment in polysomes of mRNAs encoding proteins involved in cell cycle progression, S phase and mitosis.

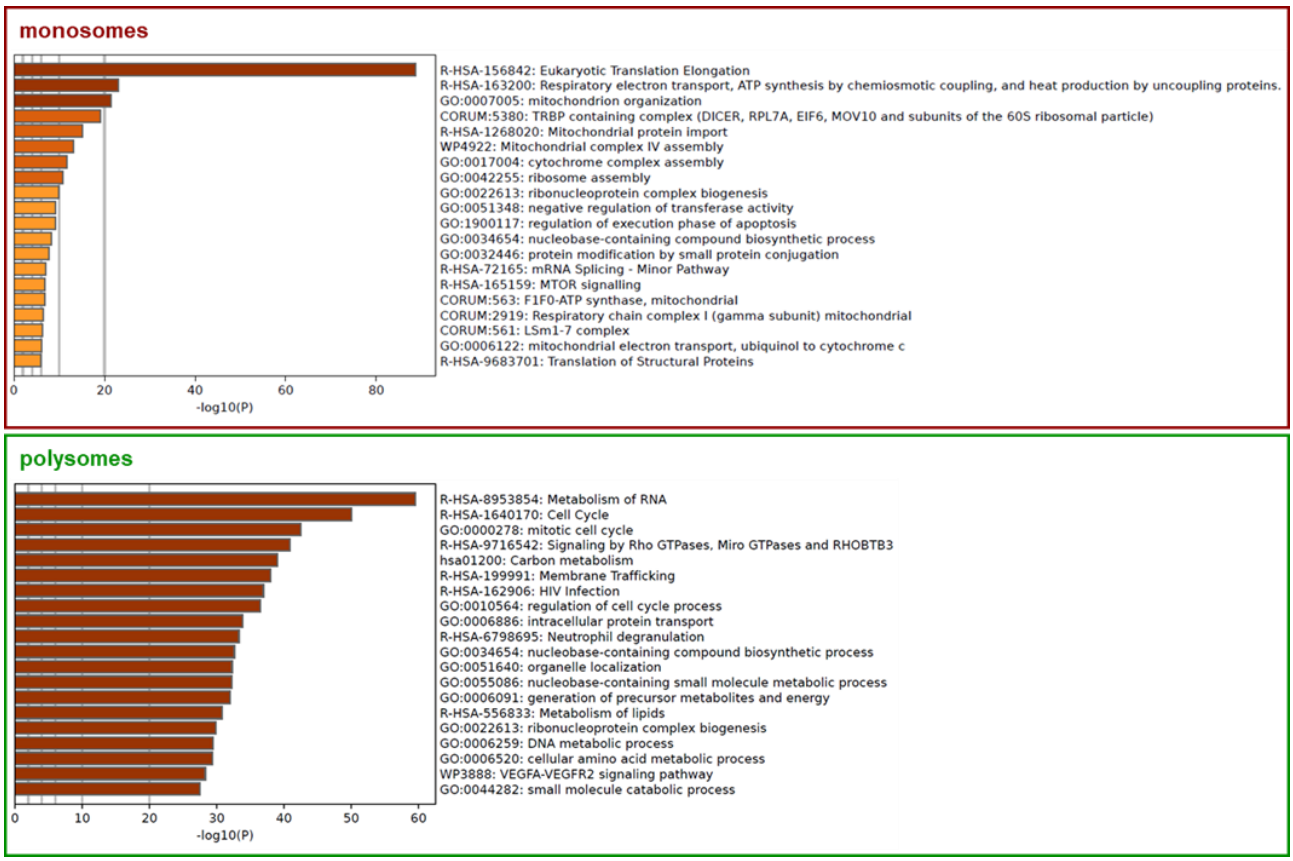


Figure 3: mRNAs preferentially translated by monosomes or polysomes are involved in different biological processes. Functional enrichment analysis of mRNAs preferentially translated by monosomes or polysomes. Analysis realized with the Metascape resource (http://metascape.org) (Zhou et al., 2019). Color scale represents statistical significance.

In order to gain further insights into a potential link between monosome/polysome-mediated translation and cell cycle, we compared the list of mRNAs identified from our polysome profiling study to transcriptomic data obtained by Liu et al., who identified differentially expressed genes between the main phases of MCF7 cell cycle (Liu et al., 2017). In this study, cells were harvested in G0/G1 phase following starvation, whereas the synchronization in S (S phase entry) and M phases was performed with a double thymidine block or with a thymidine-nocodazole treatment, respectively. From their lists of genes differentially expressed between G0/S, G0/M and S/M phases, we established new lists of genes this time specifically expressed in G0, S and M phases by selecting common and conserved genes between each pairs of conditions analysis. Importantly, 65 to 74% of these cell-phase specific mRNAs identified in the transcriptome from Liu et al. were detected in our post-polysome profiling data (Figure 4A), enabling the comparison of both datasets. Amongst the mRNAs common to both datasets, 30 to 40% were preferentially translated by one of the ribosomal populations. As shown in Figure 4B, the percentage of mRNAs specifically enriched in monosomes was greater in quiescent cells and decreases in S and M phases in favor of mRNAs detected in both monosomes and polysomes. In parallel, the proportion of mRNAs specifically enriched in polysomes remains relatively unchanged between the three phases of the cell cycle. These results together indicate that more mRNAs are translated at the polysome level during S and M phases.

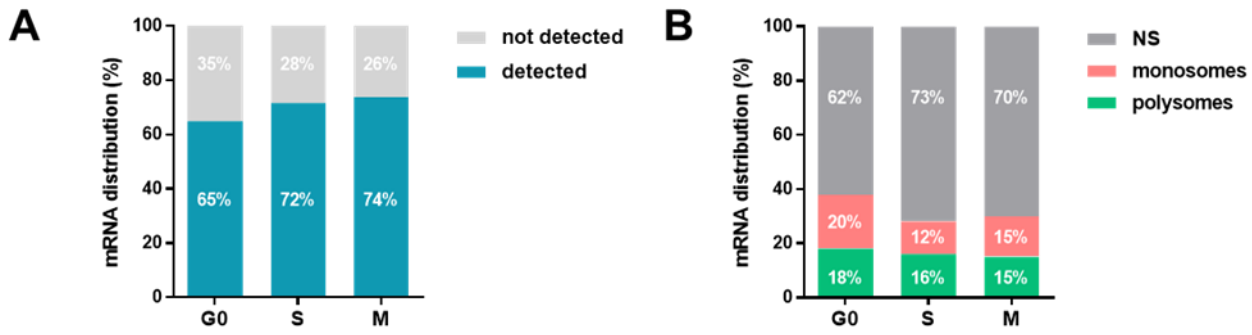


Figure 4: mRNAs specifically expressed during cell cycle are less translated by monosomes. (**A**) Proportion of mRNAs specifically expressed in G0/G1, S or M phase from the study of Liu et al. (Liu et al., 2017), that are detected or not by RNA sequencing following the polysome profiling performed in our MCF7 cells. (**B**) Proportion of mRNAs specifically expressed in G0/G1, S or M phase, that are more detected in monosomes, polysomes or without differential detection (NS) (P-value <0.05; fold-change >1.5).

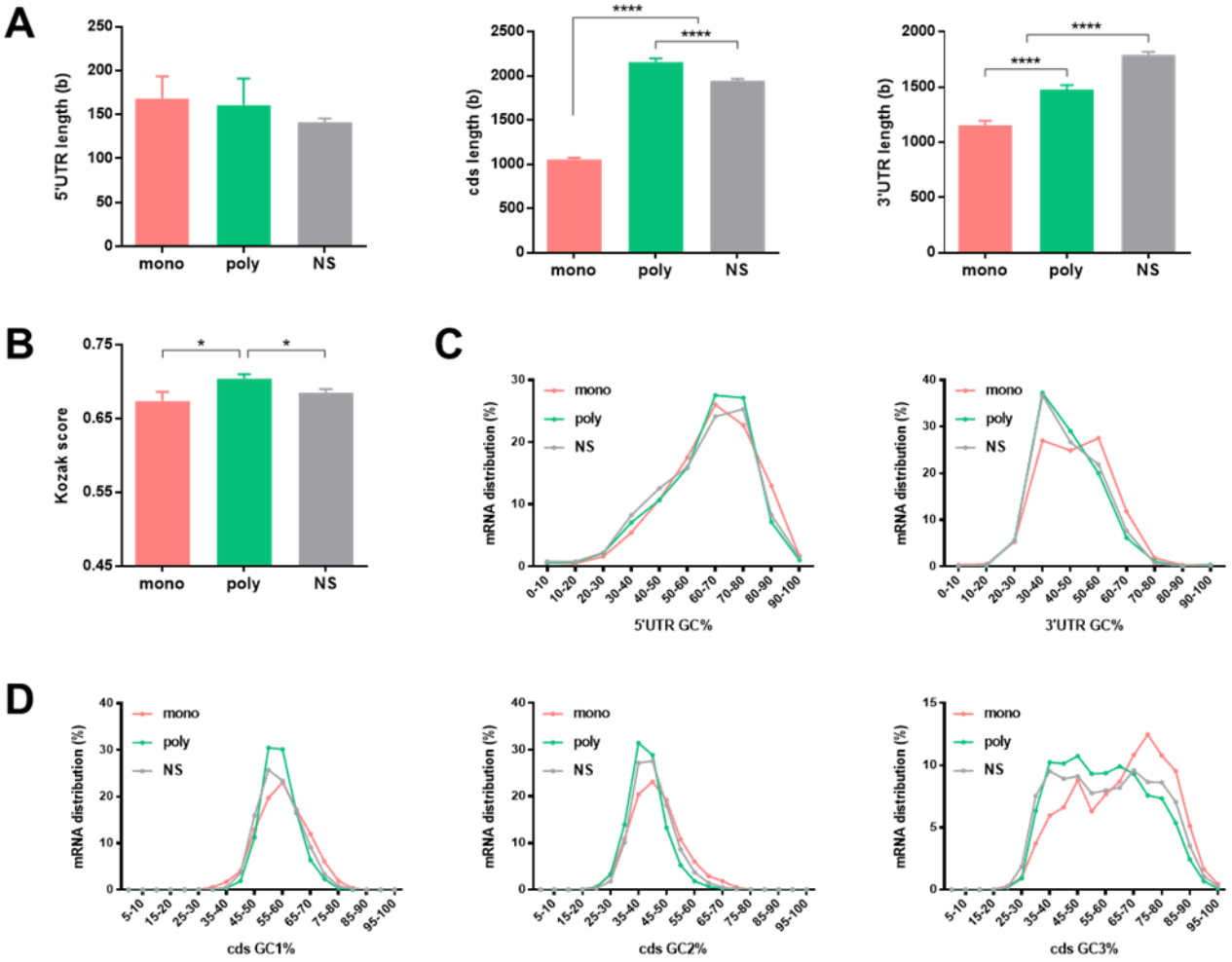


Figure 5: mRNAs preferentially translated by monosomes or polysomes present distinct intrinsic features. (**A**) Mean length (± SEM) of the 5'UTR, coding sequence (cds) and 3'UTR of mRNAs more detected in monosomes, polysomes or without differential detection (NS) by RNA sequencing following polysome profiling in MCF7 cells. Significant differences are assessed by a Student's t-test with p-value <0.0001 (****). (**B**) Mean Kozak score (± SEM) reflecting the similarity of mRNA Kozak sequence with the consensus one, depending on their preferential detection in monosomes, polysomes, or without differential detection (NS). Significant differences are assessed by a Student's t-test with P-value <0.05 (*). (**C**) Proportion of mRNA that are more detected in monosomes, polysomes or without differential detection (NS) depending on the global GC% of their 5'UTR or 3'UTR. (**D**) Proportion of mRNA that are more detected in monosomes, polysomes or without differential detection (NS) depending on the GC% of the first (GC1), second (GC2) or third (GC3) base of the codons in their coding sequence (cds).

3.4. mRNAs preferentially translated by monosomes or polysomes present distinct features including a bias in the GC3 enrichment

mRNAs preferentially translated by monosomes, polysomes or both were then investigated for intrinsic mRNAs parameters such as 5'UTR, 3'UTR and coding sequence length, Kozak sequence consensus level (Kozak score) or GC enrichment of the 5'UTR, 3'UTR and coding sequence. The coding sequences of mRNAs preferentially translated by monosomes were found significantly shorter than those of mRNAs enriched in polysomes or having no specific translation preference for monosomes/polysomes (Figure 5A, middle panel). Similar but less pronounced dynamics were also observed for the 3'UTR length, while no major difference was detected for the 5'UTR length (Figure 5A, right and left panels, respectively). Concerning the Kozak sequence, messengers enriched in polysomes display a higher correspondence to the consensus than mRNAs enriched in monosomes or with no preferential translation (Figure 5B). Finally, coding sequences of mRNAs that are more detected in monosomes are enriched in GC3, when those more detected in polysomes are slightly enriched in AU3 (Figure 5D, right panel). This bias in codon composition is specific to the third codon base, as no difference in the distribution of the three mRNA categories was found depending on their GC1 or GC2% (Figure 5D, left and middle panels). A minor enrichment in global GC% was also found in the 3'UTR region of mRNAs preferentially translated by monosomes, whereas no difference was observed for their 5'UTR sequences (Figure 5C, right and left panels, respectively). It should be noted that 5'UTR sequences were globally more GC-rich than 3'UTR. Altogether, these results indicate that mRNAs preferentially translated by monosomes or polysomes present distinct features, with a particular enrichment linked to codon composition and transcript length. Schematically, long, AU3-rich mRNAs are more translated by polysomes, while shorter, GC3-rich mRNAs are more translated by monosomes.

3.5. Preferential translation by monosomes or polysomes is linked to mRNA GC3 enrichment and promoter activity

To ascertain that a preferential translation by monosomes or polysomes can be linked to mRNA GC3%, the translational profile of synonymous versions of ERα was investigated. To do so, three ERα-coding sequences were used: the one of the wild-type ERα (ERα WT) and coding sequences in which all codons of the wild-type sequence were replaced by codons ending with G or C (ERα SYN 100% GC3), or with A or T (ERα SYN 0% GC3) (chapter 2). Each cDNA produces the same amino acid sequence corresponding to the wild-type estrogen receptor. Plasmids coding these synonymous versions of ERα were transfected in MCF7 cells before performing polysome profiling and analyzing ERα mRNA distribution along the sucrose gradient by RT-qPCR. Surprisingly, the translational profile of transiently transfected ERα WT differs from the one of endogenous ERα, with a clear enrichment in monosomes, whereas the same coding sequence is translated (Figure 6A, blue and pink curves, respectively). Expressing ERα with codons ending only with GC3 (ERα SYN 100% GC3) or AT3 (ERα SYN 0% GC3) led to the same translational profiles as transiently transfected ERα WT, suggesting that GC3% only is not sufficient to direct translation towards monosomes or polysomes (Figure 6A). Those mRNAs were expressed under the control of a CMV promoter. Yet, we have previously shown that translation

of AU3-rich ERα mRNAs was more efficient by using a promoter specifically active in the middle of the cell cycle and particularly in S phase (chapter 2). Therefore, the translational profile of ERα synonymous variants was also investigated after replacing the CMV promoter by the cyclin A2 (CCNA2) promoter known to be active in S phase (Henglein et al., 1994). Whereas the translational profile of ERα WT and ERα SYN 100% GC3 mRNAs was not modified in these new conditions, translation of ERα SYN 0% GC3 mRNAs under the transcriptional control of cyclin A2 promoter became more driven by polysomes than mRNAs under the control of the CMV promoter (Figure 6B). These results might suggest a close association between the promoter activity in cell cycle, the GC3 enrichment of mRNAs and a monosome/polysome-mediated translation.

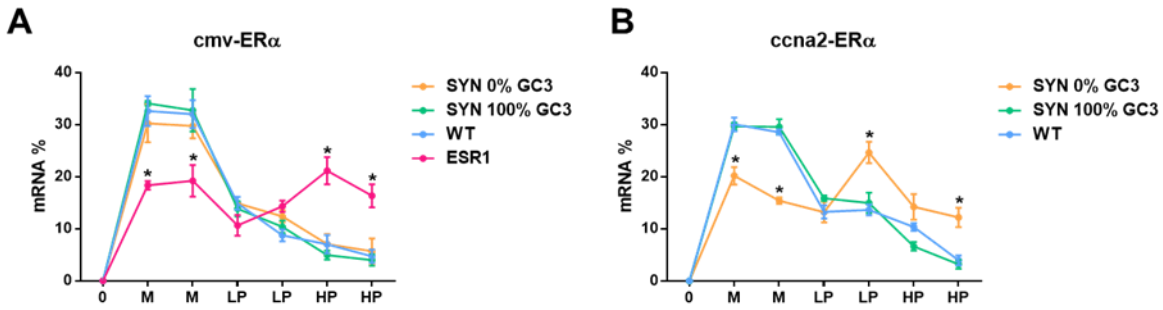


Figure 6: Preferential translation by monosomes or polysomes is linked to mRNA GC3 enrichment and promoter activity. (**A**) Proportion of ERα mRNA detected by RT-qPCR along the sucrose gradient following polysome profiling from MCF7 cells transiently transfected to express ERα WT (blue line), ERα SYN 100% GC3 (green line) or ERα SYN 0% GC3 (orange line) under a CMV promoter. The detection of endogenous ERα (ESR1) is indicated in pink, with data corresponding to the mean values from all transfected conditions ± SEM. (**B**) Proportion of ERα mRNA detected by RT-qPCR along the sucrose gradient following polysome profiling from MCF7 cells transiently transfected to express ERα WT (blue line), ERα SYN 100% GC3 (green line) or ERα SYN 0% GC3 (orange line) under the cyclin A2 (CCNA2) promoter. Data correspond to the mean values from triplicate or quadruplicate experiments ± SEM. Significant differences are assessed by a Student's t-test with P-value <0.05 (*). 0: first sample containing no ribosome; M: monosomes; LP: light polysomes; HP: heavy polysomes.

3.6. The composition of the translational machinery differs according to the cell state of MCF7 cells

Giving our previous results, MCF7 cells in S phase exhibit more polysomes, which in part translate a different set of mRNAs than those translated by monosomes. Moreover, this preference appears to be partly correlated with GC3 enrichment and length of mRNA coding sequences. These observations suggest an evolution of the translational regulation during cell cycle that would be required to produce specific proteins at a given time. To determine whether this translation change relies on variations in the composition of the translational machinery, mass spectrometry was performed on ribosomal purification from MCF7 cells synchronized in G0/G1 or S phases. In parallel, this approach was also applied on MCF7 stably expressing the constitutively active form of MRTFA (MRTFA-ΔN200). Indeed, MRTFA is a master regulator of the epithelial-

mesenchymal transition (EMT) whose activity is regulated by the dynamics of the actin cytoskeleton (Medjkane et al., 2009; Miralles et al., 2003). Deletion of its interaction with actin (ΔN200) renders it constitutively active, causing MCF7 cells to shift towards a "basal-like" phenotype, increasing cell proliferation and invasion. Importantly, functional properties of ERα also change to gain similarities to those it exerts in MCF7 cells in S phase (unpublished data) (Jehanno et al., 2020). Furthermore, as recently reported, the MRTFA-ΔN200 MCF7 cell line displays high polysome content (Fernández-Calero et al., 2020). The use of this cell model in addition to MCF7 cells in S phase aims to better identify potential modifications in the translational machinery between quiescent and proliferating cells. Mass spectrometry was performed after total ribosome purification, with no dissociation of monosomes from polysomes, to generate a global picture of the translational machinery in various cell states. The use of cycloheximide was then avoided because this translation elongation inhibitor could modify the interactions between ribosomes and associated proteins (Shen et al., 2021). As shown in Figure 7A and in agreement with previous studies (Fernández-Calero et al., 2020), MRTFA-ΔN200-overexpressing MCF7 cells display high polysome content, even higher than that observed in MCF7 cells in S phase. Concerning proteomic analysis, a total of 4,000 proteins were detected by mass spectrometry in overall samples. When comparing the composition of the translational machinery between the different conditions, very few differences in quality and quantity were observed between the G0/G1 and S phases of MCF7 cells. In contrast, many proteins were differentially detected between quiescent and MRTFA-ΔN200-overexpressing MCF7 cells, certainly due to the striking phenotype induced by MRTFA-ΔN200 protein (Figure 7B). For that reason, we decided first to focus our analysis on this last and main divergence in order to identify proteins that could play a major role in the translational regulation of mRNAs according to their GC3 enrichment and cell state. Then, we extended this analysis to the comparison between G0/G1 and S phases of MCF7 cells in order to identify common dynamics in terms of fold change.

Among the 80 ribosomal proteins (RPs), we identified with high coverage all of the 33 proteins of the small subunit, and only 4 RPs were missing for the large subunit (eL39, eL40, eL41 and P1/P2). Paralogs were also identified for both subunits: eS27L and eL22L, uL24L, eL42L, uL30L (Supplementary Table 3). Interestingly, one RP, eS26, was highly enriched in MRTFA-ΔN200-overexpressing MCF7 cells as compared to quiescent G0/G1 cells (Figure 7C). When comparing G0/G1 and S phases of MCF7 cells, it appeared that eS26 tends also to be enriched in S phase (Supplementary Figure S2A). Besides RPs, we also identified extra-ribosomal proteins, of which around 100 were significantly more detected in MRTFA-ΔN200-overexpressing MCF7 cells as compared to quiescent cells, whilst 400 were enriched in MCF7 at G0/G1 phase (Figure 7B and 7D). Similar tendency was observed in the G0/G1 vs. S phase comparison, with the proteins enriched in MRTFA-ΔN200-overexpressing MCF7 cells being more detected in MCF7 cells in S than in G0/G1 phase (Supplementary Figure S2C). Although the discrepancies between G0/G1 and S phases are generally not significant at P-value <0.05, certainly due to a less pronounced phenotype for S phase cells than cells expressing MRTFA-ΔN200, this tendency might suggest a similar mechanism of translation in S phase and in MRTFA-ΔN200 MCF7 cells, which differs from the translation in quiescent MCF7 cells.

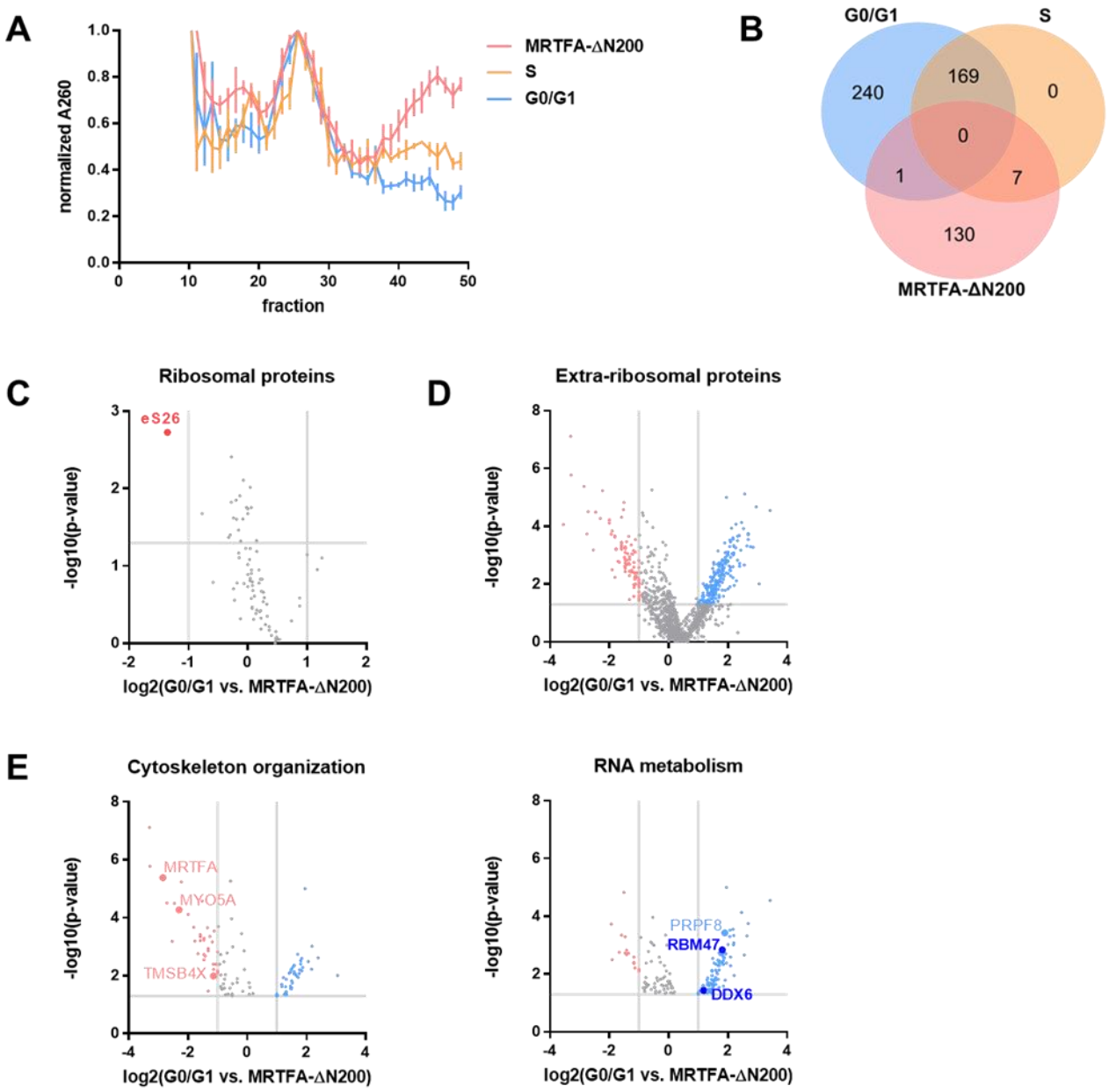


Figure 7: The composition of the translational machinery differs according to the cell state of MCF7 cells. (**A**) Polysome profiling experiment performed from MCF7 cells synchronized in G0/G1 and S phases as well as from MCF7 stably expressing MRTFA-ΔN200. Data correspond to the mean values from triplicate experiments ± SEM, after normalizing the absorbance at 260 nm to the monosome peak. (**B**) Venn diagram of the proteins significantly detected by mass spectrometry following ribosome purification in the three indicated MCF7 samples. Peptides from quadruplicate experiments were analyzed and quantified as described in the materials and methods section (P-value <0.05; fold-change >2). (**C**) Volcano plot of the ribosomal proteins (RPs) detected by mass spectrometry following ribosome purification in quiescent MCF7 cells (G0/G1) or after tetracycline induction of MRTFA-ΔN200. RPs whose detection was higher in MCF7 overexpressing MRTFA-ΔN200 are colored in red, or without differential detection in gray. (**D**) Volcano plot of the extra-ribosomal proteins detected by mass spectrometry following ribosome purification in quiescent MCF7 cells (G0/G1) or after tetracycline induction of MRTFA-ΔN200. Proteins whose detection was higher in MCF7 overexpressing MRTFA-ΔN200 appear in red, in quiescent MCF7 in blue, or without differential

detection in gray. (**E**) Volcano plot of the proteins involved in cytoskeleton organization or RNA metabolism detected by mass spectrometry following ribosome purification in quiescent MCF7 cells (G0/G1) or after tetracycline induction of MRTFA-ΔN200. Proteins whose detection was higher in MCF7 overexpressing MRTFA-ΔN200 are colored in red, in quiescent MCF7 in blue, or without differential detection in gray. Darker colors indicate proteins investigated in this study.

Enriched proteins in MRTFA-ΔN200-expressing cells were found notably to be involved in cytoskeleton organization, whereas proteins more detected in quiescent cells are linked to RNA metabolism (Figure 7E).

Because of the association observed before between GC3 enrichment in the coding sequence of mRNAs and their preferential translation by either monosomes or polysomes, we looked for extra-ribosomal proteins that would have a role in codon recognition in our proteomic analysis. Interestingly, within the proteins associated with RNA metabolism we found RNA-helicase DDX6, a protein known to be implicated in mRNA decay and translation inhibition (Minshall et al., 2009), to be more detected in quiescent G0/G1 MCF7 vs. MRTFA-ΔN200-expressing MCF7 cells (Figure 7E, right panel). According to the study of Courel et al., DDX6 is notably involved in P-bodies formation that hinders AU-rich mRNAs translation (Courel et al., 2019). Re-analysis of the data from this study on HEK293 cell transcriptomes obtained before and after P-bodies purification confirmed that mRNAs stored in P-bodies are indeed highly AU3-rich, contrary to the transcripts that are depleted from P-bodies (Figure 8A) (Hubstenberger et al., 2017). The comparison of these data with our post-polysome profiling analysis revealed that around 40% of mRNAs outside P-bodies are preferentially translated by monosomes (24%) or polysomes (19%) in MCF7 cells. In contrast, while around 19% of mRNAs detected in P-bodies are still preferentially translated by polysomes, only 7% of the transcripts are enriched in monosomes from MCF7 cells (Figure 8B). DDX6 could thus participate to the storage and translational repression of AU3-rich mRNAs whose translation is preferentially mediated by polysomes.

We observed that the Metascape resource used for annotating our data was not totally complete, as proteins interacting with RNA were not annotated as involved in RNA metabolism. For instance, a keyword-based search for “RNA-binding” on our sets of identified proteins retrieved an additional protein, RBM47, which belongs to the family of RNA-binding motif protein, whose members are involved in many RNA metabolism processes (Z. Li et al., 2021). Being absent from the RNA metabolism Volcano Plot (Figure 7E, right panel), we added RBM47 to this graph; to better visualize its differential expression between samples. The study of Vanharanta et al. identified around 2,500 transcripts bound by RBM47 (Vanharanta et al., 2014). The analysis of the GC3 composition of these transcripts revealed that they are highly AU3-rich (Figure 9A, red line). Vanharanta et al.’s study also identified 194 transcripts whose expression was modified depending on RBM47 expression, and it appears that these mRNAs are more GC3-rich than the pool of RBM47-bound mRNAs (Figure 9A, blue line). When segregating those who are up- or down-regulated by RBM47, no difference was also observed (Figure 9B). RBM47 could thus be involved in the translational regulation of mRNAs according to their GC3 bias and cell state.

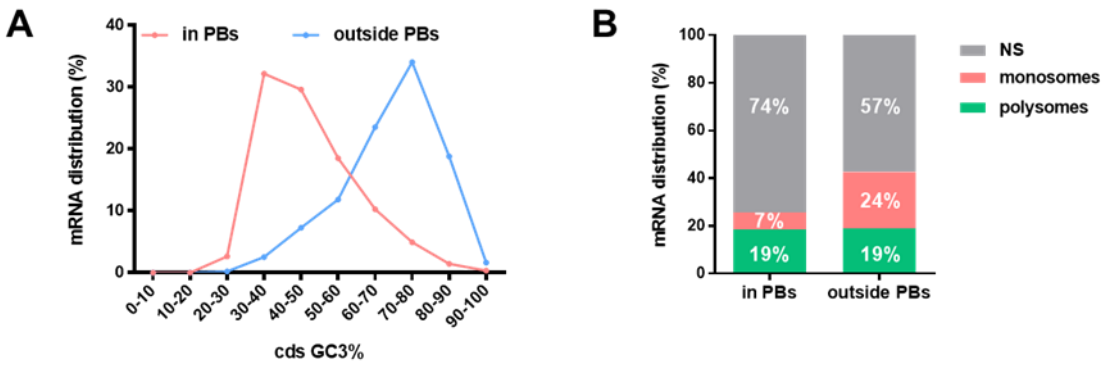


Figure 8: mRNAs stored in P-bodies are AU3-rich and more translated by polysomes. (**A**) Proportion of mRNA that are enriched (red line) or depleted (blue line) from P-bodies (PBs) depending on the percentage in GC-ending codons (GC3%) of their coding sequence (cds). Performed in HEK293 cells, data were obtained from the study of Hubstenberger et al. (Hubstenberger et al., 2017). (**B**) Proportion of mRNAs enriched or depleted from PBs that are more detected in monosomes, polysomes or without differential detection (NS) in our RNA sequencing analysis performed from polysome profiling in MCF7 cells.

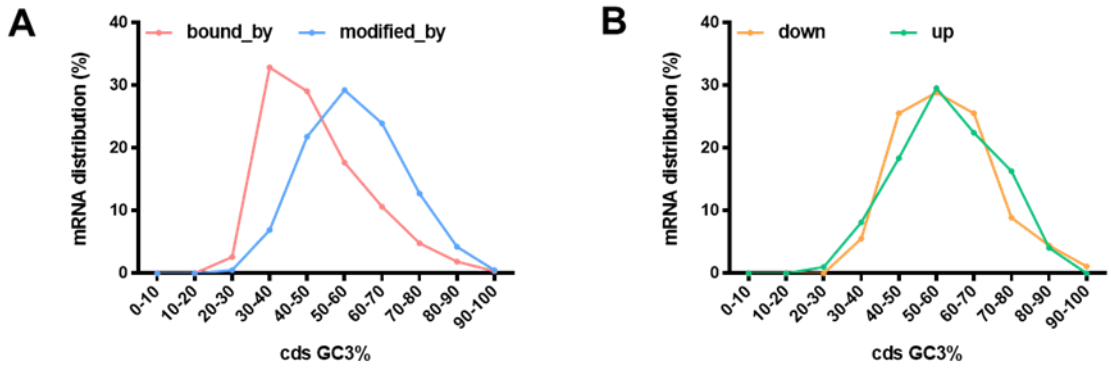


Figure 9: RBM47 binds AU3-rich mRNAs. (**A**) Percentage in GC-ending codons (GC3%) of the coding sequence (cds) of mRNAs identified in the study of Vanharanta et al. as being bound by RBM47 (red line) or affected by RBM47 levels of expression (blue line) (Vanharanta et al., 2014). (**B**) Proportion of mRNA that are up- or down-regulated by RBM47 depending on the percentage in GC-ending codons (GC3%) of their coding sequence (cds).

In conclusion, the mass spectrometry analysis of ribosome purified from MCF7 cells in G0/G1 or S phases, or after inducing the expression of MRTFA-ΔN200 in an MCF7 stable cell line, identified variations in the composition of the translational machinery between quiescent and proliferating breast cancer cells. These variations concern the ribosome composition, with a differential detection of eS26, as well as RNA-binding proteins such as DDX6 and RBM47.

Discussion

Our results altogether evidence the existence of a dynamic in polysome-mediated translation during cell cycle. Such a link between cell cycle and translation being more mediated by polysomes was early suggested in 1975 by Eremenko and Volpe, who demonstrated a modification of the polysomal pattern during the cell cycle of HeLa cells (Eremenko and Volpe, 1975). This dynamics of global translation appears to be less important in our model, probably because MCF7 cells are more differentiated than HeLa cells. MCF7 cells are indeed differentiated cancer cells that do not excessively proliferate, which suggests a moderated translational activity that would explain the low level of polysomes relating to monosomes in our experiments. An increase of monosomes to polysomes ratio was in fact demonstrated during the differentiation of murine erythroid cells (Papagiannopoulos et al., 2022). These results suggest then the existence of a regulatory mechanism leading to the modification of the translational machinery during cell cycle, resulting in an increase in polysomes-mediated translation during S phase. Among the mRNAs translated in MCF7 cells, we identified a pool of transcripts preferentially translated by monosomes or polysomes. These preferential translations are modified during cell cycle, with less mRNAs specifically translated by monosomes than polysomes in S phase. As more polysomes are notably detected in S phase, the mRNAs preferentially translated by polysomes could then result in a higher protein production in S phase, suggesting that this polysome-enrichment could play a regulatory role regarding the rate of specific protein production. Intrinsic mRNAs parameters such as the length and the Kozak sequence appear to play a role in these preferential translations, with shorter mRNAs being enriched in monosomes whereas a more optimal Kozak sequence is associated with a polysome-mediated translation. The combination of these parameters is in agreement with previous studies suggesting that shorter mRNAs enable the binding of less ribosomes than longer mRNAs, whereas displaying an optimal Kozak sequence would lead to an increased translation initiation rate, resulting in the binding of additional ribosomes before the completion of a round of translation. The regulation of translation speed is however more complex than proposed by these models, depending on the equilibrium between translation initiation and elongation rates (Arava et al., 2003; Biever et al., 2020; Heyer and Moore, 2016). Whereas the association of mRNA length and translation initiation rate with a preferential translation by monosomes or polysomes is not entirely new, we identified a role played by the GC3% of mRNAs in their translational pattern. GC3-rich mRNAs were indeed more translated by monosomes, whereas AU3-rich transcripts were enriched in polysomes. This observation is perfectly in line with our previous results highlighting the fact that mRNAs specifically expressed during cell cycle or

transformation are AU3-rich (chapter 2). The transfection of synonymous versions of ERα further enabled us to ascertain that GC3-rich mRNAs are preferentially translated by monosomes, whereas the translation of the AU3-rich version of ERα tends towards polysomes. This latter observation was however performed when expressing these cDNAs with a promoter specific to the S phase, suggesting that the combination of an adequate promoter with a particular GC3% is necessary to alter the translational profile of an mRNA. In case of ERα SYN 0% GC3, this polysome-mediated translation enabled by using a cyclin A2 promoter resulted in functional differences compared to expressing this cDNA with a CMV promoter (chapter 2). According to this model, an AU3-rich mRNA expressed in quiescent cells is translated by monosomes and leads to an impaired protein expression, whereas expressing the same mRNA in S phase favors its translation by polysomes and improves protein production. All these results highlight the existence of a regulatory mechanism enabling the coordination of mRNA transcription and translation throughout cell cycle, depending on the nature of the third codon base.

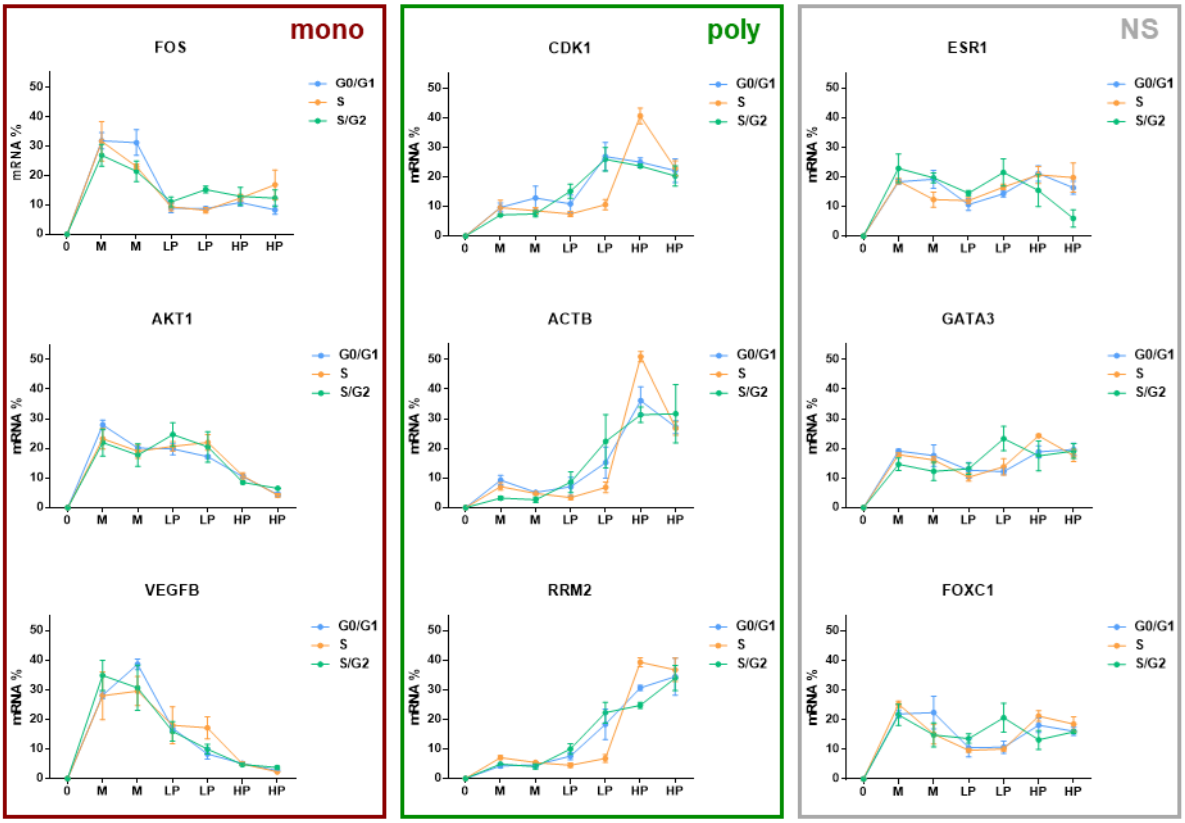
In light of these results, we hypothesized that the composition of the translational machinery could be modified during cell cycle. Performing mass spectrometry in quiescent versus proliferating breast cancer cells enabled the identification of differentially expressed proteins between conditions. First of all, we were interested in the composition of ribosomes. Only four RPs of the large subunit were missing in our analysis, but their absence can be explained by their localization at the ribosome surface, which increases the risk to be lost during sample preparation (notably eL40 and P1/P2), or their low molecular weight, that could results in peptides not detectable by mass spectrometry (especially eL39 and eL41), as it was reported in previous studies (Khatter et al., 2014; Papagiannopoulos et al., 2022). One RP in the small subunit, eS26, was differentially detected between conditions, with an enrichment in more proliferating MCF7 cells, which was also assessed at the transcriptional level in a previous study (Fernández-Calero et al., 2020). This RP is located at the exit of the mRNA-binding channel and would be the last to be incorporated in the 40S subunit of S.cerevisiae ribosomes, maybe already in polysomes. This could explain that this protein is more detected in our samples enriched in polysomes (Ben-Shem et al., 2011; Strunk et al., 2012). In agreement with this hypothesis, eS26 is one of the RPs exhibiting a dynamic association with mature ribosomes in a study investigating the dynamic exchange of RPs in cultured rat neurons (Fusco et al., 2021). This dynamics is an argument in favor of the existence of specialized ribosomes that would display specific translation. In fact in yeast, eS26-containing ribosomes enable the translation of specific mRNAs, depending on their Kozak sequence (Ferretti et al., 2017). These results indicate then that ribosomes composing polysomes differ from the monosomes in MCF7 cells at least with the presence of the RP eS26, which could play a role in the preferential translation of specific mRNAs based on nucleotide recognition. In fact, no differential detection of other RPs does not exclude the existence of differences regarding their post-translational modifications that could not be identified here, as well as rRNA modifications (Emmott et al., 2019; Monaco et al., 2018; Penzo and Montanaro, 2018). This result contributes then to the controversy regarding specialized ribosomes. Concerning extra-ribosomal proteins, only 614 of the detected proteins are known to interact

with ribosomes, according to a list of around 950 proteins in common between two studies investigating the riboproteome (Imami et al., 2018; Reschke et al., 2013). All proteins detected following ribosome purification may indeed not be directly interacting with ribosomes or the translational machinery, as the sample can be contaminated by proteins that co-sediment with ribosomes during sucrose gradient ultracentrifugation. For example, mitochondrial RPs were identified, as reported in other studies (Imami et al., 2018). This caveat does not exclude however the presence of proteins interacting with the translational machinery that were not previously characterized. Also, even if a part of the detected proteins do not interact with ribosomes, these proteins still convey information about biological processes that can differ between cell conditions. According to our mass spectrometry analysis, few proteins (13%) are differentially expressed between the riboproteomes of MRTFA-ΔN200-overexpressing and quiescent MCF7 cells, which suggests a subtle regulation of the translational machinery whereas both cell population display strong phenotype differences. Among the proteins differentially detected between both conditions, we identified a functional enrichment linked to the cytoskeleton after inducing MRTFA-ΔN200, which is consistent with the phenotype of this stable cell line. MRTFA activation indeed induces actin polymerization, leading to a global cytoskeleton remodeling and the stimulation of cell migration (Medjkane et al., 2009). The cytoskeleton is in fact implicated in many signaling processes, and it was shown that components of the translational machinery interact with it, notably polysomes with actin. We detected indeed more β-actin in heavy polysomes than in light polysomes or monosomes when performing Western blot with proteins purified after polysome profiling from quiescent MCF7 cells (data not shown). In addition, ribosomes and other translation factors were observed at focal adhesions in migrating cells (Lenk et al., 1977; Simpson et al., 2020). These observations highlight then a link between the regulation of translation and the cytoskeleton, explaining why cytoskeleton actors were more detected with ribosomes after inducing MRTFA-ΔN200 in MCF7 cells. On the other side, proteins linked to RNA metabolism were enriched in the riboproteome of quiescent MCF7 cells. We identified notably DDX6 that could participate to the repression of AU3-rich mRNAs translation in G0 phase, whereas its reduction in more proliferating cells would release such mRNAs, enabling their translation. In addition, it was shown that DDX6 interacts with the Ataxin-2 cytoplasmic polyadenylation complex, resulting in an increase in target mRNA stability and translation (Inagaki et al., 2021). As these mRNAs are GC-rich, this study reveals another mechanism by which DDX6 could participate to the translational regulation of mRNAs depending on their GC%, by supporting GC-rich mRNA translation in quiescent MCF7 cells. We also identified RBM47 as an interesting candidate in the translational regulation of mRNAs. RBM47 was identified as an mRNA-binding protein that modulates their splicing and stability, and whose expression loss is associated with breast cancer progression and metastasis (Vanharanta et al., 2014). Our data are consistent with this observation as RBM47 is more detected in quiescent MCF7 cells than after expressing MRTFA-ΔN200 which induces a basal-like phenotype. It appears in fact that many RBM proteins bind AU-rich elements, notably in the 3'UTR of mRNAs (Z. Li et al., 2021). We determined indeed that RBM47 binds AU3-rich mRNAs, without altering their expression. A hypothesis is that the AU3-rich mRNAs bound by RBM47 are sequestered in the cytoplasm and, when RBM47 is lost in more aggressive cancer cells, these mRNAs are liberated and suitable for translation.

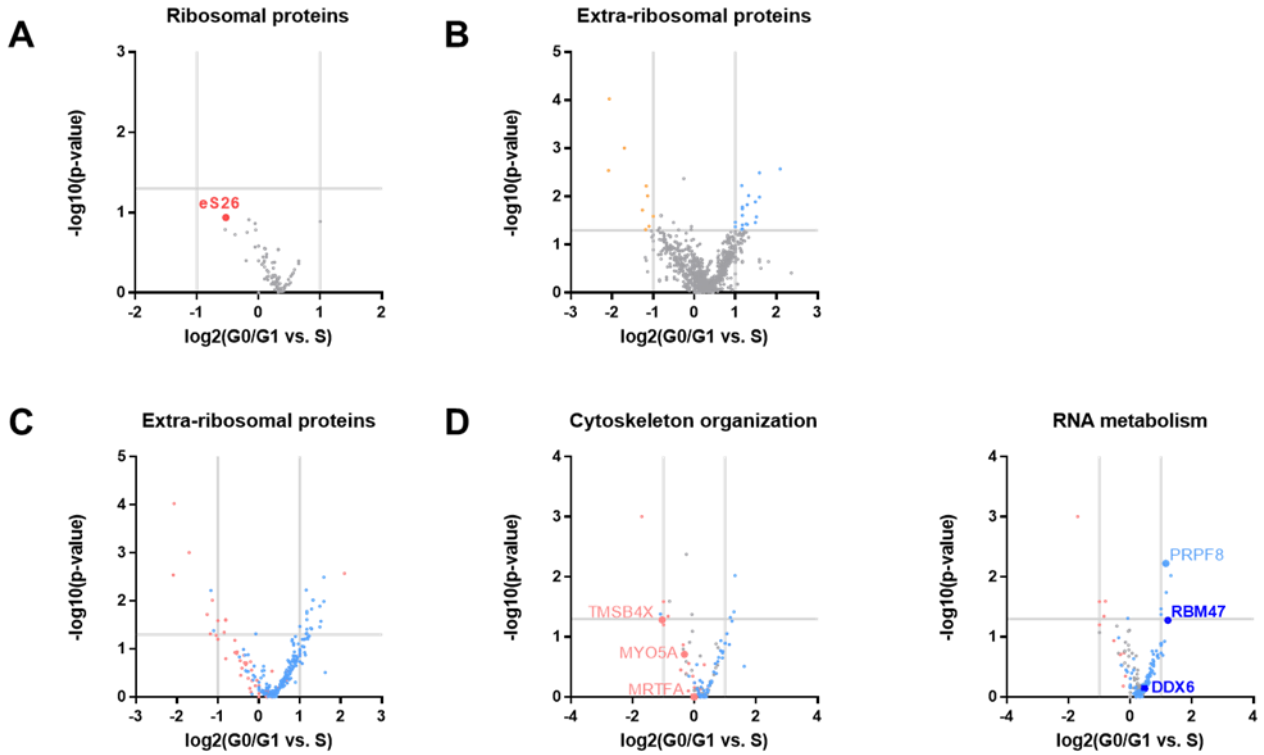
Another member of the RBM protein family, RBM33, appeared in the contrary to be enriched after MRTFA-ΔN200 induction compared to quiescent MCF7 cells. This protein is poorly characterized, but a recent study highlighted its role in nuclear export of GC-rich RNAs (Thomas et al., 2022). RBM33 did not seem to be involved in the stability of the long non-coding RNA investigated in the previous study once exported in the cytoplasm, but data are lacking concerning mRNAs recognized by RBM33. Such RNA-binding proteins could then play a role in the translational regulation depending on GC% that it would be worth investigating.

Altogether, this study highlights the existence of a dynamics in global translation during cell cycle and transformation, with an increase towards polysome-mediated translation. As polysomes translate in part specific mRNAs that are AU3-rich and implicated in cell cycle regulation, this translational dynamics seems to play a functional role necessary for the coordination of gene and protein expression of key factors. Such a regulation of translation may be linked to modifications of the translational machinery composition, notably through the existence of specialized ribosomes forming polysomes. The cytoskeleton and RNA-binding proteins are also players worth to investigating, which paves the way for understanding new regulatory processes of the translation depending on GC3% and cell state. Such a regulation could then alter the co-translational folding of proteins such as ERα, playing a role in the modification of its functions during cell cycle and transformation, leading to endocrine resistance in breast cancer.

Supporting information



Supplementary Figure S1: mRNAs translational profiles remain the same after cell synchronization. Examples of mRNA detected by RT-qPCR along the sucrose gradient following polysome profiling in MCF7 cells in G0/G1 (blue lines), S (orange lines) or S/G2 (green lines) phase. In red, green and grey squares are mRNAs more detected in monosomes, polysomes or without differential detection (NS) through RNA sequencing in quiescent MCF7 cells, respectively. Data correspond to the mean values from duplicate or triplicate experiments ± SEM. 0: first sample containing no ribosome; M: monosomes; LP: light polysomes; HP: heavy polysomes.



Supplementary Figure S2: Difference in the composition of the translational machinery between quiescent and proliferating (S phase) MCF7 cells. Proteins were analyzed by mass spectrometry following ribosome purification in MCF7 cells in G0/G1 and S phases. Peptides from quadrulicate experiments were quantified and analyzed as described in the materials and methods section. (**A-D**) Volcano plots of proteins detected by mass spectrometry following ribosome purification in MCF7 cell samples. (**A**) Ribosomal proteins (RPs); in red is highlighted the RP more detected in MCF7 overexpressing MRTFA-ΔN200 (from Figure 7C). (**B**) Extra-ribosomal proteins; in orange are depicted those enriched in S phase, in G0/G1 phase in blue, or without differential detection in gray. (**C**) Extra-ribosomal proteins, with highlighting only those who were more detected in G0/G1 phase (in blue) or after MRTFA-ΔN200 induction (in red), from Figure 7D. (**D**) Proteins involved in cytoskeleton organization or RNA metabolism; those whose the detection was higher in MCF7 overexpressing MRTFA-ΔN200 appear in red, in quiescent MCF7 in blue, or without differential detection in gray (from Figure 7E). Darker colors indicate proteins investigated in this study.

Supplementary Table 1: Fractions pooled after polysome profiling before extracting RNA.

Sample	Pooled fractions	Ribosomes contained
0	3 + 4	None
M	24 + 25	Monosomes
M	26 + 27	Monosomes
LP	34 + 35	Light polysomes
LP	37 + 38	Light polysomes
HP	44 + 45	Heavy polysomes
HP	47 + 48	Heavy polysomes

Supplementary Table 2: Primers used to detect target genes by RT-qPCR.

Target gene	Forward primer	Reverse Primer
Actin β (ACTB)	CATCCGCAAAGACCTGTACG	CCTGCTTGCTGATCCACATC
AKT serine/threonine kinase 1 (AKT1)	GTGTACGAGAAGAAGCTCAG	GTGTGATGGTGATCATCTGG
Amphiregulin (AREG)	CCTGGCTATATTGTCGATTCA	GTATTTTCACTTTCCGTCTTGTTTTG
Cyclin dependent kinase 1 (CDK1)	TTGGAAATTGAGCGGAGAGC	GGTATGGTAGATCCCGGCTT
Estrogen receptor α (ESR1)	TTTATGGGAAAAGGCTCAAA	GACAAAACCGAGTCACATCA
Transfected ERα	ACCCACTGCTTACTGGCTTA	GCTTAAGTTTAAACGCTAGCCAG
Fos proto-oncogene, AP-1 transcription factor subunit (FOS)	GAATTAACCTGGTGCTGGAT	GAACACACTATTGCCAGGAA
Forkhead box C1 (FOXC1)	CCCCCATGAGCGTGTACTC	ATGCCGTTCAGGGTGATCTTC
GATA binding protein 3 (GATA3)	GCCGTTGAGGGTTTCAGAGA	TCCGAGCACAACCACCTTAG
Glucocorticoid receptor (NR3C1)	CAGTGGAAGGACAGCACAAT	TTCCAGGTTCATTCCAGCCT
Ribosomal protein L3 like (RPL3L)	ATTACGCTGAGAAAGTCCCT	AACTTGGAGGTGGTGTCAAT
Ribonucleotide reductase regulatory subunit M2 (RRM2)	ACACAAACCATCGGAGGAGA	CCAATGAGCTTCACAGGCAA
Vascular endothelial growth factor B (VEGFB)	ACCAGAGGAAAGTGGTGTC	CCATGAGCTCCACAGTCAA

Supplementary Table 3: Ribosomal proteins identified by mass spectrometry. Paralogs are indicated in purple, and proteins that were not detected are highlighted in blue.

Large subunit proteins			Small subunit proteins		
New nomenclature	**Old nomenclature**	**Uniprot ID**	**New nomenclature**	**Old nomenclature**	**Uniprot ID**
uL1	L10A	P62906	eS1	S3A	P61247
uL2	L8	P62917	uS2	SA	P08865
uL3	L3	P39023	uS3	S3	P23396
uL4	L4	P36578	uS4	S9	P46781
uL5	L11	P62913	eS4	S4	P62701
uL6	L9	P32969	uS5	S2	P15880
eL6	L6	Q02878	eS6	S6	P62753
eL8	L7A	P62424	uS7	S5	P46782
uL11	L12	P30050	eS7	S7	P62081
uL13	L13A	P40429	uS8	S15A	P62244
eL13	L13	P26373	eS8	S8	P62241
uL14	L23	P62829	uS9	S16	P62249
eL14	L14	P50914	uS10	S20	P60866
uL15	L27A	P46776	eS10	S10	P46783
eL15	L15	P61313	uS11	S14	P62263
uL16	L10	P27635	uS12	S23	P62266
uL18	L5	P46777	eS12	S12	P25398
eL18	L18	Q07020	uS13	S18	P62269
eL19	L19	P84098	uS14	S29	P62273
eL20	L18A	Q02543	uS15	S13	P62277
eL21	L21	P46778	uS17	S11	P62280
uL22	L17	P18621	eS17	S17	P08708
eL22	L22	P35268	uS19	S15	P62841
uL23	L23A	P62750	eS19	S19	P39019
uL24	L26	P61254	eS21	S21	P63220
eL24	L24	P83731	eS24	S24	P62847
eL27	L27	P61353	eS25	S25	P62851
eL28	L28	P46779	eS26	S26	P62854
uL29	L35	P42766	eS27	S27	P42677
eL29	L29	P47914	eS28	S28	P62857
uL30	L7	P18124	eS30	S30	P62861
eL30	L30	P62888	eS31	S27A	P62979

eL31	L31	P62899	RACK1	RACK1	P63244
eL32	L32	P62910	**eS27L**	**S27L**	**Q71UM5**
eL33	L35A	P18077			
eL34	L34	P49207			
eL36	L36	Q9Y3U8			
eL37	L37	P61927			
eL38	L38	P63173			
eL39	**L39**	**P62891**			
eL40	**L40**	**P62987**			
eL41	**L41**	**P62945**			
eL43	L37A	P61513			
eL44	L36A	P83881			
P1/P2	**LP1**	**P05386**			
P2	LP2	P05387			
P0	LP0	P05388			
eL22L	**L22L**	**Q6P5R6**			
uL24L	**L26L**	**Q9UNX3**			
eL42L	**L36L**	**Q969Q0**			
uL30L	**L7L**	**Q6DKI1**			

References

Arava Y, Wang Y, Storey JD, Liu CL, Brown PO, and Herschlag D (2003) Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae *Proc Natl Acad Sci* **100**, 3889–3894.

Arnal J-F, Lenfant F, Metivier R, Flouriot G, Henrion D, Adlanmerini M, … Katzenellenbogen J (2017) Membrane and Nuclear Estrogen Receptor Alpha Actions: From Tissue Specificity to Medical Implications *Physiol Rev* **97**, 1045–1087.

Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, and Yusupov M (2011) The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution *Science* **334**, 1524–1529.

Biever A, Glock C, Tushev G, Ciirdaeva E, Dalmay T, Langer JD, and Schuman EM (2020) Monosomes actively translate synaptic mRNAs in neuronal processes *Science* **367**, eaay4991.

Bouyssié D, Hesse A-M, Mouton-Barbosa E, Rompais M, Macron C, Carapito C, … Bruley C (2020) Proline: an efficient and user-friendly software suite for large-scale proteomics *Bioinformatics* **36**, 3148–3155.

Courel M, Clément Y, Bossevain C, Foretek D, Vidal Cruchez O, Yi Z, … Weil D (2019) GC content shapes mRNA storage and decay in human cells *ELife* **8**, e49708.

Couté Y, Bruley C, and Burger T (2020) Beyond Target–Decoy Competition: Stable Validation of Peptide and Protein Identifications in Mass Spectrometry-Based Discovery Proteomics *Anal Chem* **92**, 14898–14906.

Dalla Venezia N, Vincent A, Marcel V, Catez F, and Diaz J-J (2019) Emerging Role of Eukaryote Ribosomes in Translational Control *Int J Mol Sci* **20**, 1226.

Dittmar KA, Goodenbour JM, and Pan T (2005) Tissue Specific Differences in Human Transfer RNA Expression *PLoS Genet* **preprint**, e221.

Emmott E, Jovanovic M, and Slavov N (2019) Ribosome Stoichiometry: From Form to Function *Trends Biochem Sci* **44**, 95–109.

Eremenko T, and Volpe P (1975) Polysome Translational State during the Cell Cycle *Eur J Biochem* **52**, 203–210.

Fermin D, Basrur V, Yocum AK, and Nesvizhskii AI (2011) Abacus: A computational tool for extracting and pre-processing spectral count data for label-free quantitative proteomic analysis *PROTEOMICS* **11**, 1340–1345.

Fernández-Calero T, Davyt M, Perelmuter K, Chalar C, Bampi G, Persson H, … Marín M (2020) Fine-tuning the metabolic rewiring and adaptation of translational machinery during an epithelial-mesenchymal transition in breast cancer cells *Cancer Metab* **8**, 8.

Ferretti MB, Ghalei H, Ward EA, Potts EL, and Karbstein K (2017) Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements *Nat Struct Mol Biol* **24**, 700–707.

Flouriot G, Huet G, Demay F, Pakdel F, Boujrad N, and Michel D (2014) The actin/MKL1 signalling pathway influences cell growth and gene expression through large-scale chromatin reorganization and histone post-translational modifications *Biochem J* **461**, 257–268.

Fornasiero EF, and Rizzoli SO (2019) Pathological changes are associated with shifts in the employment of synonymous codons at the transcriptome level *BMC Genomics* **20**, 566.

Fusco CM, Desch K, Dörrbaum AR, Wang M, Staab A, Chan ICW, … Schuman EM (2021) Neuronal ribosomes exhibit dynamic and context-dependent exchange of ribosomal proteins *Nat Commun* **12**, 6127.

Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, … Pilpel Y (2014) A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation *Cell* **158**, 1281–1292.

Goodarzi H, Nguyen HCB, Zhang S, Dill BD, Molina H, and Tavazoie SF (2016) Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression *Cell* **165**, 1416–1427.

Guimaraes JC, Mittal N, Gnann A, Jedlinski D, Riba A, Buczak K, … Zavolan M (2020) A rare codon-based translational program of cell proliferation *Genome Biol* **21**, 44.

Guzzi N, Cieśla M, Ngoc PCT, Lang S, Arora S, Dimitriou M, … Bellodi C (2018) Pseudouridylation of tRNA-Derived Fragments Steers Translational Control in Stem Cells *Cell* **173**, 1204-1216.e26.

Hanker AB, Sudhan DR, and Arteaga CL (2020) Overcoming Endocrine Resistance in Breast Cancer *Cancer Cell* **37**, 496–513.

Henglein B, Chenivesse X, Wang J, Eick D, and Bréchot C (1994) Structure and cell cycle-regulated transcription of the human cyclin A gene. *Proc Natl Acad Sci* **91**, 5490–5494.

Heyer EE, and Moore MJ (2016) Redefining the Translational Status of 80S Monosomes *Cell* **164**, 757–769.

Hia F, Yang SF, Shichino Y, Yoshinaga M, Murakawa Y, Vandenbon A, … Takeuchi O (2019) Codon bias confers stability to human mRNAs *EMBO Rep* **20**.

Hubstenberger A, Courel M, Bénard M, Souquere S, Ernoult-Lange M, Chouaib R, … Weil D (2017) P-Body Purification Reveals the Condensation of Repressed mRNA Regulons *Mol Cell* **68**, 144-157.e5.

Huet G, Mérot Y, Percevault F, Tiffoche C, Arnal J-F, Boujrad N, … Flouriot G (2009) Repression of the Estrogen Receptor-α Transcriptional Activity by the Rho/Megakaryoblastic Leukemia 1 Signaling Pathway *J Biol Chem* **284**, 33729–33739.

Imami K, Milek M, Bogdanow B, Yasuda T, Kastelic N, Zauber H, … Selbach M (2018) Phosphorylation of the Ribosomal Protein RPL12/uL11 Affects Translation during Mitosis *Mol Cell* **72**, 84-98.e9.

Inagaki H, Hosoda N, and Hoshino S (2021) DDX6 is a positive regulator of Ataxin-2/PAPD4 cytoplasmic polyadenylation machinery *Biochem Biophys Res Commun* **553**, 9–16.

Jehanno C, Fernandez-Calero T, Habauzit D, Avner S, Percevault F, Jullion E, … Flouriot G (2020) Nuclear accumulation of MKL1 in luminal breast cancer cells impairs genomic activity of ERα and is associated with endocrine resistance *Biochim Biophys Acta BBA - Gene Regul Mech* **1863**, 194507.

Khatter H, Myasnikov AG, Mastio L, Billas IML, Birck C, Stella S, and Klaholz BP (2014) Purification, characterization and crystallization of the human 80S ribosome *Nucleic Acids Res* **42**, e49–e49.

Kudla G, Lipinski L, Caffin F, Helwak A, and Zylicz M (2006) High Guanine and Cytosine Content Increases mRNA Levels in Mammalian Cells *PLoS Biol* **4**, e180.

Langmead B, and Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2 *Nat Methods* **9**, 357–359.

Lenk R, Ransom L, Kaufmann Y, and Penman S (1977) A cytoskeletal structure with associated polyribosomes obtained from HeLa cells *Cell* **10**, 67–78.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, … 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools *Bioinformatics* **25**, 2078–2079.

Li Z, Guo Q, Zhang J, Fu Z, Wang Y, Wang T, and Tang J (2021) The RNA-Binding Motif Protein Family in Cancer: Friend or Foe? *Front Oncol* **11**, 757135.

Liao Y., Smyth GK, and Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features *Bioinformatics* **30**, 923–930.

Liao Yang, Smyth GK, and Shi W (2019) The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads *Nucleic Acids Res* **47**, e47–e47.

Liu Y, Chen S, Wang S, Soares F, Fischer M, Meng F, … He HH (2017) Transcriptional landscape of the human cell cycle *Proc Natl Acad Sci* **114**, 3473–3478.

Love MI, Huber W, and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 *Genome Biol* **15**, 550.

Manavathi B, Dey O, Gajulapalli VNR, Bhatia RS, Bugide S, and Kumar R (2013) Derailed Estrogen Signaling and Breast Cancer: An Authentic Couple *Endocr Rev* **34**, 1–32.

Medjkane S, Perez-Sanchez C, Gaggioli C, Sahai E, and Treisman R (2009) Myocardin-related transcription factors and SRF are required for cytoskeletal dynamics and experimental metastasis *Nat Cell Biol* **11**, 257–268.

Meyer D, Kames J, Bar H, Komar AA, Alexaki A, Ibla J, … Kimchi-Sarfaty C (2021) Distinct signatures of codon and codon pair usage in 32 primary tumor types in the novel database CancerCoCoPUTs for cancer-specific codon usage *Genome Med* **13**, 122.

Minshall N, Kress M, Weil D, and Standart N (2009) Role of p54 RNA Helicase Activity and Its C-terminal Domain in Translational Repression, P-body Localization and Assembly *Mol Biol Cell* **20**, 2464–2472.

Miralles F, Posern G, Zaromytidou A-I, and Treisman R (2003) Actin Dynamics Control SRF Activity by Regulation of Its Coactivator MAL *Cell* **113**, 329–342.

Monaco P, Marcel V, Diaz J-J, and Catez F (2018) 2'-O-Methylation of Ribosomal RNA: Towards an Epitranscriptomic Control of Translation? *Biomolecules* **8**, 106.

Panda AC, Martindale JL, and Gorospe M (2017) Polysome Fractionation to Analyze mRNA Distribution Profiles *Bio-Protoc* **7**, e2126.

Papagiannopoulos CI, Kyritsis KA, Psatha K, Mavridou D, Chatzopoulou F, Orfanoudaki G, … Vizirianakis IS (2022) Invariable Ribosome Stoichiometry During Murine Erythroid Differentiation: Implications for Understanding Ribosomopathies *Front Mol Biosci* **9**, 805541.

Penzo M, and Montanaro L (2018) Turning Uridines around: Role of rRNA Pseudouridylation in Ribosome Biogenesis and Ribosomal Function *Biomolecules* **8**, 38.

Pham TH, Page YL, Percevault F, Ferrière F, Flouriot G, and Pakdel F (2021) Apigenin, a Partial Antagonist of the Estrogen Receptor (ER), Inhibits ER-Positive Breast Cancer Cell Proliferation through Akt/FOXM1 Signaling *Int J Mol Sci* **22**, 470.

Pham TV, Piersma SR, Warmoes M, and Jimenez CR (2010) On the beta-binomial model for analysis of spectral count data in label-free tandem mass spectrometry-based proteomics *Bioinformatics* **26**, 363–369.

Poulard C, Treilleux I, Lavergne E, Bouchekioua-Bouzaghou K, Goddard-Léon S, Chabaud S, … Le Romancer M (2012) Activation of rapid oestrogen signalling in aggressive human breast cancers *EMBO Mol Med* **4**, 1200–1213.

Pouyet F, Mouchiroud D, Duret L, and Sémon M (2017) Recombination, meiotic expression and human codon usage *ELife* **6**, e27344.

Rau A, Gallopin M, Celeux G, and Jaffrézic F (2013) Data-based filtering for replicated high-throughput transcriptome sequencing experiments *Bioinformatics* **29**, 2146–2152.

Reschke M, Clohessy JG, Seitzer N, Goldstein DP, Breitkopf SB, Schmolze DB, … Pandolfi PP (2013) Characterization and Analysis of the Composition and Dynamics of the Mammalian Riboproteome *Cell Rep* **4**, 1276–1287.

Rudolph KLM, Schmitt BM, Villar D, White RJ, Marioni JC, Kutter C, and Odom DT (2016) Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States *PLOS Genet* **12**, e1006024.

Rusidzé M, Adlanmérini M, Chantalat E, Raymond-Letron I, Cayre S, Arnal J-F, … Lenfant F (2021) Estrogen receptor-α signaling in post-natal mammary development and breast cancers *Cell Mol Life Sci* **78**, 5681–5705.

Shen L, Su Z, Yang K, Wu C, Becker T, Bell-Pedersen D, … Sachs MS (2021) Structure of the translating Neurospora ribosome arrested by cycloheximide *Proc Natl Acad Sci* **118**, e2111862118.

Simoes Eugénio M, Faurez F, Kara-Ali GH, Lagarrigue M, Uhart P, Bonnet MC, … Dimanche-Boitrel M-T (2021) TRIM21, a New Component of the TRAIL-Induced Endogenous Necrosome Complex *Front Mol Biosci* **8**, 645134.

Simpson LJ, Tzima E, and Reader JS (2020) Mechanical Forces and Their Effect on the Ribosome and Protein Translation Machinery *Cells* **9**, 650.

Stein KC, and Frydman J (2019) The stop-and-go traffic regulating protein biogenesis: How translation kinetics controls proteostasis *J Biol Chem* **294**, 2076–2084.

Strunk BS, Novak MN, Young CL, and Karbstein K (2012) A Translation-Like Cycle Is a Quality Control Checkpoint for Maturing 40S Ribosome Subunits *Cell* **150**, 111–121.

Thomas A, Rehfeld F, Zhang H, Chang T-C, Goodarzi M, Gillet F, and Mendell JT (2022) RBM33 directs the nuclear export of transcripts containing GC-rich elements *Genes Dev* **36**, 550–565.

Vanharanta S, Marney CB, Shu W, Valiente M, Zou Y, Mele A, … Massagué J (2014) Loss of the multifunctional RNA-binding protein RBM47 as a source of selectable metastatic traits in breast cancer *ELife* **3**, e02734.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, … Chanda SK (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets *Nat Commun* **10**, 1523.

Discussion et perspectives

Le traitement du cancer du sein est un enjeu de santé publique majeur. En effet, celui-ci est le plus diagnostiqué dans le monde, et le plus meurtrier chez la femme. Dans 70% des cas, le cancer du sein est de type luminal, exprimant le récepteur aux œstrogènes ERα. Bien qu'un traitement spécifique existe dans ce cas, visant à priver la tumeur en œstrogènes ou directement inhiber l'activité d'ERα, un phénomène de résistance a lieu pour 30% des patientes. Dans ce cas, les cellules cancéreuses continuent à proliférer indépendamment du contrôle hormonal, se dédifférencient, et peuvent mener à des métastases. Une altération de l'expression ou activité d'ERα est au cœur de cette résistance. Celle-ci peut être due à des mutations activatrices, un déséquilibre entre coactivateurs et corépresseurs, ou la suractivation de voies de signalisation de facteurs de croissance. Considérant la capacité d'ERα à avoir deux activités distinctes, génomique et non-génomique, en dehors de toute modification de sa séquence polypeptidique, il est probable que cette protéine soit sensible au phénomène de repliement co-traductionnel. Dans ce cas, l'évolution de l'environnement cellulaire au cours de la transformation tumorale pourrait modifier la traduction d'ERα et ainsi impacter sa conformation, d'où l'altération de ses propriétés fonctionnelles menant à la résistance hormonale. Le principal paramètre régulant le repliement co-traductionnel étant la vitesse de traduction, les acteurs les plus probablement impliqués sont alors le ribosome et l'usage des codons.

Il a notamment été montré que les codons synonymes n'étaient pas utilisés à une même fréquence au sein des séquences codantes, ceci étant alors nommé biais d'usage des codons. Plus surprenant, ce biais d'usage des codons serait différent pour exprimer des protéines spécifiques des processus de différentiation ou de prolifération [99]. De telles variations dans l'usage des codons étant également retrouvées en comparant des tissus sains et cancéreux [120], cela suggère que la composition en codons d'un ARNm joue un rôle fonctionnel dans son expression et/ou traduction. Le rôle joué par l'usage des codons pourrait être lié à leur adéquation avec le stock d'ARNt disponible, mais des résultats contradictoires ont été obtenus concernant de potentielles variations d'ARNt entre différentes conditions biologiques [99,121–124]. Nos propres résultats ayant montré une absence de modification majeure entre les stocks d'ARNt de cellules MCF7 natives et transformées, nous considérons que le rôle joué par l'usage des codons dans le repliement co-traductionnel n'est pas lié à son adéquation aux ARNt disponibles. Nous souhaitons donc comprendre quels paramètres dans la composition en codons participent à la régulation de la traduction pour la production de protéines fonctionnelles en fonction de l'état cellulaire.

Par ailleurs, de plus en plus d'études mettent en évidence l'existence de ribosomes spécialisés. Ceux-ci sont caractérisés par des modifications de leur composition en protéines, modifications post-traductionnelles, ou modifications d'ARNr, menant à la coexistence de ribosomes différents au sein d'une même cellule. Ces ribosomes présenteraient alors des propriétés particulières menant à la traduction d'ARNm spécifiques en fonction de l'état cellulaire. Ce concept reste controversé, de nombreuses études concluant à l'absence de

modifications d'un des paramètres composant le ribosome (sa composition en protéines par exemple), n'excluant pas pour autant la possibilité qu'un paramètre non mesuré soit lui modifié dans les conditions étudiées (comme des modifications d'ARNr) [125]. Déterminer si des ribosomes spécialisés existent en cellules de cancer du sein ERα+, pouvant jouer un rôle dans la traduction d'ARNm spécifiques en fonction de leur composition en codons, représente alors un intérêt majeur pour mieux comprendre le fonctionnement de ces cellules cancéreuses.

L'objectif du projet étant d'élucider le rôle joué par la composition en codons dans la traduction de l'ARNm d'ERα en protéine fonctionnelle, nous avons tout d'abord généré une séquence synonyme d'ERα, dont les codons étaient optimisés en respectant les fréquences déterminées par Gingold et al. [99]. ERα étant normalement exprimé en cellules quiescentes, différenciées, ses codons ne semblaient en effet pas adaptés, d'un point de vue de leur fréquence d'usage, à un contexte cellulaire plus prolifératif. Cette optimisation de codons a en effet permis de produire un récepteur fonctionnel en cellules cancéreuses mammaires, en restaurant par ailleurs un phénotype plus répressif. L'activité transcriptionnelle d'ERα médiée par la fonction AF1 était en effet améliorée, de meilleures capacités répressives en présence de corépresseurs étaient observées, et les interactions avec des kinases étaient diminuées. Aussi, la version synonyme d'ERα réduisait la prolifération induite par les œstrogènes en abolissant leur effet anti-apoptotique en cellules MCF7. Enfin, nous avons montré que ces modifications fonctionnelles étaient associées à des modifications de la conformation d'ERα, ce qui a mis en évidence un rôle joué par l'usage des codons dans la conformation et la fonctionnalité d'ERα en fonction de l'état cellulaire (chapitre 1).

Pour aller plus loin, nous avons remarqué que le biais d'usage des codons en fonction de l'état cellulaire reposait sur la nature de la troisième base des codons. Plus précisément, les gènes spécifiquement exprimés dans des cellules quiescentes ou en conditions différenciées ont tendance à être enrichis en codons terminant par G ou C (GC3), alors que les gènes spécifiquement exprimés en phase S à M ou en cellules dédifférenciées sont enrichis en codons terminant par A ou T (AT3). Cet enrichissement a été confirmé en explorant des données transcriptomiques issues de cellules plus ou moins différenciées, mais aussi de types de cancers du sein plus ou moins prolifératifs. Outre un effet sur le niveau d'expression des messagers, cet enrichissement en GC3 apparaissait également corrélé à l'efficacité de leur traduction et la quantité de protéine produite en fonction du cycle cellulaire. Ces résultats ont ainsi mis en évidence une coordination entre la transcription et la traduction pour la production optimale de certaines protéines selon leur enrichissement en GC3 en fonction de l'état cellulaire (chapitre 2). Un autre paramètre jouant certainement un rôle dans cette régulation du protéome est la stabilité des ARNm. Bien que n'ayant pas nous-même mesuré ce paramètre, une étude a effet montré qu'un enrichissement en GC3 était lié à une meilleure stabilité des messagers [126,127]. Suite à ces observations, nous avons confirmé le rôle joué par la troisième base des codons dans la régulation de l'expression et activité des protéines en générant de nouvelles séquences synonymes d'ERα, dont 100, 50 ou 0% des codons terminent par G ou C. En accord avec les observations réalisées à partir d'études omiques, il est apparu que les ERα SYN WT et 100% GC3 étaient

mieux produits en cellules quiescentes, tandis que la production des SYN 50 et 0% GC3 n'était possible qu'en phase S. Ces derniers SYN ne semblaient cependant pas fonctionnels vis-à-vis de leur activité transcriptionnelle, suggérant que l'enrichissement en codons terminant par A ou T modifie le repliement co-traductionnel du récepteur. La conformation de ces versions riches en AT3 ne semblait cependant pas dramatiquement altérée, puisque ces protéines étaient toujours reconnues par des anticorps ciblant les extrémités N- et C-terminales du récepteur. Des expériences de digestion ménagées seraient alors nécessaires pour étudier la structure de ces versions d'ERα, comme réalisé lors du chapitre 1. Aussi, d'autres propriétés fonctionnelles de ces versions synonymes d'ERα pourraient être caractérisées, pour mieux comprendre l'impact de l'enrichissement en AT3 sur la fonctionnalité d'ERα. Enfin, de nouvelles versions synonymes d'ERα pourraient être étudiées, dont la séquence codante ne serait enrichie en AT3 qu'au niveau de certaines régions de la protéine, afin de plus finement déterminer l'impact de cet usage de codons sur l'expression ou différentes fonctions du récepteur.

Notre étude a donc mis en évidence une coordination entre la transcription et la traduction pour la production optimale de certaines protéines au cours du cycle en fonction de leur composition en codons, en particulier de la nature de la troisième base des codons. Ayant mis de côté le potentiel rôle joué par le stock d'ARNt dans une telle régulation, nous nous sommes intéressés au concept de ribosome spécialisé pour expliquer les résultats obtenus.

Dans un premier temps, nous avons étudié le rôle joué par les modifications des ARNr, en particulier la 2'-O-méthylation médiée par la fibrillarin (FBL), et la pseudouridylation réalisée par la dyskerin (DKC1). Il a notamment été montré que les niveaux de méthylation des ARNr évoluent lors de la tumorigenèse mammaire, correspondant à une surexpression de FBL en cellules de cancer du sein, cette surexpression étant associée à une réduction de la survie des patientes [92,128,129]. Un taux élevé de FBL serait en effet impliqué dans une augmentation de la traduction initiée par IRES, et une baisse de la fidélité de la traduction, ce qui participerait à la tumorigenèse mammaire entre autre [89]. La régulation de la méthylation semblant hétérogène selon les sites d'ARNr, notamment ceux proches des sites A et P du ribosome, cette modification d'ARNr pourrait en effet jouer un rôle fonctionnel lors de la traduction [130]. Concernant DKC1, une faible expression serait favorable à la tumorigenèse mammaire de par une diminution de la fidélité de la traduction et une reprogrammation fine de la traduction de certains messagers [90,91,131–133]. D'après des études réalisées chez la levure, certains sites de pseudouridylation des ARNr sont proches du lieu de reconnaissance des codons par les ARNt, suggérant un potentiel rôle dans la reconnaissance codon/anticodon [134]. Peu d'études existent cependant concernant le rôle de ces modifications d'ARNr dans les cellules de mammifères.

Nous avons donc réalisé des expériences d'inhibition de ces deux enzymes en utilisant de petits ARN interférents (siRNA) en lignées de cancer du sein plus ou moins différenciées, et mesuré par la suite l'activité transcriptionnelle d'ERα sauvage et synonyme. En déterminant si l'une de ces modifications d'ARNr pouvait jouer un rôle dans l'activité d'ERα dans ces conditions, cela aurait permis d'identifier une caractéristique des ribosomes impliquée dans le repliement co-traductionnel du récepteur, et donc la régulation de sa traduction

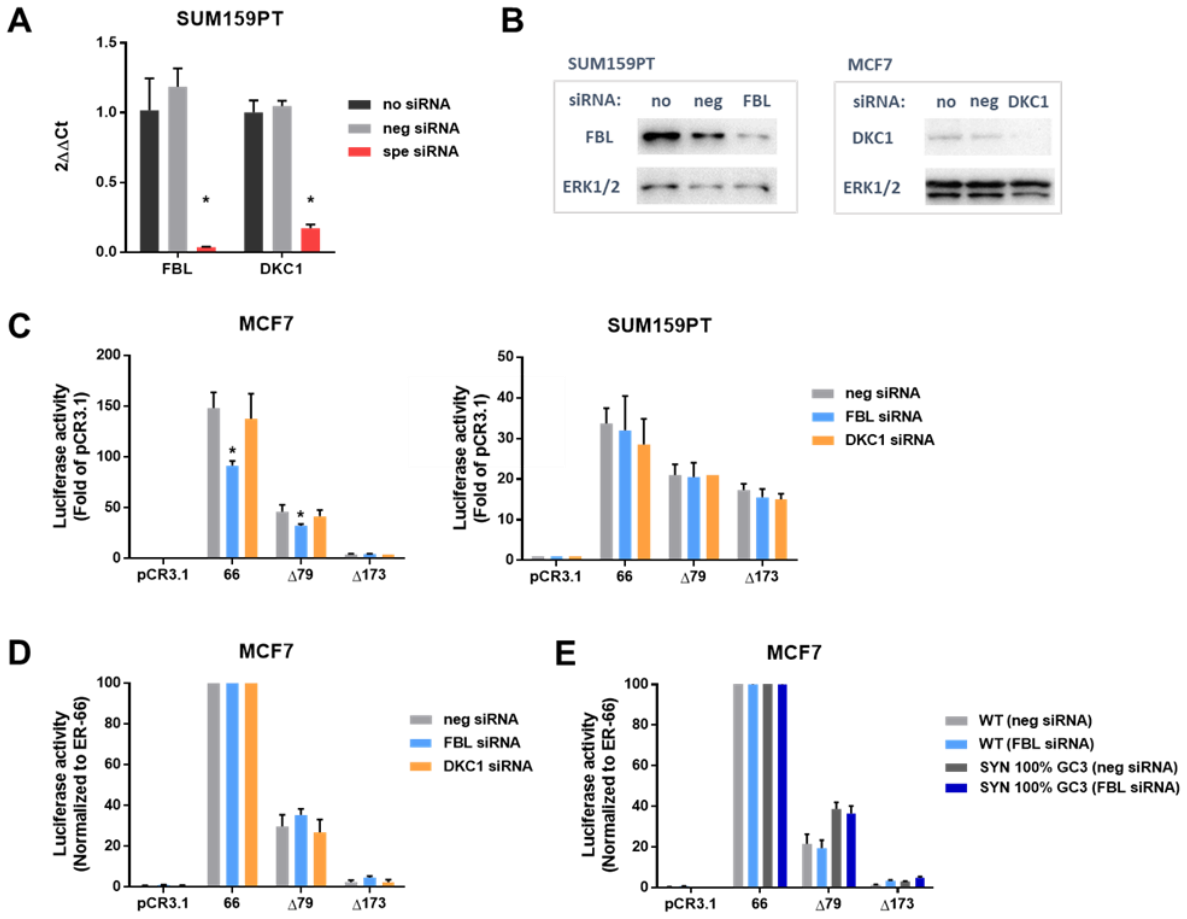


Figure 29 : Neither FBL nor DKC1 knockdown alters ERα activity in MCF7 and SUM159PT cells.

MCF7 and SUM159PT cells were routinely maintained as explained in chapter 1. Cells were plated in 6-well or 24-well plates for RT-qPCR and Western blot or luciferase assay, respectively. After 24 h, siRNA transfections were performed in dsFCS with jetPRIME (Polyplus transfection) following manufacturer's instructions. 24 h later, C3-Luc, CMV-βgal, and ERα-coding plasmids were transfected with JetPEI (Polyplus transfection) according to manufacturer's instructions. Before luciferase assay, cells were treated with 10 nM 17β-estradiol for 24 h. siRNAs targeting FBL, DKC1 or with no specific target (neg siRNA) were purchased from Sigma. Plasmids expressing full-length (ER-66) WT or SYN 100% GC3 ERα and their deleted forms (Δ79 and Δ173) were previously described (chapters 1 and 2). RT-qPCR, Western blot and luciferase assay were performed as described in chapter 1. The primary antibodies used for Western blotting were as follows: from Abcam against FBL (4A4; ab218846) and DKC1 (EPR10399; ab156877); from Cell Signaling Technology against ERK1/2 (137F5; #4695). (**A**) Validation of FBL and DKC1 inhibition at the mRNA level by RT-qPCR in SUM159PT. (**B**) Validation of FBL and DKC1 inhibition at the protein level by Western Blot. Representative images obtained from SUM159PT and MCF7 cells are showed. (**C**) ERα transcriptional activity in MCF7 and SUM159PT cells following FBL or DKC1 inhibition, expressed as fold change from pCR3.1 control. (**D**) ERα transcriptional activity in MCF7 cells following FBL or DKC1 inhibition, expressed as a percentage of full-length ERα activity. (**E**) ERα WT and SYN 100% GC3 transcriptional activities in MCF7 cells following FBL inhibition, expressed as a percentage of full-length ERα activity. Data correspond to the mean values from triplicate experiments ± SEM (*p-value <0.05, Student's t-test).

en fonction de sa composition en codons et de l'état cellulaire. Malgré l'efficacité d'inhibition de FBL et DKC1 par siRNA, aucune modification majeure de l'activité transcriptionnelle d'ERα n'a été détectée dans nos modèles cellulaires (Figure 29). Nous avons donc mis de côté l'hypothèse d'un rôle joué par ces modifications d'ARNr dans la régulation de la traduction en fonction de la composition en codon. Il se pourrait en réalité que la méthylation ou pseudouridylation des ARNr impacte plutôt l'initiation de la traduction et non son élongation, or c'est la phase d'élongation de la traduction qui est majoritairement impliquée dans la régulation du repliement co-traductionnel des protéines [130,131].

Suite à ces résultats, nous avons pris connaissance de l'étude d'Imami et al. mettant en évidence l'existence de ribosomes spécialisés, de par la phosphorylation d'une protéine ribosomale (RPL12 / uL11), impliqués dans la traduction d'ARNm spécifiquement exprimés en phase M. Par ailleurs, ces ribosomes spécialisés étaient préférentiellement détectés en monosomes, et non en polysomes [81].

Nous avons alors initié des expériences de « polysome profiling » afin d'identifier les ARNm spécifiquement traduits par des monosomes ou par des polysomes, et déterminer si cette traduction préférentielle était associée à des différences d'enrichissement en GC3. Les résultats obtenus ont en effet mis en évidence qu'une partie du translatome des cellules MCF7 était préférentiellement traduit par des monosomes ou polysomes, avec des messagers enrichis en GC3 en monosomes, et en AU3 en polysomes. Ces ARNm riches en AU3 correspondent à des messagers spécifiquement exprimés en phase S, et une dynamique de la traduction au cours du cycle cellulaire a été mise en évidence, avec une augmentation du niveau de polysomes en phase S. Ces observations suggèrent donc qu'une traduction préférentielle par des monosomes ou polysomes présente un rôle fonctionnel, impliqué dans la production optimale de certaines protéines au cours du cycle cellulaire, en fonction de leur composition en codons.

Afin d'identifier des acteurs de cette régulation de la traduction, nous avons eu recours à la spectrométrie de masse pour caractériser la composition des ribosomes et protéines associées selon l'état cellulaire. Les résultats obtenus ont cette fois-ci mis en évidence l'existence de ribosomes spécialisés en cellules MCF7 transformées, de par une présence plus abondante de protéine ribosomale eS26. D'autres protéines composant le riboprotéome étaient différentiellement détectées entre nos échantillons, avec notamment un enrichissement d'acteurs du cytosquelette chez les cellules transformées, et certaines protéines liant l'ARN en cellules MCF7 quiescentes.

Ces résultats mettent ainsi en évidence des paramètres pouvant réguler la traduction en fonction de la composition en codon et l'état cellulaire, qui pourraient alors être impliqués dans la régulation du repliement co-traductionnel d'ERα et de ses propriétés fonctionnelles au cours du cycle cellulaire ou de la transformation tumorale. Afin de mieux comprendre le rôle joué par chacun de ces acteurs dans une telle régulation de la traduction, eS26 pourrait être réprimé en cellules MCF7 transformées avant de déterminer si les propriétés fonctionnelles d'ERα en cellules MCF7 natives sont ainsi restaurées. Les ribosomes spécialisés tel qu'identifié dans ces cellules transformées pourraient également être purifiés, afin de réaliser des expériences de traduction in vitro d'ARNm de différents niveaux d'enrichissement en GC3, et étudier ainsi les propriétés de

ces ribosomes [135]. Considérant le potentiel rôle joué par le cytosquelette dans la régulation de la traduction, il serait intéressant d'investiguer par immunofluorescence une potentielle localisation particulière des ribosomes en fonction de l'état cellulaire. Des interactions avec certains membres du cytosquelette pourraient également être recherchées par test de proximité (« proximity ligation assay »). Enfin, les protéines telles que DDX6 et RBM47 pourraient être surexprimées en cellules MCF7 transformées, avant de mettre en œuvre les protocoles de « polysome profiling » et séquençage d'ARN pour déterminer si l'augmentation de ces protéines mène à une réduction de la traduction des messagers riches en AU3.

L'analyse du riboprotéome des cellules MCF7 a également mis en évidence la présence de cinq paralogues de protéines ribosomales : eS27L, eL22L, uL24L, eL42L et uL30L. Alors que 80 protéines constituent le ribosome eucaryote, plus de 2000 pseudogènes ont en effet été identifiés dans le génome humain. Il a d'abord été montré que trois d'entre eux permettaient en effet la production de protéines : eL42L, eL39L et uL16L [136]. Depuis, de nombreux autres paralogues de protéines ribosomales ont été détectés, mais tous n'ont pas encore fait l'objet de travaux permettant de confirmer leur incorporation au sein du ribosome ou de potentielles fonctions extra-ribosomales. Une expression et activité spécifiques aux cellules testiculaires et musculaires ont notamment été démontrées pour les paralogues eL39L et uL3L, respectivement [137]. Concernant les protéines identifiées dans nos échantillons, eS27L jouerait un rôle dans la stabilité de p53 et de l'ADN, son inactivation permettant d'augmenter la mort des cellules induite par des radiations [138]. Une forte expression de eS27L aurait par ailleurs été associée à un meilleur pronostic chez des patients atteint de cancer colorectal [139]. La protéine eL22L a quant à elle été identifiée comme un oncogène contribuant aux phénomènes de métastases et résistance aux traitements dans différents cancers tels que le cancer colorectal ou ceux de l'ovaire, de la prostate, ou du foie [140–143]. Une surexpression de uL24L a été détectée dans des cancers du sein, et pourrait être un marqueur de mauvais pronostic, de par une compétition avec sa protéine homologue uL24 qui favoriserait la traduction de l'anti-oncogène p53 [137]. Enfin, un rôle structurel a seulement été étudié pour eL42L, dont la localisation proche du centre de transfert peptidique (PTC) permettrait la formation de contacts avec les ARNt déacétylés en transition des sites P à E du ribosome [144]. Déterminer si ces protéines homologues présentent un rôle fonctionnel dans le cancer du sein serait alors intéressant dans la compréhension des mécanismes de la tumorigenèse mammaire.

Cette dernière étude ouvre ainsi la voie à de nouvelles pistes pour mieux comprendre la régulation de la traduction en fonction de la composition en codon et de l'état cellulaire. Ce nouveau niveau de complexité entrant en jeu dans l'expression des gènes en protéines pourrait alors être impliqué dans de nombreux phénomènes pathologiques en plus du cancer du sein et de la résistance hormonale. Aussi, de tels paramètres seraient à prendre en compte lors de projets de production de protéines in vitro. Une récente étude a ainsi montré par exemple que l'optimisation de codons pouvait impacter les propriétés thérapeutiques et immunogènes du facteur IX dans le traitement de l'hémophilie [145].

Références

1. Harbeck, N.; Penault-Llorca, F.; Cortes, J.; Gnant, M.; Houssami, N.; Poortmans, P.; Ruddy, K.; Tsang, J.; Cardoso, F. Breast Cancer. *Nat. Rev. Dis. Primer* **2019**, *5*, 66, doi:10.1038/s41572-019-0111-2.
2. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **2021**, *71*, 209–249, doi:10.3322/caac.21660.
3. Akram, M.; Iqbal, M.; Daniyal, M.; Khan, A.U. Awareness and Current Knowledge of Breast Cancer. *Biol. Res.* **2017**, *50*, 33, doi:10.1186/s40659-017-0140-9.
4. Liang, Y.; Zhang, H.; Song, X.; Yang, Q. Metastatic Heterogeneity of Breast Cancer: Molecular Mechanism and Potential Therapeutic Targets. *Semin. Cancer Biol.* **2020**, *60*, 14–27, doi:10.1016/j.semcancer.2019.08.012.
5. Tan, P.H.; Ellis, I.; Allison, K.; Brogi, E.; Fox, S.B.; Lakhani, S.; Lazar, A.J.; Morris, E.A.; Sahin, A.; Salgado, R.; et al. The 2019 World Health Organization Classification of Tumours of the Breast. *Histopathology* **2020**, *77*, 181–185, doi:10.1111/his.14091.
6. Perou, C.M.; Sørlie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; Rees, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; et al. Molecular Portraits of Human Breast Tumours. *Nature* **2000**, *406*, 747–752, doi:10.1038/35021093.
7. Trenner, A.; Sartori, A.A. Harnessing DNA Double-Strand Break Repair for Cancer Treatment. *Front. Oncol.* **2019**, *9*, 1388, doi:10.3389/fonc.2019.01388.
8. Łukasiewicz, S.; Czeczelewski, M.; Forma, A.; Baj, J.; Sitarz, R.; Stanisławek, A. Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review. *Cancers* **2021**, *13*, 4287, doi:10.3390/cancers13174287.
9. Garrido-Castro, A.C.; Lin, N.U.; Polyak, K. Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer Discov.* **2019**, *9*, 176–198, doi:10.1158/2159-8290.CD-18-1177.
10. Sun, Y.-S.; Zhao, Z.; Yang, Z.-N.; Xu, F.; Lu, H.-J.; Zhu, Z.-Y.; Shi, W.; Jiang, J.; Yao, P.-P.; Zhu, H.-P. Risk Factors and Preventions of Breast Cancer. *Int. J. Biol. Sci.* **2017**, *13*, 1387–1397, doi:10.7150/ijbs.21635.
11. McAnulty, J.; DiFeo, A. The Molecular 'Myc-Anisms' behind Myc-Driven Tumorigenesis and the Relevant Myc-Directed Therapeutics. *Int. J. Mol. Sci.* **2020**, *21*, 9486, doi:10.3390/ijms21249486.
12. Fallah, Y.; Brundage, J.; Allegakoen, P.; Shajahan-Haq, A.N. MYC-Driven Pathways in Breast Cancer Subtypes. *Biomolecules* **2017**, *7*, 53, doi:10.3390/biom7030053.
13. Venkitaraman, A.R. How Do Mutations Affecting the Breast Cancer Genes BRCA1 and BRCA2 Cause Cancer Susceptibility? *DNA Repair* **2019**, *81*, 102668, doi:10.1016/j.dnarep.2019.102668.
14. Csolle, M.P.; Ooms, L.M.; Papa, A.; Mitchell, C.A. PTEN and Other PtdIns(3,4,5)P3 Lipid Phosphatases in Breast Cancer. *Int. J. Mol. Sci.* **2020**, *21*, 9189, doi:10.3390/ijms21239189.
15. Christopoulos, P.F.; Msaouel, P.; Koutsilieris, M. The Role of the Insulin-like Growth Factor-1 System in Breast Cancer. *Mol. Cancer* **2015**, *14*, 43, doi:10.1186/s12943-015-0291-7.
16. Steelman, L.S.; Chappell, W.H.; Akula, S.M.; Abrams, S.L.; Cocco, L.; Manzoli, L.; Ratti, S.; Martelli, A.M.; Montalto, G.; Cervello, M.; et al. Therapeutic Resistance in Breast Cancer Cells Can Result from Deregulated EGFR Signaling. *Adv. Biol. Regul.* **2020**, *78*, 100758, doi:10.1016/j.jbior.2020.100758.
17. Silwal-Pandit, L.; Langerød, A.; Børresen-Dale, A.-L. TP53 Mutations in Breast and Ovarian Cancer. *Cold Spring Harb. Perspect. Med.* **2017**, *7*, a026252, doi:10.1101/cshperspect.a026252.
18. Slepicka, P.F.; Cyrill, S.L.; dos Santos, C.O. Pregnancy and Breast Cancer: Pathways to Understand Risk and Prevention. *Trends Mol. Med.* **2019**, *25*, 866–881, doi:10.1016/j.molmed.2019.06.003.

19. Winters, S.; Martin, C.; Murphy, D.; Shokar, N.K. Breast Cancer Epidemiology, Prevention, and Screening. In *Progress in Molecular Biology and Translational Science*; Elsevier, 2017; Vol. 151, pp. 1–32 ISBN 978-0-12-812772-8.
20. Calaf, G.; Ponce-Cusi, R.; Aguayo, F.; Muñoz, J.; Bleak, T. Endocrine Disruptors from the Environment Affecting Breast Cancer. *Oncol. Lett.* **2020**, doi:10.3892/ol.2020.11566.
21. Koual, M.; Tomkiewicz, C.; Cano-Sancho, G.; Antignac, J.-P.; Bats, A.-S.; Coumoul, X. Environmental Chemicals, Breast Cancer Progression and Drug Resistance. *Environ. Health* **2020**, *19*, 117, doi:10.1186/s12940-020-00670-2.
22. Turner, N.C.; Kingston, B.; Kilburn, L.S.; Kernaghan, S.; Wardley, A.M.; Macpherson, I.R.; Baird, R.D.; Roylance, R.; Stephens, P.; Oikonomidou, O.; et al. Circulating Tumour DNA Analysis to Direct Therapy in Advanced Breast Cancer (PlasmaMATCH): A Multicentre, Multicohort, Phase 2a, Platform Trial. *Lancet Oncol.* **2020**, *21*, 1296–1308, doi:10.1016/S1470-2045(20)30444-7.
23. Bortolini Silveira, A.; Bidard, F.-C.; Tanguy, M.-L.; Girard, E.; Trédan, O.; Dubot, C.; Jacot, W.; Goncalves, A.; Debled, M.; Levy, C.; et al. Multimodal Liquid Biopsy for Early Monitoring and Outcome Prediction of Chemotherapy in Metastatic Breast Cancer. *Npj Breast Cancer* **2021**, *7*, 115, doi:10.1038/s41523-021-00319-4.
24. Messinis, I.E.; Messini, C.I.; Dafopoulos, K. Novel Aspects of the Endocrinology of the Menstrual Cycle. *Reprod. Biomed. Online* **2014**, *28*, 714–722, doi:10.1016/j.rbmo.2014.02.003.
25. Brisken, C. Progesterone Signalling in Breast Cancer: A Neglected Hormone Coming into the Limelight. *Nat. Rev. Cancer* **2013**, *13*, 385–396, doi:10.1038/nrc3518.
26. Gruber, C.J.; Tschugguel, W.; Schneeberger, C.; Huber, J.C. Production and Actions of Estrogens. *N. Engl. J. Med.* **2002**, *346*, 340–352, doi:10.1056/NEJMra000471.
27. Brisken, C.; O'Malley, B. Hormone Action in the Mammary Gland. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a003178–a003178, doi:10.1101/cshperspect.a003178.
28. Nagy, B.; Szekeres-Barthó, J.; Kovács, G.L.; Sulyok, E.; Farkas, B.; Várnagy, Á.; Vértes, V.; Kovács, K.; Bódis, J. Key to Life: Physiological Role and Clinical Implications of Progesterone. *Int. J. Mol. Sci.* **2021**, *22*, 11039, doi:10.3390/ijms222011039.
29. Rusidzé, M.; Adlanmérini, M.; Chantalat, E.; Raymond-Letron, I.; Cayre, S.; Arnal, J.-F.; Deugnier, M.-A.; Lenfant, F. Estrogen Receptor-α Signaling in Post-Natal Mammary Development and Breast Cancers. *Cell. Mol. Life Sci.* **2021**, *78*, 5681–5705, doi:10.1007/s00018-021-03860-4.
30. Hanker, A.B.; Sudhan, D.R.; Arteaga, C.L. Overcoming Endocrine Resistance in Breast Cancer. *Cancer Cell* **2020**, *37*, 496–513, doi:10.1016/j.ccell.2020.03.009.
31. Kharb, R.; Haider, K.; Neha, K.; Yar, M.S. Aromatase Inhibitors: Role in Postmenopausal Breast Cancer. *Arch. Pharm. (Weinheim)* **2020**, *353*, 2000081, doi:10.1002/ardp.202000081.
32. Patel, H.K.; Bihani, T. Selective Estrogen Receptor Modulators (SERMs) and Selective Estrogen Receptor Degraders (SERDs) in Cancer Treatment. *Pharmacol. Ther.* **2018**, *186*, 1–24, doi:10.1016/j.pharmthera.2017.12.012.
33. Wang, L.; Sharma, A. The Quest for Orally Available Selective Estrogen Receptor Degraders (SERDs). *ChemMedChem* **2020**, *15*, 2072–2097, doi:10.1002/cmdc.202000473.
34. Brufsky, A.M.; Dickler, M.N. Estrogen Receptor-Positive Breast Cancer: Exploiting Signaling Pathways Implicated in Endocrine Resistance. *The Oncologist* **2018**, *23*, 528–539, doi:10.1634/theoncologist.2017-0423.
35. Herrera-Martínez, M.; Orozco-Samperio, E.; Montaño, S.; Ariza-Ortega, J.A.; Flores-García, Y.; López-Contreras, L. Vorinostat as Potential Antiparasitic Drug. *Eur. Rev. Med. Pharmacol. Sci.* **2020**, *24*, 7521–7528, doi:10.26355/eurrev_202007_21909.
36. Lin, X.; Xiang, H.; Luo, G. Targeting Estrogen Receptor α for Degradation with PROTACs: A Promising Approach to Overcome Endocrine Resistance. *Eur. J. Med. Chem.* **2020**, *206*, 112689, doi:10.1016/j.ejmech.2020.112689.

37. Abdel-Hafiz, H.A.; Horwitz, K.B. Role of Epigenetic Modifications in Luminal Breast Cancer. *Epigenomics* **2015**, *7*, 847–862, doi:10.2217/epi.15.10.
38. Wawruszak, A.; Borkiewicz, L.; Okon, E.; Kukula-Koch, W.; Afshan, S.; Halasa, M. Vorinostat (SAHA) and Breast Cancer: An Overview. *Cancers* **2021**, *13*, 4700, doi:10.3390/cancers13184700.
39. Mal, R.; Magner, A.; David, J.; Datta, J.; Vallabhaneni, M.; Kassem, M.; Manouchehri, J.; Willingham, N.; Stover, D.; Vandeusen, J.; et al. Estrogen Receptor Beta (ERβ): A Ligand Activated Tumor Suppressor. *Front. Oncol.* **2020**, *10*, 587386, doi:10.3389/fonc.2020.587386.
40. Khalid, A.B.; Krum, S.A. Estrogen Receptors Alpha and Beta in Bone. *Bone* **2016**, *87*, 130–135, doi:10.1016/j.bone.2016.03.016.
41. Uddin, Md.S.; Rahman, Md.M.; Jakaria, Md.; Rahman, Md.S.; Hossain, Md.S.; Islam, A.; Ahmed, M.; Mathew, B.; Omar, U.M.; Barreto, G.E.; et al. Estrogen Signaling in Alzheimer's Disease: Molecular Insights and Therapeutic Targets for Alzheimer's Dementia. *Mol. Neurobiol.* **2020**, *57*, 2654–2670, doi:10.1007/s12035-020-01911-8.
42. Uenoyama, Y.; Inoue, N.; Nakamura, S.; Tsukamura, H. Kisspeptin Neurons and Estrogen–Estrogen Receptor α Signaling: Unraveling the Mystery of Steroid Feedback System Regulating Mammalian Reproduction. *Int. J. Mol. Sci.* **2021**, *22*, 9229, doi:10.3390/ijms22179229.
43. Bernasochi, G.B.; Bell, J.R.; Simpson, E.R.; Delbridge, L.M.D.; Boon, W.C. Impact of Estrogens on the Regulation of White, Beige, and Brown Adipose Tissue Depots. *Compr. Physiol.* **2019**, *9*, 457–475, doi:10.1002/cphy.c180009.
44. Aryan, L.; Younessi, D.; Zargari, M.; Banerjee, S.; Agopian, J.; Rahman, S.; Borna, R.; Ruffenach, G.; Umar, S.; Eghbali, M. The Role of Estrogen Receptors in Cardiovascular Disease. *Int. J. Mol. Sci.* **2020**, *21*, 4314, doi:10.3390/ijms21124314.
45. Ranganathan, P.; Nadig, N.; Nambiar, S. Non-Canonical Estrogen Signaling in Endocrine Resistance. *Front. Endocrinol.* **2019**, *10*, 708, doi:10.3389/fendo.2019.00708.
46. Zhou, Y.; Liu, X. The Role of Estrogen Receptor Beta in Breast Cancer. *Biomark. Res.* **2020**, *8*, 39, doi:10.1186/s40364-020-00223-2.
47. Molina, L.; Figueroa, C.D.; Bhoola, K.D.; Ehrenfeld, P. GPER-1/GPR30 a Novel Estrogen Receptor Sited in the Cell Membrane: Therapeutic Coupling to Breast Cancer. *Expert Opin. Ther. Targets* **2017**, *21*, 755–766, doi:10.1080/14728222.2017.1350264.
48. Arnal, J.-F.; Lenfant, F.; Metivier, R.; Flouriot, G.; Henrion, D.; Adlanmerini, M.; Fontaine, C.; Gourdy, P.; Chambon, P.; Katzenellenbogen, B.; et al. Membrane and Nuclear Estrogen Receptor Alpha Actions: From Tissue Specificity to Medical Implications. *Physiol. Rev.* **2017**, *97*, 1045–1087, doi:10.1152/physrev.00024.2016.
49. Le Romancer, M.; Poulard, C.; Cohen, P.; Sentis, S.; Renoir, J.-M.; Corbo, L. Cracking the Estrogen Receptor's Posttranslational Code in Breast Tumors. *Endocr. Rev.* **2011**, *32*, 597–622, doi:10.1210/er.2010-0016.
50. Jeffreys, S.A.; Powter, B.; Balakrishnar, B.; Mok, K.; Soon, P.; Franken, A.; Neubauer, H.; de Souza, P.; Becker, T.M. Endocrine Resistance in Breast Cancer: The Role of Estrogen Receptor Stability. *Cells* **2020**, *9*, 2077, doi:10.3390/cells9092077.
51. Clusan, L.; Le Goff, P.; Flouriot, G.; Pakdel, F. A Closer Look at Estrogen Receptor Mutations in Breast Cancer and Their Implications for Estrogen and Antiestrogen Responses. *Int. J. Mol. Sci.* **2021**, *22*, 756, doi:10.3390/ijms22020756.
52. Manavathi, B.; Dey, O.; Gajulapalli, V.N.R.; Bhatia, R.S.; Bugide, S.; Kumar, R. Derailed Estrogen Signaling and Breast Cancer: An Authentic Couple. *Endocr. Rev.* **2013**, *34*, 1–32, doi:10.1210/er.2011-1057.
53. Farcas, A.M.; Nagarajan, S.; Cosulich, S.; Carroll, J.S. Genome-Wide Estrogen Receptor Activity in Breast Cancer. *Endocrinology* **2021**, *162*, bqaa224, doi:10.1210/endocr/bqaa224.
54. Altwegg, K.A.; Vadlamudi, R.K. Role of Estrogen Receptor Coregulators in Endocrine Resistant Breast Cancer. *Explor. Target. Anti-Tumor Ther.* **2021**, 385–400, doi:10.37349/etat.2021.00052.
55. Mayayo-Peralta, I.; Prekovic, S.; Zwart, W. Estrogen Receptor on the Move: Cistromic Plasticity and Its Implications in Breast Cancer. *Mol. Aspects Med.* **2021**, *78*, 100939, doi:10.1016/j.mam.2020.100939.

56. Welboren, W.-J.; Sweep, F.C.G.J.; Span, P.N.; Stunnenberg, H.G. Genomic Actions of Estrogen Receptor α: What Are the Targets and How Are They Regulated? *Endocr. Relat. Cancer* **2009**, *16*, 1073–1089, doi:10.1677/ERC-09-0086.

57. Jehanno, C.; Fernandez-Calero, T.; Habauzit, D.; Avner, S.; Percevault, F.; Jullion, E.; Le Goff, P.; Coissieux, M.M.; Muenst, S.; Marin, M.; et al. Nuclear Accumulation of MKL1 in Luminal Breast Cancer Cells Impairs Genomic Activity of ERα and Is Associated with Endocrine Resistance. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **2020**, *1863*, 194507, doi:10.1016/j.bbagrm.2020.194507.

58. Mérot, Y.; Métivier, R.; Penot, G.; Manu, D.; Saligaut, C.; Gannon, F.; Pakdel, F.; Kah, O.; Flouriot, G. The Relative Contribution Exerted by AF-1 and AF-2 Transactivation Functions in Estrogen Receptor α Transcriptional Activity Depends upon the Differentiation Stage of the Cell. *J. Biol. Chem.* **2004**, *279*, 26184–26191, doi:10.1074/jbc.M402148200.

59. Huet, G.; Mérot, Y.; Le Dily, F.; Kern, L.; Ferrière, F.; Saligaut, C.; Boujrad, N.; Pakdel, F.; Métivier, R.; Flouriot, G. Loss of E-Cadherin-Mediated Cell Contacts Reduces Estrogen Receptor Alpha (ERα) Transcriptional Efficiency by Affecting the Respective Contribution Exerted by AF1 and AF2 Transactivation Functions. *Biochem. Biophys. Res. Commun.* **2008**, *365*, 304–309, doi:10.1016/j.bbrc.2007.10.178.

60. Boonyaratanakornkit, V.; Hamilton, N.; Márquez-Garbán, D.C.; Pateetin, P.; McGowan, E.M.; Pietras, R.J. Extranuclear Signaling by Sex Steroid Receptors and Clinical Implications in Breast Cancer. *Mol. Cell. Endocrinol.* **2018**, *466*, 51–72, doi:10.1016/j.mce.2017.11.010.

61. Madak-Erdogan, Z.; Kieser, K.J.; Kim, S.H.; Komm, B.; Katzenellenbogen, J.A.; Katzenellenbogen, B.S. Nuclear and Extranuclear Pathway Inputs in the Regulation of Global Gene Expression by Estrogen Receptors. *Mol. Endocrinol.* **2008**, *22*, 2116–2127, doi:10.1210/me.2008-0059.

62. Dawson, M.A. The Cancer Epigenome: Concepts, Challenges, and Therapeutic Opportunities. *Science* **2017**, *355*, 1147–1152, doi:10.1126/science.aam7304.

63. Hazra, A.; Bose, P.; Sunita, P.; Pattanayak, S.P. Molecular Epigenetic Dynamics in Breast Carcinogenesis. *Arch. Pharm. Res.* **2021**, *44*, 741–763, doi:10.1007/s12272-021-01348-0.

64. Lee, L.J.; Papadopoli, D.; Jewer, M.; del Rincon, S.; Topisirovic, I.; Lawrence, M.G.; Postovit, L.-M. Cancer Plasticity: The Role of MRNA Translation. *Trends Cancer* **2021**, *7*, 134–145, doi:10.1016/j.trecan.2020.09.005.

65. Silvera, D.; Formenti, S.C.; Schneider, R.J. Translational Control in Cancer. *Nat. Rev. Cancer* **2010**, *10*, 254–266, doi:10.1038/nrc2824.

66. Fabbri, L.; Chakraborty, A.; Robert, C.; Vagner, S. The Plasticity of MRNA Translation during Cancer Progression and Therapy Resistance. *Nat. Rev. Cancer* **2021**, *21*, 558–577, doi:10.1038/s41568-021-00380-y.

67. Hinnebusch, A.G. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu. Rev. Biochem.* **2014**, *83*, 779–812, doi:10.1146/annurev-biochem-060713-035802.

68. Knight, J.R.P.; Garland, G.; Pöyry, T.; Mead, E.; Vlahov, N.; Sfakianos, A.; Grosso, S.; De-Lima-Hedayioglu, F.; Mallucci, G.R.; von der Haar, T.; et al. Control of Translation Elongation in Health and Disease. *Dis. Model. Mech.* **2020**, *13*, dmm043208, doi:10.1242/dmm.043208.

69. Hellen, C.U.T. Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a032656, doi:10.1101/cshperspect.a032656.

70. Truitt, M.L.; Ruggero, D. New Frontiers in Translational Control of the Cancer Genome. *Nat. Rev. Cancer* **2016**, *16*, 288–304, doi:10.1038/nrc.2016.27.

71. Collart, M.A.; Weiss, B. Ribosome Pausing, a Dangerous Necessity for Co-Translational Events. *Nucleic Acids Res.* **2020**, *48*, 1043–1055, doi:10.1093/nar/gkz763.

72. Thommen, M.; Holtkamp, W.; Rodnina, M.V. Co-Translational Protein Folding: Progress and Methods. *Curr. Opin. Struct. Biol.* **2017**, *42*, 83–89, doi:10.1016/j.sbi.2016.11.020.

73. Liutkute, M.; Samatova, E.; Rodnina, M.V. Cotranslational Folding of Proteins on the Ribosome. *Biomolecules* **2020**, *10*, 97, doi:10.3390/biom10010097.

74. Cassaignau, A.M.E.; Włodarski, T.; Chan, S.H.S.; Woodburn, L.F.; Bukvin, I.V.; Streit, J.O.; Cabrita, L.D.; Waudby, C.A.; Christodoulou, J. Interactions between Nascent Proteins and the Ribosome Surface Inhibit Co-Translational Folding. *Nat. Chem.* **2021**, doi:10.1038/s41557-021-00796-x.

75. Genuth, N.R.; Barna, M. The Discovery of Ribosome Heterogeneity and Its Implications for Gene Regulation and Organismal Life. *Mol. Cell* **2018**, *71*, 364–374, doi:10.1016/j.molcel.2018.07.018.

76. Wilson, D.N.; Doudna Cate, J.H. The Structure and Function of the Eukaryotic Ribosome. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a011536–a011536, doi:10.1101/cshperspect.a011536.

77. Narla, A.; Ebert, B.L. Ribosomopathies: Human Disorders of Ribosome Dysfunction. *Blood* **2010**, *115*, 3196–3205, doi:10.1182/blood-2009-10-178129.

78. Venturi, G.; Montanaro, L. How Altered Ribosome Production Can Cause or Contribute to Human Disease: The Spectrum of Ribosomopathies. *Cells* **2020**, *9*, 2300, doi:10.3390/cells9102300.

79. Yusupova, G.; Yusupov, M. Crystal Structure of Eukaryotic Ribosome and Its Complexes with Inhibitors. *Philos. Trans. R. Soc. B Biol. Sci.* **2017**, *372*, 20160184, doi:10.1098/rstb.2016.0184.

80. Slavov, N.; Semrau, S.; Airoldi, E.; Budnik, B.; van Oudenaarden, A. Differential Stoichiometry among Core Ribosomal Proteins. *Cell Rep.* **2015**, *13*, 865–873, doi:10.1016/j.celrep.2015.09.056.

81. Imami, K.; Milek, M.; Bogdanow, B.; Yasuda, T.; Kastelic, N.; Zauber, H.; Ishihama, Y.; Landthaler, M.; Selbach, M. Phosphorylation of the Ribosomal Protein RPL12/UL11 Affects Translation during Mitosis. *Mol. Cell* **2018**, *72*, 84-98.e9, doi:10.1016/j.molcel.2018.08.019.

82. Simsek, D.; Barna, M. An Emerging Role for the Ribosome as a Nexus for Post-Translational Modifications. *Curr. Opin. Cell Biol.* **2017**, *45*, 92–101, doi:10.1016/j.ceb.2017.02.010.

83. Kondrashov, N.; Pusic, A.; Stumpf, C.R.; Shimizu, K.; Hsieh, A.C.; Xue, S.; Ishijima, J.; Shiroishi, T.; Barna, M. Ribosome-Mediated Specificity in Hox MRNA Translation and Vertebrate Tissue Patterning. *Cell* **2011**, *145*, 383–397, doi:10.1016/j.cell.2011.03.028.

84. Ferretti, M.B.; Ghalei, H.; Ward, E.A.; Potts, E.L.; Karbstein, K. Rps26 Directs MRNA-Specific Translation by Recognition of Kozak Sequence Elements. *Nat. Struct. Mol. Biol.* **2017**, *24*, 700–707, doi:10.1038/nsmb.3442.

85. Shi, Z.; Fujii, K.; Kovary, K.M.; Genuth, N.R.; Röst, H.L.; Teruel, M.N.; Barna, M. Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of MRNAs Genome-Wide. *Mol. Cell* **2017**, *67*, 71-83.e7, doi:10.1016/j.molcel.2017.05.021.

86. Salim, D.; Gerton, J.L. Ribosomal DNA Instability and Genome Adaptability. *Chromosome Res.* **2019**, *27*, 73–87, doi:10.1007/s10577-018-9599-7.

87. Natchiar, S.K.; Myasnikov, A.G.; Kratzat, H.; Hazemann, I.; Klaholz, B.P. Visualization of Chemical Modifications in the Human 80S Ribosome Structure. *Nature* **2017**, *551*, 472–477, doi:10.1038/nature24482.

88. Sharma, S.; Lafontaine, D.L.J. 'View From A Bridge': A New Perspective on Eukaryotic RRNA Base Modification. *Trends Biochem. Sci.* **2015**, *40*, 560–575, doi:10.1016/j.tibs.2015.07.008.

89. Monaco, P.; Marcel, V.; Diaz, J.-J.; Catez, F. 2'-O-Methylation of Ribosomal RNA: Towards an Epitranscriptomic Control of Translation? *Biomolecules* **2018**, *8*, 106, doi:10.3390/biom8040106.

90. Penzo, M.; Montanaro, L. Turning Uridines around: Role of RRNA Pseudouridylation in Ribosome Biogenesis and Ribosomal Function. *Biomolecules* **2018**, *8*, 38, doi:10.3390/biom8020038.

91. Montanaro, L.; Brigotti, M.; Clohessy, J.; Barbieri, S.; Ceccarelli, C.; Santini, D.; Taffurelli, M.; Calienni, M.; Teruya-Feldstein, J.; Trerè, D.; et al. Dyskerin Expression Influences the Level of Ribosomal RNA Pseudo-Uridylation and Telomerase RNA Component in Human Breast Cancer. *J. Pathol.* **2006**, *210*, 10–18, doi:10.1002/path.2023.

92. Su, H.; Xu, T.; Ganapathy, S.; Shadfan, M.; Long, M.; Huang, T.H.-M.; Thompson, I.; Yuan, Z.-M. Elevated SnoRNA Biogenesis Is Essential in Breast Cancer. *Oncogene* **2014**, *33*, 1348–1358, doi:10.1038/onc.2013.89.

93. Fu, J.; Dang, Y.; Counter, C.; Liu, Y. Codon Usage Regulates Human KRAS Expression at Both Transcriptional and Translational Levels. *J. Biol. Chem.* **2018**, *293*, 17929–17940, doi:10.1074/jbc.RA118.004908.

94. Nieuwkoop, T.; Finger-Bou, M.; van der Oost, J.; Claassens, N.J. The Ongoing Quest to Crack the Genetic Code for Protein Production. *Mol. Cell* **2020**, *80*, 193–209, doi:10.1016/j.molcel.2020.09.014.

95. Hussmann, J.A.; Patchett, S.; Johnson, A.; Sawyer, S.; Press, W.H. Understanding Biases in Ribosome Profiling Experiments Reveals Signatures of Translation Dynamics in Yeast. *PLOS Genet.* **2015**, *11*, e1005732, doi:10.1371/journal.pgen.1005732.

96. Weinberg, D.E.; Shah, P.; Eichhorn, S.W.; Hussmann, J.A.; Plotkin, J.B.; Bartel, D.P. Improved Ribosome-Footprint and MRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Rep.* **2016**, *14*, 1787–1799, doi:10.1016/j.celrep.2016.01.043.

97. Gamble, C.E.; Brule, C.E.; Dean, K.M.; Fields, S.; Grayhack, E.J. Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. *Cell* **2016**, *166*, 679–690, doi:10.1016/j.cell.2016.05.070.

98. Leininger, S.E.; Rodriguez, J.; Vu, Q.V.; Jiang, Y.; Li, M.S.; Deutsch, C.; O'Brien, E.P. Ribosome Elongation Kinetics of Consecutively Charged Residues Are Coupled to Electrostatic Force. *Biochemistry* **2021**, acs.biochem.1c00507, doi:10.1021/acs.biochem.1c00507.

99. Gingold, H.; Tehler, D.; Christoffersen, N.R.; Nielsen, M.M.; Asmar, F.; Kooistra, S.M.; Christophersen, N.S.; Christensen, L.L.; Borre, M.; Sørensen, K.D.; et al. A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell* **2014**, *158*, 1281–1292, doi:10.1016/j.cell.2014.08.011.

100. Santos, M.; Fidalgo, A.; Varanda, A.S.; Oliveira, C.; Santos, M.A.S. TRNA Deregulation and Its Consequences in Cancer. *Trends Mol. Med.* **2019**, *25*, 853–865, doi:10.1016/j.molmed.2019.05.011.

101. Hanson, G.; Coller, J. Codon Optimality, Bias and Usage in Translation and MRNA Decay. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 20–30, doi:10.1038/nrm.2017.91.

102. Marín, M.; Fernández-Calero, T.; Ehrlich, R. Protein Folding and TRNA Biology. *Biophys. Rev.* **2017**, *9*, 573–588, doi:10.1007/s12551-017-0322-2.

103. Liu, Y. A Code within the Genetic Code: Codon Usage Regulates Co-Translational Protein Folding. *Cell Commun. Signal.* **2020**, *18*, 145, doi:10.1186/s12964-020-00642-6.

104. Yu, C.-H.; Dang, Y.; Zhou, Z.; Wu, C.; Zhao, F.; Sachs, M.S.; Liu, Y. Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-Translational Protein Folding. *Mol. Cell* **2015**, *59*, 744–754, doi:10.1016/j.molcel.2015.07.018.

105. Buhr, F.; Jha, S.; Thommen, M.; Mittelstaet, J.; Kutz, F.; Schwalbe, H.; Rodnina, M.V.; Komar, A.A. Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations. *Mol. Cell* **2016**, *61*, 341–351, doi:10.1016/j.molcel.2016.01.008.

106. Zhou, M.; Wang, T.; Fu, J.; Xiao, G.; Liu, Y. Nonoptimal Codon Usage Influences Protein Structure in Intrinsically Disordered Regions. *Mol. Microbiol.* **2015**, *97*, 974–987, doi:10.1111/mmi.13079.

107. Oldfield, C.J.; Peng, Z.; Uversky, V.N.; Kurgan, L. Codon Selection Reduces GC Content Bias in Nucleic Acids Encoding for Intrinsically Disordered Proteins. *Cell. Mol. Life Sci.* **2020**, *77*, 149–160, doi:10.1007/s00018-019-03166-6.

108. Peng, Y.; Cao, S.; Kiselar, J.; Xiao, X.; Du, Z.; Hsieh, A.; Ko, S.; Chen, Y.; Agrawal, P.; Zheng, W.; et al. A Metastable Contact and Structural Disorder in the Estrogen Receptor Transactivation Domain. *Structure* **2019**, *27*, 229-240.e4, doi:10.1016/j.str.2018.10.026.

109. Fernández-Calero, T.; Astrada, S.; Alberti, Á.; Horjales, S.; Arnal, J.F.; Rovira, C.; Bollati-Fogolín, M.; Flouriot, G.; Marin, M. The Transcriptional Activities and Cellular Localization of the Human Estrogen Receptor Alpha Are Affected by the Synonymous Ala87 Mutation. *J. Steroid Biochem. Mol. Biol.* **2014**, *143*, 99–104, doi:10.1016/j.jsbmb.2014.02.016.

110. Zhang, G.; Ignatova, Z. Generic Algorithm to Predict the Speed of Translational Elongation: Implications for Protein Biogenesis. *PLoS ONE* **2009**, *4*, e5036, doi:10.1371/journal.pone.0005036.

111. Bitran, A.; Jacobs, W.M.; Zhai, X.; Shakhnovich, E. Cotranslational Folding Allows Misfolding-Prone Proteins to Circumvent Deep Kinetic Traps. *Proc. Natl. Acad. Sci.* **2020**, *117*, 1485–1495, doi:10.1073/pnas.1913207117.

112. Pechmann, S.; Frydman, J. Evolutionary Conservation of Codon Optimality Reveals Hidden Signatures of Cotranslational Folding. *Nat. Struct. Mol. Biol.* **2013**, *20*, 237–243, doi:10.1038/nsmb.2466.

113. Jacobs, W.M.; Shakhnovich, E.I. Evidence of Evolutionary Selection for Cotranslational Folding. *Proc. Natl. Acad. Sci.* **2017**, *114*, 11434–11439, doi:10.1073/pnas.1705772114.

114. Chaney, J.L.; Steele, A.; Carmichael, R.; Rodriguez, A.; Specht, A.T.; Ngo, K.; Li, J.; Emrich, S.; Clark, P.L. Widespread Position-Specific Conservation of Synonymous Rare Codons within Coding Sequences. *PLOS Comput. Biol.* **2017**, *13*, e1005531, doi:10.1371/journal.pcbi.1005531.

115. Kirchner, S.; Cai, Z.; Rauscher, R.; Kastelic, N.; Anding, M.; Czech, A.; Kleizen, B.; Ostedgaard, L.S.; Braakman, I.; Sheppard, D.N.; et al. Alteration of Protein Function by a Silent Polymorphism Linked to TRNA Abundance. *PLOS Biol.* **2017**, *15*, e2000779, doi:10.1371/journal.pbio.2000779.

116. Alexaki, A.; Hettiarachchi, G.K.; Athey, J.C.; Katneni, U.K.; Simhadri, V.; Hamasaki-Katagiri, N.; Nanavaty, P.; Lin, B.; Takeda, K.; Freedberg, D.; et al. Effects of Codon Optimization on Coagulation Factor IX Translation and Structure: Implications for Protein and Gene Therapies. *Sci. Rep.* **2019**, *9*, 15449, doi:10.1038/s41598-019-51984-2.

117. Sauna, Z.E.; Kimchi-Sarfaty, C. Understanding the Contribution of Synonymous Mutations to Human Disease. *Nat. Rev. Genet.* **2011**, *12*, 683–691, doi:10.1038/nrg3051.

118. Bartoszewski, R.A.; Jablonsky, M.; Bartoszewska, S.; Stevenson, L.; Dai, Q.; Kappes, J.; Collawn, J.F.; Bebok, Z. A Synonymous Single Nucleotide Polymorphism in ΔF508 CFTR Alters the Secondary Structure of the MRNA and the Expression of the Mutant Protein. *J. Biol. Chem.* **2010**, *285*, 28741–28748, doi:10.1074/jbc.M110.154575.

119. Oliver, K.E.; Rauscher, R.; Mijnders, M.; Wang, W.; Wolpert, M.J.; Maya, J.; Sabusap, C.M.; Kesterson, R.A.; Kirk, K.L.; Rab, A.; et al. Slowing Ribosome Velocity Restores Folding and Function of Mutant CFTR. *J. Clin. Invest.* **2019**, *129*, 5236–5253, doi:10.1172/JCI124282.

120. Meyer, D.; Kames, J.; Bar, H.; Komar, A.A.; Alexaki, A.; Ibla, J.; Hunt, R.C.; Santana-Quintero, L.V.; Golikov, A.; DiCuccio, M.; et al. Distinct Signatures of Codon and Codon Pair Usage in 32 Primary Tumor Types in the Novel Database CancerCoCoPUTs for Cancer-Specific Codon Usage. *Genome Med.* **2021**, *13*, 122, doi:10.1186/s13073-021-00935-6.

121. Goodarzi, H.; Nguyen, H.C.B.; Zhang, S.; Dill, B.D.; Molina, H.; Tavazoie, S.F. Modulated Expression of Specific TRNAs Drives Gene Expression and Cancer Progression. *Cell* **2016**, *165*, 1416–1427, doi:10.1016/j.cell.2016.05.046.

122. Guimaraes, J.C.; Mittal, N.; Gnann, A.; Jedlinski, D.; Riba, A.; Buczak, K.; Schmidt, A.; Zavolan, M. A Rare Codon-Based Translational Program of Cell Proliferation. *Genome Biol.* **2020**, *21*, 44, doi:10.1186/s13059-020-1943-5.

123. Pouyet, F.; Mouchiroud, D.; Duret, L.; Sémon, M. Recombination, Meiotic Expression and Human Codon Usage. *eLife* **2017**, *6*, e27344, doi:10.7554/eLife.27344.

124. Rudolph, K.L.M.; Schmitt, B.M.; Villar, D.; White, R.J.; Marioni, J.C.; Kutter, C.; Odom, D.T. Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States. *PLOS Genet.* **2016**, *12*, e1006024, doi:10.1371/journal.pgen.1006024.

125. Barna, M.; Karbstein, K.; Tollervey, D.; Ruggero, D.; Brar, G.; Greer, E.L.; Dinman, J.D. The Promises and Pitfalls of Specialized Ribosomes. *Mol. Cell* **2022**, *82*, 2179–2184, doi:10.1016/j.molcel.2022.05.035.

126. Hia, F.; Yang, S.F.; Shichino, Y.; Yoshinaga, M.; Murakawa, Y.; Vandenbon, A.; Fukao, A.; Fujiwara, T.; Landthaler, M.; Natsume, T.; et al. Codon Bias Confers Stability to Human MRNAs. *EMBO Rep.* **2019**, *20*, doi:10.15252/embr.201948220.

127. Kudla, G.; Lipinski, L.; Caffin, F.; Helwak, A.; Zylicz, M. High Guanine and Cytosine Content Increases MRNA Levels in Mammalian Cells. *PLoS Biol.* **2006**, *4*, e180, doi:10.1371/journal.pbio.0040180.

128. Belin, S.; Beghin, A.; Solano-Gonzàlez, E.; Bezin, L.; Brunet-Manquat, S.; Textoris, J.; Prats, A.-C.; Mertani, H.C.; Dumontet, C.; Diaz, J.-J. Dysregulation of Ribosome Biogenesis and Translational Capacity Is Associated with Tumor Progression of Human Breast Cancer Cells. *PLoS ONE* **2009**, *4*, e7147, doi:10.1371/journal.pone.0007147.

129. Marcel, V.; Ghayad, S.E.; Belin, S.; Therizols, G.; Morel, A.-P.; Solano-Gonzàlez, E.; Vendrell, J.A.; Hacot, S.; Mertani, H.C.; Albaret, M.A.; et al. P53 Acts as a Safeguard of Translational Control by Regulating Fibrillarin and RRNA Methylation in Cancer. *Cancer Cell* **2013**, *24*, 318–330, doi:10.1016/j.ccr.2013.08.013.

130. Erales, J.; Marchand, V.; Panthu, B.; Gillot, S.; Belin, S.; Ghayad, S.E.; Garcia, M.; Laforêts, F.; Marcel, V.; Baudin-Baillieu, A.; et al. Evidence for RRNA 2'-O-Methylation Plasticity: Control of Intrinsic Translational Capabilities of Human Ribosomes. *Proc. Natl. Acad. Sci.* **2017**, *114*, 12934–12939, doi:10.1073/pnas.1707674114.

131. Penzo, M.; Rocchi, L.; Brugiere, S.; Carnicelli, D.; Onofrillo, C.; Couté, Y.; Brigotti, M.; Montanaro, L. Human Ribosomes from Cells with Reduced Dyskerin Levels Are Intrinsically Altered in Translation. *FASEB J.* **2015**, *29*, 3472–3482, doi:10.1096/fj.15-270991.

132. Rocchi, L.; Pacilli, A.; Sethi, R.; Penzo, M.; Schneider, R.J.; Treré, D.; Brigotti, M.; Montanaro, L. Dyskerin Depletion Increases VEGF MRNA Internal Ribosome Entry Site-Mediated Translation. *Nucleic Acids Res.* **2013**, *41*, 8308–8318, doi:10.1093/nar/gkt587.

133. Montanaro, L.; Calienni, M.; Bertoni, S.; Rocchi, L.; Sansone, P.; Storci, G.; Santini, D.; Ceccarelli, C.; Taffurelli, M.; Carnicelli, D.; et al. Novel Dyskerin-Mediated Mechanism of P53 Inactivation through Defective MRNA Translation. *Cancer Res.* **2010**, *70*, 4767–4777, doi:10.1158/0008-5472.CAN-09-4024.

134. King, T.H.; Liu, B.; McCully, R.R.; Fournier, M.J. Ribosome Structure and Activity Are Altered in Cells Lacking SnoRNPs That Form Pseudouridines in the Peptidyl Transferase Center. *Mol. Cell* **2003**, *11*, 425–435, doi:10.1016/S1097-2765(03)00040-6.

135. Panthu, B.; Décimo, D.; Balvay, L.; Ohlmann, T. In Vitro Translation in a Hybrid Cell Free Lysate with Exogenous Cellular Ribosomes. *Biochem. J.* **2015**, *467*, 387–398, doi:10.1042/BJ20141498.

136. Gupta, V.; Warner, J.R. Ribosome-Omics of the Human Ribosome. *RNA* **2014**, *20*, 1004–1013, doi:10.1261/rna.043653.113.

137. Guimaraes, J.C.; Zavolan, M. Patterns of Ribosomal Protein Expression Specify Normal and Malignant Human Cells. *Genome Biol.* **2016**, *17*, 236, doi:10.1186/s13059-016-1104-z.

138. Zhao, Y.; Tan, M.; Liu, X.; Xiong, X.; Sun, Y. Inactivation of Ribosomal Protein S27-like Confers Radiosensitivity via the Mdm2-P53 and Mdm2–MRN–ATM Axes. *Cell Death Dis.* **2018**, *9*, 145, doi:10.1038/s41419-017-0192-3.

139. Huang, C.-J.; Yang, S.-H.; Lee, C.-L.; Cheng, Y.-C.; Tai, S.-Y.; Chien, C.-C. Ribosomal Protein S27-Like in Colorectal Cancer: A Candidate for Predicting Prognoses. *PLoS ONE* **2013**, *8*, e67043, doi:10.1371/journal.pone.0067043.

140. Rao, S.; Peri, S.; Hoffmann, J.; Cai, K.Q.; Harris, B.; Rhodes, M.; Connolly, D.C.; Testa, J.R.; Wiest, D.L. RPL22L1 Induction in Colorectal Cancer Is Associated with Poor Prognosis and 5-FU Resistance. *PLOS ONE* **2019**, *14*, e0222392, doi:10.1371/journal.pone.0222392.

141. Wu, N.; Wei, J.; Wang, Y.; Yan, J.; Qin, Y.; Tong, D.; Pang, B.; Sun, D.; Sun, H.; Yu, Y.; et al. Ribosomal L22-Like1 (RPL22L1) Promotes Ovarian Cancer Metastasis by Inducing Epithelial-to-Mesenchymal Transition. *PLOS ONE* **2015**, *10*, e0143659, doi:10.1371/journal.pone.0143659.

142. Liang, Z.; Mou, Q.; Pan, Z.; Zhang, Q.; Gao, G.; Cao, Y.; Gao, Z.; Pan, Z.; Feng, W. Identification of Candidate Diagnostic and Prognostic Biomarkers for Human Prostate Cancer: RPL22L1 and RPS21. *Med. Oncol.* **2019**, *36*, 56, doi:10.1007/s12032-019-1283-z.

143. Zhang, D.; Zhou, Y.; Ma, Y.; Jiang, P.; Lv, H.; Liu, S.; Mu, Y.; Zhou, C.; Xiao, S.; Ji, G.; et al. Ribosomal Protein L22-Like1 (RPL22L1) Mediates Sorafenib Sensitivity via ERK in Hepatocellular Carcinoma. *Cell Death Discov.* **2022**, *8*, 365, doi:10.1038/s41420-022-01153-8.

144. Hountondji, C.; Bulygin, K.; Créchet, J.-B.; Woisard, A.; Tuffery, P.; Nakayama, J.; Frolova, L.; Knud H, N.; Karpova, G.; Baouz, S. The CCA-End of P-TRNA Contacts Both the Human RPL36AL and the A-Site Bound Translation Termination Factor ERF1 at the Peptidyl Transferase Center of the Human 80S Ribosome. *Open Biochem. J.* **2014**, *8*, 52–67, doi:10.2174/1874091X01408010052.

145. Katneni, U.K.; Alexaki, A.; Hunt, R.C.; Hamasaki-Katagiri, N.; Hettiarachchi, G.K.; Kames, J.M.; McGill, J.R.; Holcomb, D.D.; Athey, J.C.; Lin, B.; et al. Structural, Functional, and Immunogenicity Implications of F9 Gene Recoding. *Blood Adv.* **2022**, *6*, 3932–3944, doi:10.1182/bloodadvances.2022007094.

Annexe

EXAMEN DES MUTATIONS DU RECEPTEUR AUX ŒSTROGENES DANS LE CANCER DU SEIN ET LEURS IMPLICATIONS DANS LA REPONSE AUX ŒSTROGENES ET ANTI-ŒSTROGENES

Le cancer du sein est le cancer le plus répandu chez la femme dans le monde. Plus de 70% des cas expriment le récepteur aux œstrogènes ERα, un facteur de transcription central qui stimule la prolifération des cellules cancéreuses mammaires, en général en présence d'œstrogènes. Tandis que la majorité des cas de cancer du sein ERα-positif (ERα+) répond initialement aux thérapies anti-œstrogéniques, un pourcentage élevé de patientes développe une résistance au traitement au cours du temps. Des formes mutées d'ERα résultant en l'activation constitutive du récepteur ont notamment été identifiées comme jouant un rôle dans ce phénomène de résistance. La majorité de ces mutations est retrouvée au niveau du domaine de liaison au ligand, où elles induisent des modifications de la conformation du récepteur, le stabilisant dans sa forme active, ne requérant plus alors de liaison par les œstrogènes pour permettre son activité. De telles mutations sont majoritairement retrouvées dans des cas de cancer métastatique, suite à de premières phases de traitements anti-œstrogéniques, où un environnement pauvre en œstrogènes exerce une pression de sélection permettant l'émergence de ces mutations. Cette découverte chez des stades métastatiques résistants de cancer du sein a alors justifié et permis le développement de nouveaux anti-œstrogènes, permettant de cibler ces formes mutées du récepteur. La combinaison d'anti-œstrogènes avec d'autres inhibiteurs de voies de signalisation spécifiques constitue également des traitements alternatifs pour améliorer la prise en charge des patientes en cas de cancer du sein ERα+ métastatique résistant. Dans cette revue, nous résumons les dernières avancées concernant l'implication particulière des mutations ponctuelles d'ERα dans la résistance hormonale. Nous discutons également des mutations synonymes du récepteur, et leur rapport avec le repliement co-traductionnel d'ERα. Bien qu'aucun consensus n'ait encore été atteint concernant l'association de certains polymorphismes d'ERα avec le risque de développer un cancer du sein, de récentes études mettent en évidence le rôle des mutations synonymes dans diverses pathologies. La composition en codons d'un ARNm peut en effet modifier sa vitesse de traduction, en fonction de son adéquation avec le stock d'ARNt par exemple, et résulter en des modifications du repliement co-traductionnel des protéines. ERα pourrait notamment être sensible à ce phénomène, qui mériterait d'être étudié au niveau structurel et fonctionnel pour déterminer sa potentielle implication dans le développement du cancer du sein et la résistance aux anti-œstrogènes.

International Journal of
Molecular Sciences

MDPI

Review

A Closer Look at Estrogen Receptor Mutations in Breast Cancer and Their Implications for Estrogen and Antiestrogen Responses

Léa Clusan, Pascale Le Goff, Gilles Flouriot and Farzad Pakdel *

Inserm, EHESP, Irset (Institut de Recherche en Santé, Environnement et Travail)-UMR_S1085, Rennes University, F-35000 Rennes, France; lea.clusan@univ-rennes1.fr (L.C.); pascale.le-goff@univ-rennes1.fr (P.L.G.); gilles.flouriot@univ-rennes1.fr (G.F.)
* Correspondence: farzad.pakdel@univ-rennes1.fr; Tel.: +33-(0)22-323-5132

Abstract: Breast cancer (BC) is the most common cancer among women worldwide. More than 70% of BC cases express estrogen receptor alpha (ERα), a central transcription factor that stimulates the proliferation of breast cancer cells, usually in the presence of estrogen. While most cases of ER-positive BC initially respond to antiestrogen therapies, a high percentage of cases develop resistance to treatment over time. The recent discovery of mutated forms of ERα that result in constitutively active forms of the receptor in the metastatic-resistance stage of BC has provided a strong rationale for the development of new antiestrogens. These molecules targeting clinically relevant ERα mutants and a combination with other pharmacological inhibitors of specific pathways may constitute alternative treatments to improve clinical practice in the fight against metastatic-resistant ER-positive BC. In this review, we summarize the latest advances regarding the particular involvement of point mutations of ERα in endocrine resistance. We also discuss the involvement of synonymous ERα mutations with respect to co-translational folding of the receptor and ribosome biogenesis in breast carcinogenesis.

Keywords: estrogen receptor; breast cancer; endocrine resistance; mutation; receptor folding

check for updates

Citation: Clusan, L.; Le Goff, P.; Flouriot, G.; Pakdel, F. A Closer Look at Estrogen Receptor Mutations in Breast Cancer and Their Implications for Estrogen and Antiestrogen Responses. *Int. J. Mol. Sci.* **2021**, *22*, 756. https://doi.org/10.3390/ijms22020756

Received: 26 November 2020
Accepted: 31 December 2020
Published: 13 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Breast Cancer Types

Millions of women develop breast cancer worldwide, representing a major health issue. Notably, breast cancer also exists in men but is very rare, accounting for fewer than one percent of cases. This disease generally arises from the proliferation of epithelial cells in the lobules or lactiferous ducts of the mammary gland and is a very heterogeneous malignancy. According to histopathological data, breast cancers are classified as lobular or ductal, in situ or invasive carcinomas preferentially colonizing bone, liver, lung or brain. In addition to histological grade and cancer stage determination, the development of molecular techniques has shed light on the heterogeneity of molecular profiles and gene expression across the types of breast cancer, resulting in more than 20 subtypes of breast carcinoma being characterized [1,2]. Molecular markers rely primarily on the expression of relevant receptors, including estrogen receptor alpha (ERα), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2). Various expressions of these receptors by breast cancer cells are correlated with different degrees of differentiation and aggressiveness of the tumor. Knowing these characteristics enables improved prognosis and selection of the most relevant therapy [3,4]. The simplest classification relies on three major subtypes: luminal, HER2-enriched, and triple-negative, as shown in Table 1 [5].

The luminal type of breast cancer is divided into two subclasses: A and B. Both are characterized by the expression of ERα, but differ in terms of aggressiveness: luminal A breast cancers are usually low-grade, whereas luminal B cancers display overexpression of HER2, reduced ERα expression, and increased proliferation. Such breast cancers are

predominant, especially luminal A, and convey a better prognosis because these tumors depend on estrogen for their growth, and specifically targeting estrogen through endocrine therapy to block its proliferative action, which is an effective strategy. As reviewed by Jensen and Jordan, the identification of the estrogen receptor and subsequent understanding of its implication in breast cancer paved the way for developing targeted therapy [8]. The selective estrogen receptor modulator (SERM) tamoxifen was first developed during the 1970s and became the standard of care for breast cancer, as it enabled the saving of many lives with fewer side effects than chemotherapy. Nevertheless, the use of tamoxifen presents some drawbacks, which led to the development of additional SERMs, as reviewed by Maximov et al. [9]. While SERMs are competitive inhibitors of ERα that prevent its activation by estrogens, another way to counteract ERα activity relies on using selective estrogen receptor downregulators (SERDs), whose binding to ERα results in the degradation of the receptor. Finally, another therapeutic approach aims to directly deprive the tumor of estrogen by ovariectomy or the use of aromatase inhibitors. The development of this class of agents began during the 1980s by Brodie and colleagues and had benefits for patients who do not respond to SERMs [10]. However, for 30–50% of ER-positive breast cancers, resistance to endocrine therapy occurs. For this large number of patients, the prognosis is worse, which raises real public health concerns [5].

Table 1. Breast cancer classification based on receptor expression (ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2).

Breast Cancer Type	Proportion	Biological Profile	Therapy of Choice
Luminal			
A	60%	ERα+ PR+/− HER2−	Endocrine therapy
B	10%	ERα+ PR+/− HER2+	
HER2-enriched [6]	20%	ERα+/− PR+/− HER2+	Anti-HER2 therapy
Triple negative [7]		ERα− PR− HER2−	
Basal-like	7%	+ basal markers	Chemotherapy
Non-basal-like	3%	− basal markers	

1.2. ERα Activity

Estrogen receptors are nuclear receptors that mediate estrogen actions by regulating gene expression. Two highly homologous protein isoforms exist in vertebrates, ERα and ERβ, encoded by two independent genes. However, several primarily in vivo studies have shown that their activity differs in mammary gland development and nonreproductive tissue functioning, as well as in breast cancer pathogenesis [11]. This difference is in line with their differential tissue expression. Concerning breast cancer, clinicopathological data notably reveal that ERβ levels decrease during carcinogenesis [12,13]. The precise role of ERβ remains elusive; several lines of evidence confer it tumor suppressive activity, but further studies are needed to gain insights into the mechanisms involved [14].

Much more is known about ERα. Its activity is essential for the reproductive system, and its expression is increased in most breast cancers, which has led to extensive research to understand its regulatory role. ERα is a ligand-inducible transcriptional factor. After ligand binding and dimerization, ERα is recruited to the promoter region of the target genes, either by binding, directly targeting DNA sequences called EREs (estrogen responsive elements), or by protein/protein interactions with other transcriptional factors, such as AP1 or SP1. Recent development of chromatin immunoprecipitation of DNA coupled to high-throughput sequencing (ChIP-Seq) techniques has led to the identification of the ERα cistrome in mammary adenoma carcinoma cell lines, such as MCF-7, with 5000 to 10,000 estrogen receptor binding sites (ERBS), three quarters of which are EREs [15]. This results in the recruitment of numerous coactivators (members of the p160 family, CBP/P300, members of the SWI/SNF family, members of the mediator complex, etc.) via the AF1 and AF2 transactivation domains of ERα in an ordered, cyclic and combinatorial process,

leading to transcriptional activation of target genes [16]. In addition to this activity at the genomic level, ERα also has nongenomic activity by interacting in the cytoplasm with cellular kinases that activate various signaling pathways, such as PI3K-AKT and Src-MAPK (Figure 1). These rapid actions of the receptor may ultimately result in the regulation of gene expression, highlighting the complex interrelationships between membrane and nuclear-induced events. All these modes of action ultimately lead to the regulation of cell fate, resulting in a balance between proliferation, differentiation and cell survival [11].

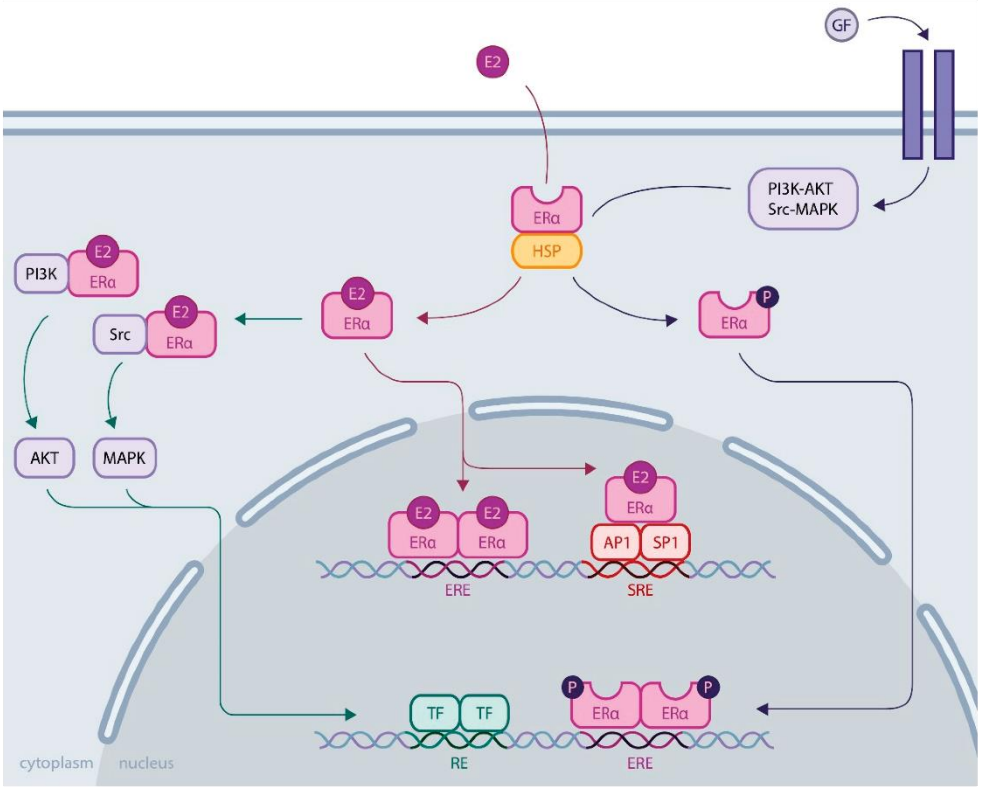


Figure 1. ERα activity through genomic and nongenomic actions. **Genomic activity**: ERα activation upon estrogen binding (E2) or after its phosphorylation by cellular kinases following growth factor (GF) receptor stimulation allows its release from heat shock proteins (HSPs). Then, ERα translocates into the nucleus, where it binds DNA by direct (through estrogen responsive elements, EREs) or indirect mechanisms (upon binding to other transcription factors, such as AP1 or SP1, that bind DNA through serum responsive elements, SREs). **Nongenomic activity**: Activated ERα interacts with cellular kinases (e.g., PI3K and Src), leading to signaling pathway stimulation involving AKT or MAPK, for example, eventually resulting in transcription factor (TF) activation. All these mechanisms induce transcriptional activation or repression of the regulation of cell fate.

1.3. Role of ERα in Breast Cancer

Many players are involved in breast carcinogenesis at the cellular level. Notable examples include tumor suppressor genes BRCA1/2, TP53 or PTEN undergoing loss-of-function mutations or decreased expression, receptor tyrosine kinases overexpression, such as EGFR, IGF1R or HER2, whose downstream signaling pathways promote cell proliferation and invasion (PI3K-AKT, RAS-MAPK, JNK) or overexpression of the oncogene c-Myc [17,18].

However, for most breast cancers, ER deregulation plays a major role, participating in aberrant cell proliferation and leading to tumor development. This deregulation of ERα function is due to multiple phenomena, implicating a shift in the balance of cofactors in

favor of coactivators, the overexpression and function of growth factor receptors whose signaling pathways (MAPK, PI3K, etc.) result in ERα activation or alteration of ERα expression at the genetic level through epigenetic regulation [19–21]. ERα can also be mutated and become constitutively active, but it is not the primary source of breast cancer development. Instead, activating mutations are acquired following estrogen deprivation therapies as a resistance mechanism of tumor cells to escape hormonal control and promote cell proliferation through ligand-independent activation of ERα [22]. Breast cancer cells have evolved into many other strategies against endocrine therapies, as reviewed by Musgrove and Sutherland [23].

Recent developments in biophysical techniques have enabled us to gain insights into ERα conformational changes related to mutations, causing increasing interest for understanding of the ERα response to estrogen and antiestrogens. Whereas several ERα alterations have been reviewed elsewhere, such as gene amplification or translocation [24], as well as splice variants [22], this review aims to summarize the latest advances relative to the particular implication of ERα point mutations in endocrine resistance.

2. ERα Missense Mutations

2.1. The Ligand-Binding Domain

Since the first discovery of a missense mutation of ERα in a breast cancer sample in 1997 [25], several mutations have been identified through cohort studies. By comparing mutations detected in samples of primary versus metastatic tumors, it has been demonstrated that most missense mutations are acquired under selective pressure of endocrine therapies that create a low-estrogen environment, such as aromatase inhibitors [26,27]. In fact, an analysis performed by The Cancer Genome Atlas Network in 2012 did not detect significant mutations in the ERα gene (*ESR1*) in primary breast cancer samples contrary to other genes, such as PIK3CA (49% of luminal A tumor samples) or TP53 (32% of luminal B patients) [28]. When investigating *ESR1* mutations in metastatic breast cancer, however, the prevalence of missense mutations expands to 20–50% [29,30]. Notably, these mutations are localized in the ligand-binding domain (LBD) of ERα (Figure 2A), and several biophysical and functional studies have enabled us to decipher their consequences on ERα activity and their role in endocrine resistance.

The LBD is a highly structured region with three layers of α-helices (h1 to h12) and two β-sheets forming a hydrophobic pocket where the ligand binds (Figure 2B). Agonist binding induces structural modifications of the receptor, where h12 plays a critical role in generating a more compact conformation of ERα. These structural rearrangements participate in coactivator recruitment for ERα transcriptional activity. Antagonist binding, however, inhibits the receptor by preventing h12 from folding properly [11]. Thus, it is not surprising that most mutations acquired by ERα, in response to antiestrogens, localize in the LBD and impact ligand binding.

The most prevalent ERα mutation is a substitution of the amino acid Y537 in S, N or C (a Y537D mutation was also observed in one patient) with a prevalence reaching 60% of mutations detected in metastatic breast cancer samples [29,31]. Such mutations result in a conformational modification of the receptor that stabilizes it in its agonist form, conferring ligand-independent, constitutive activity to the mutated receptor. According to the crystal structures of the Y537S mutant, this conformational modification is due to replacement of the Y537-N348 interaction with a S537-D351 hydrogen bonding that optimizes the h11-h12 loop in the agonist conformation [32]. In this conformation, coactivators can be recruited to the AF2 cleft, and Fanning et al. showed that this binding occurs with a high affinity for the Y537S mutant, even in the absence of estrogens. This explains why this mutation confers constitutive ligand-independent activity to ERα [33]. The same study demonstrated that conformational rearrangements occurring around the h11-12 loop of the receptor, conferring an agonist-bound-like structure to ERα, reduce its affinity for the SERM tamoxifen. In addition, it was suggested that the SERM-bound Y537S mutant adopts an altered conformation compared to the wild type receptor bound to tamoxifen,

participating in a decrease in efficacy of such therapeutic agents. SERDs such as fulvestrant, however, target h12 in a different way and still inhibit ERα, but increased therapeutic doses seem to be necessary [34,35]. Of note, the Y537S and C mutations were also detected in vitro in breast cancer cell lines after depriving them of estrogen to mimic the acquisition of endocrine resistance. In line with previous studies, these mutations confer ligand-independent activities to ERα and altered responses to endocrine therapy [36].

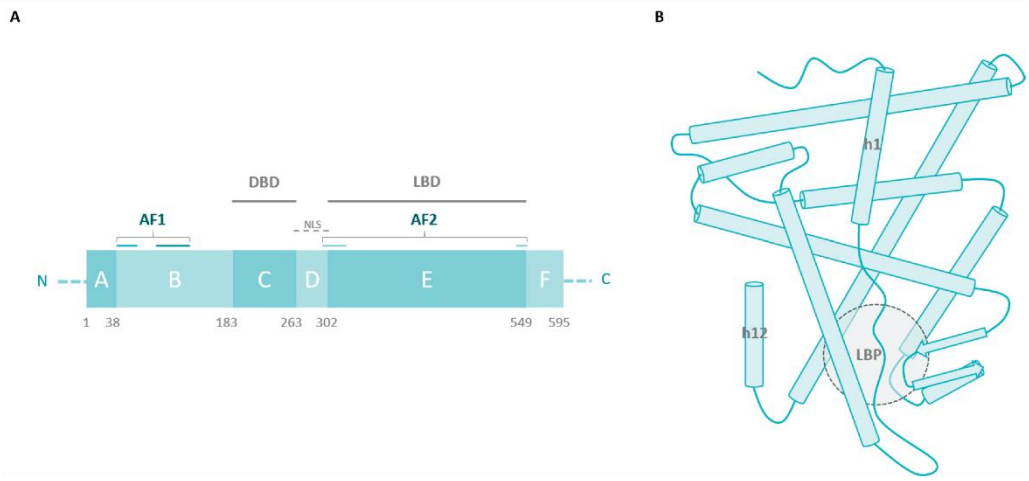


Figure 2. (**A**) ERα is composed of 595 amino acids forming 6 domains, from A to F. Transcriptional activation function 1 (AF1) is localized in the N-terminal region of the receptor, whereas a second transactivation function (AF2) is generated at the C-terminus when conformational rearrangements take place. The C and E domains contain the DNA- and ligand-binding domains, respectively (DBD, LBD). Finally, the D domain is called the hinge region, which participates in conformational changes and protein/protein interactions and contains nuclear localization signals (NLS) [11]. (**B**) A schematic representation of ERα LBD (from the crystal structure 1GWR deposited in the Protein Data Bank) with the ligand-binding pocket (LBP) depicted.

Another LBD mutation particularly observed in patients following antiestrogen treatment affects D538 (at a frequency of 20%, as reported by Katzenellenbogen et al. [32]). When substituted with G, modifications of the electrostatic environment and an increase in h12 flexibility occur, which also results in stabilization of the agonist form of ERα [32]. This leads to increased affinity for coactivators in a ligand-independent way, conferring constitutive activity to the receptor, but this activity is moderate, compared to Y537 mutations. Insights into this phenotypic difference were provided by structural studies, showing that the agonist-bound-like conformation conferred by the D538G mutation is less stable than the conformation permitted by the Y537S mutation [33]. The D538G substitution still confers increased migratory capacities to cancerous cells, which probably contribute to metastasis [37]. A study by Li and colleagues analyzed the growth of ER-positive breast tumors carrying the ER Y537S mutation in patient-derived xenografts (PDXs) after transplantation into ovariectomized mice [38]. Results showed greater tumor growth compared to tumors with wild type ERα under low estrogen conditions and an incomplete response to antiestrogenic treatments. Other studies have evaluated the response of ERα mutations (Y537S, D538G) to estradiol and antiestrogens by measuring the activation of a reporter luciferase gene [34], or endogenous ERα target genes [35]. These ERα mutants exhibited high constitutive transcriptional activation, in contrast to wild type ERα, which shows low activity in the absence of estradiol. In addition, mammary cancer MCF-7 cell lines that stably express ERα Y537S and D538G showed dramatically higher proliferation than wild type MCF-7 cells, suggesting that these ERα mutations induce a significant growth consequence in breast cancer cells [35]. Of note, Y537S and D538G mutants showed an

increase in the interaction with the transcriptional coactivators AIB1 and SRC-1 compared to the wild type receptor, which is consistent with the increase in their ligand-independent activity observed in reporter gene assays [26].

Additional mutations were detected at this hotspot in breast tumor samples, L536R, P535H, V534E etc., but were less prevalent (usually less than 5% of the mutations detected in metastatic breast cancers), and studies investigating their effect on ERα conformation and crystal structures are especially lacking [30]. As those mutations also confer ligand-independent activity to ERα, it is reasonable to speculate that the mechanisms involved rely at least partially on stabilization of the receptor in its agonist form as well.

Concerning L536, for example, its replacement with R, Q, P or H probably reduces the hydrophobicity of the environment, which enables the rearrangement of the h11-h12 loop favoring the agonist conformation of ERα in the absence of estrogens; this altered conformation increases the binding of coactivators for ligand-independent activity [32].

Finally, it is noteworthy that other LBD-activating mutations are outside the mutational hotspots previously mentioned.

E380Q notably appears to be the third prevailing ERα mutation, with a detection rate up to 14% among patients with advanced breast cancer following aromatase inhibitor treatment [29]. This would neutralize charge repulsion between residues in h5 and h12, which would favor an active conformation of the receptor without ligand binding [32]. In vitro studies indeed emphasized the constitutive activity of this mutated receptor, resulting in increased target gene transcription and cell proliferation in the absence of estrogens [39,40]. Evidence is lacking to explain the underlying mechanisms because ligand-independent coactivator binding does not seem to be involved [40]. One hint could be a defect in the corepressor PHB2 binding, which would enhance ERα signaling [41]. Additionally, increased sensitivity to estrogens has been reported, which contributes to promoting tumor growth and resistance to aromatase inhibitors [39].

Finally, the S463P mutation is less documented (with an apparent distribution inferior to 4% in metastatic breast cancer patients [29]), and only speculations can be made about the conformational modifications it might induce. Located between h9 and h10, mutation of this residue could notably affect ERα binding to heat shock proteins (HSPs) and/or dimer stability [32]. In addition to E380Q, the S463P mutation presents only slight constitutive activity assessed by target gene transcription, although no interaction with coactivators was detected without estrogen stimulation in vitro. Nonetheless, hormone-independent cell proliferation is observed when ERα contains this mutation, raising questions about its functional role in breast cancer [40].

It should be noted that several mutations are sometimes detected in the same tumor sample, but it could not be established whether they reside within the same cell population. Chandarlapaty et al., for example, identified concomitant of the D538G and Y537S mutations in the plasma of 30/541 patients (5.5%) with metastatic breast cancer [42]. In any case, this observation highlights the intrinsic heterogeneity of breast cancers and the mechanisms involved in the development of endocrine resistance [30].

Usually acquired after the first line of endocrine therapies, the mutations presented here counteract the efficacy of aromatase inhibitors and alter the inhibitory effect of SERMs and SERDs to various extents (Table 2).

Table 2. Major ERα mutations discussed in this review and a summary of their characteristics. Purple: mutations in the LBD; blue: mutations outside the LBD (AI: aromatase inhibitor; SERD: selective estrogen receptor downregulator; SERM: selective estrogen receptor modulator; E2: estrogen).

ERα Substitution	Y537S/N/C	D538G	L536R/Q/P/H	E380Q	S463P	K303R
Structural data obtained	Stabilization of the agonist conformation	Stabilization of the agonist conformation				
Ligand independent activity	↑ target genes transcription ↑ coactivator recruitment ↑ proliferation	↑ target genes transcription ↑ coactivator recruitment ↑ proliferation ↑ migratory properties	↑ target genes transcription ↑ coactivator recruitment	↑ target genes transcription ↑ proliferation	↑ target genes transcription ↑ proliferation	↑ ERα stability ↑ coactivator recruitment ↑ interactions with growth factor receptors
Estrogen and antiestrogen responses	AI resistance ↓ SERD sensitivity SERM resistance	AI resistance SERD sensitivity SERM resistance	AI resistance SERD sensitivity	AI resistance SERD sensitivity SERM sensitivity ↑ E2 sensitivity	SERD sensitivity SERM sensitivity	AI resistance SERD sensitivity SERM = agonist activity ↑ E2 sensitivity
References	[26,30–36,38,40,42,43]	[26,30–35,37,40,42,43]	[32,34,40]	[30–32,38–41]	[32,35,40]	[44–50]

According to in vitro studies, Y537S/N/C, D538G and L536Q mutations that confer ligand-independent activity to ERα remain sensitive to the SERM tamoxifen and the SERD fulvestrant when therapeutic doses are increased [34,35,37]. In contrast, in vitro and in vivo fulvestrant is effective in tumors driven by mutations such as E380Q and S463P, indicating the heterogeneity of therapeutic responses depending on the mutations driving breast cancer endocrine resistance [40]. From analysis of clinical data, it is unclear whether the SERD fulvestrant is effective for patients harboring LBD-activating mutations [31,51,52]. Using high doses of fulvestrant was then explored for patients developing resistance to aromatase inhibitors, and new therapeutic strategies were investigated to inhibit mutated ERα in a more potent and specific way. Notably, orally available SERDs with improved bioavailability are under development [53]. Additionally, Toy et al. showed by combining in vitro and in vivo studies that AZD9496 or GDC-0810 provided complete inhibition of tumors driven by mutated ERα, contrary to fulvestrant, against which the Y537S mutant was particularly resistant [40]. Likewise, elacestrant (RAD1901) inhibited ERα signaling and tumor development in PDX models harboring either wild type or mutant ERs [54]. Elacestrant is currently in a phase III trial for patients with ER-positive advanced breast cancer [55]. Patients with metastatic breast cancer carrying LBD *ESR1* mutations have poorer overall survival [31,42], which is consistent with the study by Jeselsohn et al. who showed that Y537S and D538G ERα mutants exhibit specific cistromes and transcriptomes that promote tumor metastatic phenotypes [43]. Thus, to treat cancers harboring *ESR1* LBD mutants, it will be necessary to develop combined treatments composed of SERD/SERM associated with other compounds, such as THZ1, a CDK7 inhibitor, making it possible to fight against the metastatic propensity of these mutants [43].

2.2. Outside the LBD

While less frequent than LBD-activating mutations, other ERα mutations are observed in breast cancer patients. The primary recurrent mutation is the K303R substitution, which was detected in 5–10% of invasive breast cancers, but could be more prevalent if more studies were performed with alternative sequencing techniques, according to Fuqua et al. [44]. Unlike mutations occurring in the LBD, the acquisition of this alteration in the hinge region

does not appear to result from selection under endocrine therapy [44–46]. This part of the receptor is the target of multiple posttranslational modifications (PTMs): by affecting residues 266 to 305, modifications such as acetylation, phosphorylation, methylation, ubiquitination and sumoylation enable the interaction with multiple coregulators and play a role in DNA binding to regulate transcription [47].

The K303R mutation results in hypersensitivity to estrogen and a reduction in endocrine therapy efficacy due to a combination of molecular mechanisms that have been previously reviewed [44]. This phenotype is notably due to an increase in the phosphorylation of S305 by cellular kinases, such as PKA [48], and the inhibition of other PTMs surrounding the mutated residue. These modifications of the hinge region enhance the stability of the receptor, alter coregulator binding in favor of coactivator recruitment, and favor interactions with growth factor receptors and downstream signaling pathways. All of these factors lead to increased ligand-independent activity of ERα and a better response to estrogen stimulation, allowing tumor cells to grow in a low-estrogen environment, contributing to the resistance to aromatase inhibitors [49]. Additionally, the response to SERM is altered, and the mutated receptor responds to tamoxifen as an agonist [50], which is not the case for the SERD fulvestrant [48]. Structural data are lacking for our understanding of the phenotype caused by the K303R substitution in ERα. Considering the role played by the hinge region in protein interactions and PTMs of the receptor, such a mutation could lead to a conformational modification that would be relevant for endocrine resistance. Investigating the arrangement of ERα in response to the K303R mutation would then be of great interest to conceive new therapeutic strategies.

Other missense mutations were identified outside the LBD of ERα in patients with breast cancer, such as the S47T, N69K and A86V substitutions in the AF1 domain and the L296P point mutation in the hinge domain. Cell-based assays did not demonstrate alterations of their transcriptional activity compared to the wild type receptor, but no further studies were conducted to investigate these mutations or to decipher their potential role in breast carcinogenesis and drug resistance [56].

3. ERα Synonymous Mutations

Restriction fragment length polymorphisms (RFLPs) of the ERα gene were first identified in introns, as well as exons [57]. Subsequently, the development of DNA sequencing enabled the detection of additional silent mutations in breast cancer patients [58]. Many studies have investigated the association of such polymorphisms with the risk of developing breast cancer (Table 3).

Table 3. Main polymorphisms of the ERα gene (*ESR1*) investigated in association studies. The major allele was selected in agreement with the single nucleotide polymorphism database dbSNP. (RFLP: restriction fragment length polymorphism).

RFLP	rsID	Domain	Codon	Major Allele	Minor Allele	Amino Acid
PvuII	rs2234693		397 (Intron 1)	T	C	
XbaI	rs9340799		351 (Intron 1)	A	G	
	rs2077647	A/B	10 (Exon 1)	TCT	TCC	Ser
BstUI	rs746432	A/B	87 (Exon 1)	GCG	GCC	Ala
		C	243 (Exon 3)	CGC	CGT	Arg
	rs1801132	E	325 (Exon 4)	CCG	CCC	Pro
	rs2228480		594 (Exon 8)	ACG	ACA	Thr

Despite the number of studies realized, no consensus has emerged due to several drawbacks:

Small sample size: Most studies included only a few hundred patients and controls, resulting in low statistical power for determining associations;
Control source: Some studies compared data from breast cancer patients to controls originating from the entire population, whereas other studies used data from the hospital as a control, which could induce bias in the observed associations [59];
Ethnicity: Association studies are generally performed in populations from a unique geographical origin, resulting in conflicting results between women of European and Asian or African ancestry, for example, due to diverse genetic backgrounds [60,61];
Analytic methods: The heterogeneity of methods employed to analyze the association of *ESR1* silent mutations with breast cancer development plays a role in the inconsistency of conclusions as well.

Meta-analyses are then of interest to combine the results of several studies, enabling us to increase the sample size and gather data from different ethnic groups. According to the latest studies in this field, some *ESR1* variants effectively seem to be associated with breast cancer risk, such as rs2234693 and rs9340799, but data from patients of African ancestry or environmental factors were lacking in these analyses [62,63]. New association studies with larger sample sizes and more comprehensive information about patients' lifestyles are needed to clarify the association of *ESR1* polymorphisms and synonymous mutations with breast cancer development to determine whether they are a relevant risk or prognostic factors.

Functional studies are also needed to decipher the role of synonymous ERα mutations in breast carcinogenesis. As they do not alter the sequence of the protein, experimental data are indeed lacking regarding the role that such mutations could play in ERα transcription, translation and functions.

After being ignored for a long time, synonymous mutations began to draw attention only more recently, following increasing evidence that codons used in mRNAs play a role in their translation. In fact, the use of synonymous codons can modify amino acid incorporation during translation elongation due to various parameters reviewed by Hanson and Coller [64]. First, the differential codon composition of an mRNA can alter its secondary structure and stability, which affects its translation rate [65–67]. Furthermore, the composition of the tRNA pool (concentration of each species, charging in amino acids, posttranscriptional modifications) results in differential availability of cognate tRNAs for each codon, which impacts thermodynamic parameters of anticodon-codon pairing and wobble base pairing [68]. This tRNA pool varies notably in a tissue-specific manner [69], depending on the differentiation state of the cell [70], and appears to be deregulated in cancer [71–73]. The combination of these factors can then alter the kinetics of translation during the initiation and elongation steps, which impact cotranslational folding of the emerging peptide. After being hypothesized in 1987 [74], it is indeed now admitted that many proteins begin to fold during their translation. This process already occurs in the polypeptide exit tunnel and when emerging from it through transient electrostatic interactions with the ribosome, protein folding activity of ribosomes (PFAR) and binding to chaperone proteins [68,75]. Translation speed is then a major actor of this cotranslational folding, with the use of codons translated faster or slower to characterize boundaries between protein domains, which participate in their proper folding [64,67,75]. Notably, the translational process is deregulated in cancer cells to meet their increased need for protein synthesis, enabling the expression of specific proteins for tumor growth [76]. Such a phenomenon notably relies on modifications of ribosome biogenesis in response to oncogenic signaling and is linked to the concept of specialized ribosomes [77,78]. This translational specificity of cancer cells could eventually participate in the modification of cotranslational folding of proteins as well. A mutation resulting in a synonymous codon substitution, although not modifying the amino acid sequence of the protein, can then alter mRNA stability and translation speed, resulting in the modification of protein folding and thereby

altering protein expression, stability [79], and function [68,80]. In addition, the use of a synonymous codon less suited to the cellular environment can promote the incorporation of the wrong amino acid [67]. Therefore, synonymous mutations, although not modifying the sequence of the protein, appear to play a functional role in protein translation and function (Figure 3).

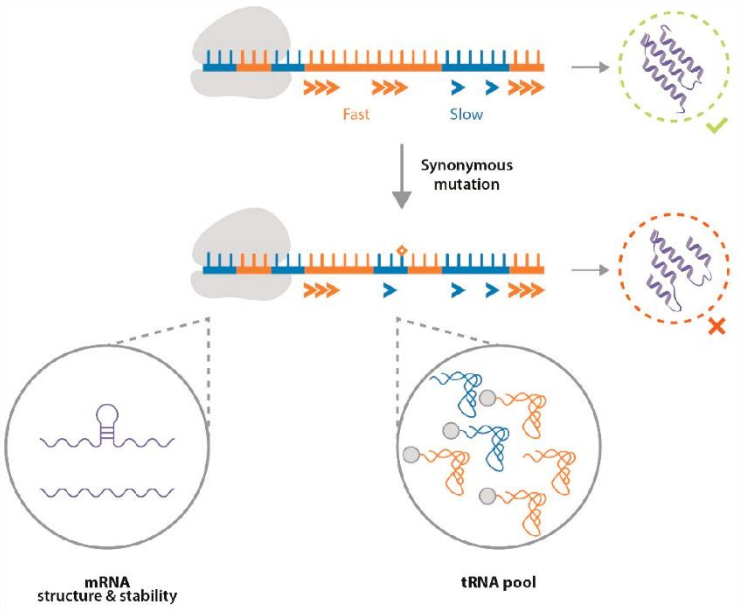


Figure 3. Synonymous mutations (represented by a star) can modify translational speed by impacting various parameters, such as the mRNA structure and the stability or availability of cognate tRNAs. This can ultimately result in the alteration of protein conformation, affecting its expression, stability and function.

It is relevant to speculate that such mutations could play a role in the development of human diseases, as well as in the efficacy of therapies [68,75,81]. This began to be notably demonstrated with the example of the CFTR gene implicated in cystic fibrosis [82], or the oncogene KRAS [83]. With the development of next-generation sequencing and the increasing amount of data freely accessible through online databases, many more polymorphisms of the human genome could be investigated for their potential implication in human health [84].

Concerning ERα, it has been shown through in vitro translation that differences in the translation machinery modify the conformation of the produced receptor, which highlights a role of the translational process in the proper folding of the protein [85]. Fernández-Calero et al. showed that the synonymous mutation Ala87 (BstUI, rs746432) in the A/B domain of ERα alters its transcriptional activity and nuclear export, suggesting a conformational modification of the receptor [86]. Finally, Hertz and colleagues reported that another *ESR1* SNP (rs9322336) could be associated with an increase in ERα gene expression in ER-positive breast cancer patients, but no significant change of protein levels was recorded [87]. In light of recent research on cotranslational folding and the first data concerning ERα, we can hypothesize that synonymous mutations of this receptor could modify its conformation, playing a role in its interactions with cofactors and ligands, thereby altering its activity and cell fate regulation. Thus, further investigating the biological implications of synonymous ERα mutations through functional and structural studies would definitely be of great interest to increase our understanding of breast carcinogenesis and endocrine resistance and to identify new therapeutic strategies.

4. Conclusions

ERα is a major player in breast cancer development. By reducing circulating estrogen or directly inhibiting ERα, endocrine therapy is an effective strategy against luminal breast cancers. However, endocrine resistance exists in a significant number of patients, raising public health concerns. One of the mechanisms enabling tumor cells to escape hormonal control is mutation of ERα. Recurrent missense mutations were notably identified in the LBD of the receptor that confer ligand-independent activity by stabilizing its agonist form without estrogen binding. The development of biophysical studies has indeed enabled insights into the conformational modifications induced by ERα mutations detected in patients. Such information is of great value to better understand the biology of endocrine resistance and design new therapeutic strategies. In fact, the development of resistance towards standard of care highlights the need for more potent and selective agents against ERα in its native and mutated forms. In this regard, news SERDs such as GDC-0810, AZD9496 or RAD1901, were proven to efficiently target *ESR1* mutants in preclinical studies, with better pharmacokinetic properties than fulvestrant [40,54]. Furthermore, such therapeutic molecules can be used in combination with inhibiting actors of the metastatic process, such as cyclins, growth factors or PI3K/AKT/mTOR signaling pathways [88]. Novel molecules continue to show promising in vitro effects in recent years, such as AF2-specific inhibitors [89], selective ER covalent antagonist (SERCA) [90], and proteolysis-targeting chimera (PROTAC) [91]. In addition, the principle of biased ligands, well developed for GPCRs, could be applied to fine-tune ER signaling in breast cancer [92].

Structural data are still lacking for many point mutations known to play a role in hormonal escape of cancerous cells. The development of biophysical techniques should then improve the characterization of those receptors to develop more potent targeted therapies. A recent study conducted by Huang et al. revealed notable new interactions between the DBD and LBD of ERα that play a role in its genomic activity [93]. A missense mutation observed in endometrial cancer (Y191H) appeared to have a structural role in this interaction, which correlates with increased ERα transcriptional activity assessed in vitro [93]. While no oncogenic role has been reported for the Y191H substitution, this study highlights the significance of investigating ERα structure to extend our understanding of the mechanisms implicated in its activity and their potential roles in breast cancer development and endocrine resistance. Improving the detection of ERα mutations (notably by noninvasive methods) is critical as well to enable better patient care by adapting the therapeutic strategy faster in a personalized way [94,95]. In fact, the relevance of analyzing circulating DNA isolated from plasma samples to select therapy for breast cancer patients was recently assessed by Turner and colleagues [96].

In addition to point mutations leading to modifications of the amino acid sequence of ERα, synonymous mutations also exist. These mutations do not alter the protein sequence and were consequently ignored for a long time. Furthermore, no consensus has emerged regarding the association of such silent mutations with the risk of developing breast cancer. The association studies performed so far present several drawbacks, and more studies are needed to investigate the potential association of synonymous ERα mutations with breast cancer. Indeed, recent advances in the biology of codon usage and translation suggest that synonymous mutations play a role in a mechanism called cotranslational folding and therefore modify the conformation of the receptor. In this way, such mutations could modify ERα interactions and ligand binding, which could participate in tumorigenesis and endocrine resistance. More studies are needed to confirm whether synonymous mutations of ERα have a functional role in breast cancer development and therapeutic response.

Moreover, taking into account such a hypothesis would enlarge the field of therapies potentially suitable to overcome endocrine resistance. If synonymous mutations contribute to breast carcinogenesis by altering ERα function through the modification of its conformation due to an alteration of the cotranslational folding process, targeting the translational machinery could represent an additional line of treatment for endocrine therapy and the new ERα antagonists discussed above. With the increasing knowledge concerning the role

played by ribosomes in tumorigenesis, this molecule has already emerged as a therapeutic target in cancer [97]. In addition, small molecules targeting the protein folding activity of ribosomes were developed as antiprion drugs [98], suggesting that more specific inhibitors of the ribosome could be designed as adjuvant therapies. In any case, investigating the role of synonymous mutations in human disease is a new area worth exploring [81,99].

Author Contributions: L.C., P.L.G., G.F. and F.P. were the major contributors in writing the manuscript. L.C., P.L.G., G.F. and F.P. read, corrected and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The PhD work of L.C. was funded by MENRT scholarships from the French Minister of High Education, Research and Innovation.

Acknowledgments: This work was supported by the Ligue Contre le Cancer, the University of Rennes 1, CNRS and INSERM.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

AF1	transcriptional activation function 1
DBD	DNA-binding domain
E2	estrogen
ERα	estrogen receptor alpha
ERBS	estrogen receptor binding site
ERE	estrogen responsive element
GF	growth factor
HER2	human epidermal growth factor receptor 2
HSP	heat shock protein
LBD	ligand-binding domain
LBP	ligand-binding pocket
PDX	patient derived xenograft
PFAR	protein folding activity of ribosomes
PR	progesterone receptor
PROTAC	proteolysis-targeting chimera
PTM	posttranslational modification
RFLP	restriction fragment length polymorphism
SERCA	selective ER covalent antagonist
SERD	selective estrogen receptor downregulator
SERM	selective estrogen receptor modulator
SRE	serum responsive element
TF	transcription factor

References

1. Lakhani, S.; Ellis, I.; Schnitt, S.; Tan, P.; van de Vijver, M. *WHO Classification of Tumours of the Breast*, 4th ed.; International Agency for Research on Cancer: Lyon, France, 2012; ISBN 978-92-832-2433-4.
2. Vuong, D.; Simpson, P.T.; Green, B.; Cummings, M.C.; Lakhani, S.R. Molecular Classification of Breast Cancer. *Virchows Arch.* **2014**, *465*, 1–14. [CrossRef]
3. Bertucci, F.; Finetti, P.; Rougemont, J.; Charafe-Jauffret, E.; Cervera, N.; Tarpin, C.; Nguyen, C.; Xerri, L.; Houlgatte, R.; Jacquemier, J.; et al. Gene Expression Profiling Identifies Molecular Subtypes of Inflammatory Breast Cancer. *Cancer Res.* **2005**, *65*, 2170–2178. [CrossRef]
4. Gajulapalli, V.N.R.; Malisetty, V.L.; Chitta, S.K.; Manavathi, B. Oestrogen Receptor Negativity in Breast Cancer: A Cause or Consequence? *Biosci. Rep.* **2016**, *36*, e00432. [CrossRef]
5. Clarke, R.; Tyson, J.J.; Dixon, J.M. Endocrine Resistance in Breast Cancer—An Overview and Update. *Mol. Cell. Endocrinol.* **2015**, *418*, 220–234. [CrossRef]
6. Costa, R.L.B.; Czerniecki, B.J. Clinical Development of Immunotherapies for HER2 + Breast Cancer: A Review of HER2-Directed Monoclonal Antibodies and Beyond. *NPJ Breast Cancer* **2020**, *6*, 10. [CrossRef]
7. Diana, A.; Carlino, F.; Franzese, E.; Oikonomidou, O.; Criscitiello, C.; De Vita, F.; Ciardiello, F.; Orditura, M. Early Triple Negative Breast Cancer: Conventional Treatment and Emerging Therapeutic Landscapes. *Cancers* **2020**, *12*, 819. [CrossRef]
8. Jensen, E.V.; Jordan, V.C. The Estrogen Receptor: A Model for Molecular Medicine. *Clin. Cancer Res.* **2003**, *9*, 1980–1989.

9. Maximov, P.; Lee, T.; Jordan, V. The Discovery and Development of Selective Estrogen Receptor Modulators (SERMs) for Clinical Practice. *CCP* **2013**, *8*, 135–155. [CrossRef]
10. Jordan, V.C.; Brodie, A.M.H. Development and Evolution of Therapies Targeted to the Estrogen Receptor for the Treatment and Prevention of Breast Cancer. *Steroids* **2007**, *72*, 7–25. [CrossRef] [PubMed]
11. Arnal, J.-F.; Lenfant, F.; Metivier, R.; Flouriot, G.; Henrion, D.; Adlanmerini, M.; Fontaine, C.; Gourdy, P.; Chambon, P.; Katzenellenbogen, B.; et al. Membrane and Nuclear Estrogen Receptor Alpha Actions: From Tissue Specificity to Medical Implications. *Physiol. Rev.* **2017**, *97*, 1045–1087. [CrossRef] [PubMed]
12. Haldosén, L.-A.; Zhao, C.; Dahlman-Wright, K. Estrogen Receptor Beta in Breast Cancer. *Mol. Cell. Endocrinol.* **2014**, *382*, 665–672. [CrossRef] [PubMed]
13. Huang, B.; Omoto, Y.; Iwase, H.; Yamashita, H.; Toyama, T.; Coombes, R.C.; Filipovic, A.; Warner, M.; Gustafsson, J.-Å. Differential Expression of Estrogen Receptor α, B1, and B2 in Lobular and Ductal Breast Cancer. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1933–1938. [CrossRef] [PubMed]
14. Ranganathan, P.; Nadig, N.; Nambiar, S. Non-Canonical Estrogen Signaling in Endocrine Resistance. *Front. Endocrinol.* **2019**, *10*, 708. [CrossRef] [PubMed]
15. Stender, J.D.; Kim, K.; Charn, T.H.; Komm, B.; Chang, K.C.N.; Kraus, W.L.; Benner, C.; Glass, C.K.; Katzenellenbogen, B.S. Genome-Wide Analysis of Estrogen Receptor α DNA Binding and Tethering Mechanisms Identifies Runx1 as a Novel Tethering Factor in Receptor-Mediated Transcriptional Activation. *MCB* **2010**, *30*, 3943–3955. [CrossRef]
16. Métivier, R.; Penot, G.; Hübner, M.R.; Reid, G.; Brand, H.; Koš, M.; Gannon, F. Estrogen Receptor-α Directs Ordered, Cyclical, and Combinatorial Recruitment of Cofactors on a Natural Target Promoter. *Cell* **2003**, *115*, 751–763. [CrossRef]
17. Sun, Y.-S.; Zhao, Z.; Yang, Z.-N.; Xu, F.; Lu, H.-J.; Zhu, Z.-Y.; Shi, W.; Jiang, J.; Yao, P.-P.; Zhu, H.-P. Risk Factors and Preventions of Breast Cancer. *Int. J. Biol. Sci.* **2017**, *13*, 1387–1397. [CrossRef]
18. Fallah, Y.; Brundage, J.; Allegakoen, P.; Shajahan-Haq, A.N. MYC-Driven Pathways in Breast Cancer Subtypes. *Biomolecules* **2017**, *7*, 53. [CrossRef]
19. Murphy, L.; Cherlet, T.; Lewis, A.; Banu, Y.; Watson, P. New Insights into Estrogen Receptor Function in Human Breast Cancer. *Ann. Med.* **2003**, *35*, 614–631. [CrossRef]
20. Manavathi, B.; Dey, O.; Gajulapalli, V.N.R.; Bhatia, R.S.; Bugide, S.; Kumar, R. Derailed Estrogen Signaling and Breast Cancer: An Authentic Couple. *Endocr. Rev.* **2013**, *34*, 1–32. [CrossRef]
21. Kerdivel, G.; Flouriot, G.; Pakdel, F. Modulation of Estrogen Receptor Alpha Activity and Expression During Breast Cancer Progression. In *Vitamins & Hormones*; Elsevier: Cambridge, MA, USA, 2013; Volume 93, pp. 135–160. ISBN 978-0-12-416673-8.
22. Barone, I.; Brusco, L.; Fuqua, S.A.W. Estrogen Receptor Mutations and Changes in Downstream Gene Expression and Signaling. *Clin. Cancer Res.* **2010**, *16*, 2702–2708. [CrossRef]
23. Musgrove, E.A.; Sutherland, R.L. Biological Determinants of Endocrine Resistance in Breast Cancer. *Nat. Rev. Cancer* **2009**, *9*, 631–643. [CrossRef] [PubMed]
24. Ma, C.X.; Reinert, T.; Chmielewska, I.; Ellis, M.J. Mechanisms of Aromatase Inhibitor Resistance. *Nat. Rev. Cancer* **2015**, *15*, 261–275. [CrossRef] [PubMed]
25. Zhang, Q.-X.; Borg, Å.; Wolf, D.M.; Oesterreich, S.; Fuqua, S.A.W. An Estrogen Receptor Mutant with Strong Hormone-Independent Activity from a Metastatic Breast Cancer. *Cancer Res.* **1997**, *57*, 1244–1249. [PubMed]
26. Jeselsohn, R.; Yelensky, R.; Buchwalter, G.; Frampton, G.; Meric-Bernstam, F.; Gonzalez-Angulo, A.M.; Ferrer-Lozano, J.; Perez-Fidalgo, J.A.; Cristofanilli, M.; Gomez, H.; et al. Emergence of Constitutively Active Estrogen Receptor-α Mutations in Pretreated Advanced Estrogen Receptor-Positive Breast Cancer. *Clin. Cancer Res.* **2014**, *20*, 1757–1767. [CrossRef]
27. Schiavon, G.; Hrebien, S.; Garcia-Murillas, I.; Cutts, R.J.; Pearson, A.; Tarazona, N.; Fenwick, K.; Kozarewa, I.; Lopez-Knowles, E.; Ribas, R.; et al. Analysis of ESR1 Mutation in Circulating Tumor DNA Demonstrates Evolution during Therapy for Metastatic Breast Cancer. *Sci. Transl. Med.* **2015**, *7*, 313ra182. [CrossRef] [PubMed]
28. The Cancer Genome Atlas Network. Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* **2012**, *490*, 61–70. [CrossRef]
29. Hamadeh, I.S.; Patel, J.N.; Rusin, S.; Tan, A.R. Personalizing Aromatase Inhibitor Therapy in Patients with Breast Cancer. *Cancer Treat. Rev.* **2018**, *70*, 47–55. [CrossRef]
30. Jeselsohn, R.; Buchwalter, G.; De Angelis, C.; Brown, M.; Schiff, R. ESR1 Mutations—A Mechanism for Acquired Endocrine Resistance in Breast Cancer. *Nat. Rev. Clin. Oncol.* **2015**, *12*, 573–583. [CrossRef]
31. Spoerke, J.M.; Gendreau, S.; Walter, K.; Qiu, J.; Wilson, T.R.; Savage, H.; Aimi, J.; Derynck, M.K.; Chen, M.; Chan, I.T.; et al. Heterogeneity and Clinical Significance of ESR1 Mutations in ER-Positive Metastatic Breast Cancer Patients Receiving Fulvestrant. *Nat. Commun.* **2016**, *7*, 11579. [CrossRef]
32. Katzenellenbogen, J.A.; Mayne, C.G.; Katzenellenbogen, B.S.; Greene, G.L.; Chandarlapaty, S. Structural Underpinnings of Oestrogen Receptor Mutations in Endocrine Therapy Resistance. *Nat. Rev. Cancer* **2018**, *18*, 377–388. [CrossRef]
33. Fanning, S.W.; Mayne, C.G.; Dharmarajan, V.; Carlson, K.E.; Martin, T.A.; Novick, S.J.; Toy, W.; Green, B.; Panchamukhi, S.; Katzenellenbogen, B.S.; et al. Estrogen Receptor Alpha Somatic Mutations Y537S and D538G Confer Breast Cancer Endocrine Resistance by Stabilizing the Activating Function-2 Binding Conformation. *eLife* **2016**, *5*, e12792. [CrossRef]
34. Robinson, D.R.; Wu, Y.-M.; Vats, P.; Su, F.; Lonigro, R.J.; Cao, X.; Kalyana-Sundaram, S.; Wang, R.; Ning, Y.; Hodges, L.; et al. Activating ESR1 Mutations in Hormone-Resistant Metastatic Breast Cancer. *Nat. Genet.* **2013**, *45*, 1446–1451. [CrossRef] [PubMed]

35. Toy, W.; Shen, Y.; Won, H.; Green, B.; Sakr, R.A.; Will, M.; Li, Z.; Gala, K.; Fanning, S.; King, T.A.; et al. ESR1 Ligand-Binding Domain Mutations in Hormone-Resistant Breast Cancer. *Nat. Genet.* **2013**, *45*, 1439–1445. [CrossRef] [PubMed]
36. Martin, L.-A.; Ribas, R.; Simigdala, N.; Schuster, E.; Pancholi, S.; Tenev, T.; Gellert, P.; Buluwela, L.; Harrod, A.; Thornhill, A.; et al. Discovery of Naturally Occurring ESR1 Mutations in Breast Cancer Cell Lines Modelling Endocrine Resistance. *Nat. Commun.* **2017**, *8*, 1865. [CrossRef] [PubMed]
37. Merenbakh-Lamin, K.; Ben-Baruch, N.; Yeheskel, A.; Dvir, A.; Soussan-Gutman, L.; Jeselsohn, R.; Yelensky, R.; Brown, M.; Miller, V.A.; Sarid, D.; et al. D538G Mutation in Estrogen Receptor-α: A Novel Mechanism for Acquired Endocrine Resistance in Breast Cancer. *Cancer Res.* **2013**, *73*, 6856–6864. [CrossRef] [PubMed]
38. Li, S.; Shen, D.; Shao, J.; Crowder, R.; Liu, W.; Prat, A.; He, X.; Liu, S.; Hoog, J.; Lu, C.; et al. Endocrine-Therapy-Resistant ESR1 Variants Revealed by Genomic Characterization of Breast-Cancer-Derived Xenografts. *Cell Rep.* **2013**, *4*, 1116–1130. [CrossRef]
39. Pakdel, F.; Reese, J.C.; Katzenellenbogen, B.S. Identification of Charged Residues in an N-Terminal Portion of the Hormone-Binding Domain of the Human Estrogen Receptor Important in Transcriptional Activity of the Receptor. *Mol. Endocrinol.* **1993**, *7*, 1408–1417. [CrossRef]
40. Toy, W.; Weir, H.; Razavi, P.; Lawson, M.; Goeppert, A.U.; Mazzola, A.M.; Smith, A.; Wilson, J.; Morrow, C.; Wong, W.L.; et al. Activating ESR1 Mutations Differentially Affect the Efficacy of ER Antagonists. *Cancer Discov.* **2017**, *7*, 277–287. [CrossRef]
41. Chigira, T.; Nagatoishi, S.; Tsumoto, K. Differential Binding of Prohibitin-2 to Estrogen Receptor α and to Drug-Resistant ERα Mutants. *Biochem. Biophys. Res. Commun.* **2015**, *463*, 726–731. [CrossRef]
42. Chandarlapaty, S.; Chen, D.; He, W.; Sung, P.; Samoila, A.; You, D.; Bhatt, T.; Patel, P.; Voi, M.; Gnant, M.; et al. Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial. *JAMA Oncol.* **2016**, *2*, 1310. [CrossRef]
43. Jeselsohn, R.; Bergholz, J.S.; Pun, M.; Cornwell, M.; Liu, W.; Nardone, A.; Xiao, T.; Li, W.; Qiu, X.; Buchwalter, G.; et al. Allele-Specific Chromatin Recruitment and Therapeutic Vulnerabilities of ESR1 Activating Mutations. *Cancer Cell* **2018**, *33*, 173–186.e5. [CrossRef] [PubMed]
44. Fuqua, S.A.W.; Gu, G.; Rechoum, Y. Estrogen Receptor (ER) α Mutations in Breast Cancer: Hidden in Plain Sight. *Breast Cancer Res. Treat.* **2014**, *144*, 11–19. [CrossRef] [PubMed]
45. Conway, K.; Parrish, E.; Edmiston, S.N.; Tolbert, D.; Tse, C.-K.; Geradts, J.; Livasy, C.A.; Singh, H.; Newman, B.; Millikan, R.C. The Estrogen Receptor-α A908G (K303R) Mutation Occurs at a Low Frequency in Invasive Breast Tumors: Results from a Population-Based Study. *Breast Cancer Res.* **2005**, *7*, R871. [CrossRef] [PubMed]
46. Abbasi, S.; Rasouli, M.; Nouri, M.; Kalbasi, S. Association of Estrogen Receptor-α A908G (K303R) Mutation with Breast Cancer Risk. *Int. J. Clin. Exp. Med.* **2013**, *6*, 39–49. [PubMed]
47. Le Romancer, M.; Poulard, C.; Cohen, P.; Sentis, S.; Renoir, J.-M.; Corbo, L. Cracking the Estrogen Receptor's Posttranslational Code in Breast Tumors. *Endocr. Rev.* **2011**, *32*, 597–622. [CrossRef]
48. Michalides, R.; Griekspoor, A.; Balkenende, A.; Verwoerd, D.; Janssen, L.; Jalink, K.; Floore, A.; Velds, A.; van't Veer, L.; Neefjes, J. Tamoxifen Resistance by a Conformational Arrest of the Estrogen Receptor α after PKA Activation in Breast Cancer. *Cancer Cell* **2004**, *5*, 597–605. [CrossRef]
49. Barone, I.; Iacopetta, D.; Covington, K.R.; Cui, Y.; Tsimelzon, A.; Beyer, A.; Andò, S.; Fuqua, S.A.W. Phosphorylation of the Mutant K303R Estrogen Receptor α at Serine 305 Affects Aromatase Inhibitor Sensitivity. *Oncogene* **2010**, *29*, 2404–2414. [CrossRef]
50. Giordano, C.; Cui, Y.; Barone, I.; Ando, S.; Mancini, M.A.; Berno, V.; Fuqua, S.A.W. Growth Factor-Induced Resistance to Tamoxifen Is Associated with a Mutation of Estrogen Receptor α and Its Phosphorylation at Serine 305. *Breast Cancer Res. Treat.* **2010**, *119*, 71–85. [CrossRef]
51. Johnston, S.R.; Kilburn, L.S.; Ellis, P.; Dodwell, D.; Cameron, D.; Hayward, L.; Im, Y.-H.; Braybrooke, J.P.; Brunt, A.M.; Cheung, K.-L.; et al. Fulvestrant plus Anastrozole or Placebo versus Exemestane Alone after Progression on Non-Steroidal Aromatase Inhibitors in Postmenopausal Patients with Hormone-Receptor-Positive Locally Advanced or Metastatic Breast Cancer (SoFEA): A Composite, Multicentre, Phase 3 Randomised Trial. *Lancet Oncol.* **2013**, *14*, 989–998. [CrossRef]
52. Cristofanilli, M.; Turner, N.C.; Bondarenko, I.; Ro, J.; Im, S.-A.; Masuda, N.; Colleoni, M.; DeMichele, A.; Loi, S.; Verma, S.; et al. Fulvestrant plus Palbociclib versus Fulvestrant plus Placebo for Treatment of Hormone-Receptor-Positive, HER2-Negative Metastatic Breast Cancer That Progressed on Previous Endocrine Therapy (PALOMA-3): Final Analysis of the Multicentre, Double-Blind, Phase 3 Randomised Controlled Trial. *Lancet Oncol.* **2016**, *17*, 425–439. [CrossRef]
53. Fanning, S.W.; Greene, G.L. Next-Generation ERα Inhibitors for Endocrine-Resistant ER+ Breast Cancer. *Endocrinology* **2019**, *160*, 759–769. [CrossRef] [PubMed]
54. Bihani, T.; Patel, H.K.; Arlt, H.; Tao, N.; Jiang, H.; Brown, J.L.; Purandare, D.M.; Hattersley, G.; Garner, F. Elacestrant (RAD1901), a Selective Estrogen Receptor Degrader (SERD), has Antitumor Activity in Multiple ER + Breast Cancer Patient-Derived Xenograft Models. *Clin. Cancer Res.* **2017**, *23*, 4793–4804. [CrossRef] [PubMed]
55. Bardia, A.; Aftimos, P.; Bihani, T.; Anderson-Villaluz, A.T.; Jung, J.; Conlan, M.G.; Kaklamani, V.G. EMERALD: Phase III Trial of Elacestrant (RAD1901) vs Endocrine Therapy for Previously Treated ER + Advanced Breast Cancer. *Future Oncol.* **2019**, *15*, 3209–3218. [CrossRef] [PubMed]
56. Herynk, M.H.; Fuqua, S.A.W. Estrogen Receptor Mutations in Human Disease. *Endocr. Rev.* **2004**, *25*, 869–898. [CrossRef] [PubMed]
57. Andersen, T.; Heimdal, K.; Skrede, M.; Tveit, K.; Berg, K.; Børresen, A.-L. Oestrogen Receptor (ESR) Polymorphisms and Breast Cancer Susceptibility. *Hum. Genet.* **1994**, *94*, 665–670. [CrossRef]

58. Roodi, N.; Bailey, L.R.; Kao, W.-Y.; Verrier, C.S.; Yee, C.J.; Dupont, W.D.; Parl, F.F. Estrogen Receptor Gene Analysis in Estrogen Receptor-Positive and Receptor-Negative Primary Breast Cancer. *J. Nat. Cancer Inst.* **1995**, *87*, 446–451. [CrossRef]
59. Zhang, Y.; Zhang, M.; Yuan, X.; Zhang, Z.; Zhang, P.; Chao, H.; Jiang, L.; Jiang, J. Association Between ESR1 PvuII, XbaI, and P325P Polymorphisms and Breast Cancer Susceptibility: A Meta-Analysis. *Med. Sci. Monit.* **2015**, *21*, 2986–2996. [CrossRef]
60. Lu, H.; Chen, D.; Hu, L.-P.; Zhou, L.-L.; Xu, H.-Y.; Bai, Y.-H.; Lin, X.-Y. Estrogen Receptor Alpha Gene Polymorphisms and Breast Cancer Risk: A Case-Control Study with Meta-Analysis Combined. *Asian Pac. J. Cancer Prev.* **2013**, *14*, 6743–6749. [CrossRef]
61. Quan, L.; Hong, C.-C.; Zirpoli, G.; Roberts, M.R.; Khoury, T.; Sucheston-Campbell, L.E.; Bovbjerg, D.H.; Jandorf, L.; Pawlish, K.; Ciupak, G.; et al. Variants of Estrogen-Related Genes and Breast Cancer Risk in European and African American Women. *Endocr. Relat. Cancer* **2014**, *21*, 853–864. [CrossRef]
62. Yang, Y.; Shu, X.; Shu, X.; Bolla, M.K.; Kweon, S.-S.; Cai, Q.; Michailidou, K.; Wang, Q.; Dennis, J.; Park, B.; et al. Re-Evaluating Genetic Variants Identified in Candidate Gene Studies of Breast Cancer Risk Using Data from Nearly 280,000 Women of Asian and European Ancestry. *EBioMedicine* **2019**, *48*, 203–211. [CrossRef]
63. Zhang, Z.; Zhang, C.; Li, Y.; Zhao, Z.; Yang, S. Association between ERα Gene Pvu II Polymorphism and Breast Cancer Susceptibility: A Meta-Analysis. *Medicine* **2018**, *97*, e0317. [CrossRef] [PubMed]
64. Hanson, G.; Coller, J. Codon Optimality, Bias and Usage in Translation and MRNA Decay. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 20–30. [CrossRef] [PubMed]
65. Hia, F.; Yang, S.F.; Shichino, Y.; Yoshinaga, M.; Murakawa, Y.; Vandenbon, A.; Fukao, A.; Fujiwara, T.; Landthaler, M.; Natsume, T.; et al. Codon Bias Confers Stability to Human MRNAs. *EMBO Rep.* **2019**, *20*, e48220. [CrossRef] [PubMed]
66. Wu, Q.; Medina, S.G.; Kushawah, G.; DeVore, M.L.; Castellano, L.A.; Hand, J.M.; Wright, M.; Bazzini, A.A. Translation Affects MRNA Stability in a Codon-Dependent Manner in Human Cells. *eLife* **2019**, *8*, e45396. [CrossRef] [PubMed]
67. Komar, A.A. The Yin and Yang of Codon Usage. *Hum. Mol. Genet.* **2016**, *25*, R77–R85. [CrossRef]
68. Marín, M.; Fernández-Calero, T.; Ehrlich, R. Protein Folding and TRNA Biology. *Biophys. Rev.* **2017**, *9*, 573–588. [CrossRef]
69. Dittmar, K.A.; Goodenbour, J.M.; Pan, T. Tissue Specific Differences in Human Transfer RNA Expression. *PLoS Genet.* **2005**, *2*, e221. [CrossRef]
70. Gingold, H.; Tehler, D.; Christoffersen, N.R.; Nielsen, M.M.; Asmar, F.; Kooistra, S.M.; Christophersen, N.S.; Christensen, L.L.; Borre, M.; Sørensen, K.D.; et al. A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. *Cell* **2014**, *158*, 1281–1292. [CrossRef]
71. Santos, M.; Fidalgo, A.; Varanda, A.S.; Oliveira, C.; Santos, M.A.S. TRNA Deregulation and Its Consequences in Cancer. *Trends Mol. Med.* **2019**, *25*, 853–865. [CrossRef]
72. Goodarzi, H.; Nguyen, H.C.B.; Zhang, S.; Dill, B.D.; Molina, H.; Tavazoie, S.F. Modulated Expression of Specific TRNAs Drives Gene Expression and Cancer Progression. *Cell* **2016**, *165*, 1416–1427. [CrossRef]
73. Pavon-Eternod, M.; Gomes, S.; Geslain, R.; Dai, Q.; Rosner, M.R.; Pan, T. TRNA Over-Expression in Breast Cancer and Functional Consequences. *Nucleic Acids Res.* **2009**, *37*, 7268–7280. [CrossRef] [PubMed]
74. Purvis, I.J.; Bettany, A.J.E.; Santiago, T.C.; Coggins, J.R.; Duncan, K.; Eason, R.; Brown, A.J.P. The Efficiency of Folding of Some Proteins is Increased by Controlled Rates of Translation In Vivo. *J. Mol. Biol.* **1987**, *193*, 413–417. [CrossRef]
75. Thommen, M.; Holtkamp, W.; Rodnina, M.V. Co-Translational Protein Folding: Progress and Methods. *Curr. Opin. Struct. Biol.* **2017**, *42*, 83–89. [CrossRef]
76. Truitt, M.L.; Ruggero, D. New Frontiers in Translational Control of the Cancer Genome. *Nat. Rev. Cancer* **2016**, *16*, 288–304. [CrossRef] [PubMed]
77. Penzo, M.; Galbiati, A.; Treré, D.; Montanaro, L. The Importance of Being (Slightly) Modified: The Role of RRNA Editing on Gene Expression Control and Its Connections with Cancer. *Biochim. Biophys. Acta (BBA) Rev. Cancer* **2016**, *1866*, 330–338. [CrossRef]
78. Shi, Z.; Fujii, K.; Kovary, K.M.; Genuth, N.R.; Röst, H.L.; Teruel, M.N.; Barna, M. Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of MRNAs Genome-Wide. *Mol. Cell* **2017**, *67*, 71–83.e7. [CrossRef]
79. Buhr, F.; Jha, S.; Thommen, M.; Mittelstaet, J.; Kutz, F.; Schwalbe, H.; Rodnina, M.V.; Komar, A.A. Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations. *Mol. Cell* **2016**, *61*, 341–351. [CrossRef]
80. Yu, C.-H.; Dang, Y.; Zhou, Z.; Wu, C.; Zhao, F.; Sachs, M.S.; Liu, Y. Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-Translational Protein Folding. *Mol. Cell* **2015**, *59*, 744–754. [CrossRef]
81. Sauna, Z.E.; Kimchi-Sarfaty, C. Understanding the Contribution of Synonymous Mutations to Human Disease. *Nat. Rev. Genet.* **2011**, *12*, 683–691. [CrossRef]
82. Rauscher, R.; Ignatova, Z. Timing during Translation Matters: Synonymous Mutations in Human Pathologies Influence Protein Folding and Function. *Biochem. Soc. Trans.* **2018**, *46*, 937–944. [CrossRef]
83. Fu, J.; Dang, Y.; Counter, C.; Liu, Y. Codon Usage Regulates Human KRAS Expression at Both Transcriptional and Translational Levels. *J. Biol. Chem.* **2018**, *293*, 17929–17940. [CrossRef] [PubMed]
84. Savas, S. Useful Genetic Variation Databases for Oncologists Investigating the Genetic Basis of Variable Treatment Response and Survival in Cancer. *Acta Oncol.* **2010**, *49*, 1217–1226. [CrossRef]
85. Horjales, S.; Cota, G.; Señorale-Pose, M.; Rovira, C.; Román, E.; Artagaveytia, N.; Ehrlich, R.; Marín, M. Translational Machinery and Protein Folding: Evidence of Conformational Variants of the Estrogen Receptor Alpha. *Arch. Biochem. Biophys.* **2007**, *467*, 139–143. [CrossRef] [PubMed]

86. Fernández-Calero, T.; Astrada, S.; Alberti, Á.; Horjales, S.; Arnal, J.F.; Rovira, C.; Bollati-Fogolín, M.; Flouriot, G.; Marin, M. The Transcriptional Activities and Cellular Localization of the Human Estrogen Receptor Alpha Are Affected by the Synonymous Ala87 Mutation. *J. Steroid Biochem. Mol. Biol.* **2014**, *143*, 99–104. [CrossRef] [PubMed]
87. Hertz, D.L.; Henry, N.L.; Kidwell, K.M.; Thomas, D.; Goddard, A.; Azzouz, F.; Speth, K.; Li, L.; Banerjee, M.; Thibert, J.N.; et al. ESR1 and PGR Polymorphisms are Associated with Estrogen and Progesterone Receptor Expression in Breast Tumors. *Physiol. Genom.* **2016**, *48*, 688–698. [CrossRef]
88. AlFakeeh, A.; Brezden-Masley, C. Overcoming Endocrine Resistance in Hormone Receptor–Positive Breast Cancer. *Curr. Oncol.* **2018**, *25*, 18. [CrossRef]
89. Singh, K.; Munuganti, R.; Lallous, N.; Dalal, K.; Yoon, J.; Sharma, A.; Yamazaki, T.; Cherkasov, A.; Rennie, P. Benzothiophenone Derivatives Targeting Mutant Forms of Estrogen Receptor-α in Hormone-Resistant Breast Cancers. *Int. J. Mol. Sci.* **2018**, *19*, 579. [CrossRef]
90. Furman, C.; Hao, M.-H.; Prajapati, S.; Reynolds, D.; Rimkunas, V.; Zheng, G.Z.; Zhu, P.; Korpal, M. Estrogen Receptor Covalent Antagonists: The Best is Yet to Come. *Cancer Res.* **2019**, *79*, 1740–1745. [CrossRef]
91. Lin, X.; Xiang, H.; Luo, G. Targeting Estrogen Receptor α for Degradation with PROTACs: A Promising Approach to Overcome Endocrine Resistance. *Eur. J. Med. Chem.* **2020**, *206*, 112689. [CrossRef]
92. Revankar, C.M.; Bologa, C.G.; Pepermans, R.A.; Sharma, G.; Petrie, W.K.; Alcon, S.N.; Field, A.S.; Ramesh, C.; Parker, M.A.; Savchuk, N.P.; et al. A Selective Ligand for Estrogen Receptor Proteins Discriminates Rapid and Genomic Signaling. *Cell Chem. Biol.* **2019**, *26*, 1692–1702.e5. [CrossRef]
93. Huang, W.; Peng, Y.; Kiselar, J.; Zhao, X.; Albaqami, A.; Mendez, D.; Chen, Y.; Chakravarthy, S.; Gupta, S.; Ralston, C.; et al. Multidomain Architecture of Estrogen Receptor Reveals Interfacial Cross-Talk between Its DNA-Binding and Ligand-Binding Domains. *Nat. Commun.* **2018**, *9*, 3520. [CrossRef] [PubMed]
94. Rakha, E.A.; Pareja, F.G. New Advances in Molecular Breast Cancer Pathology. *Semin. Cancer Biol.* **2020**, in press. [CrossRef] [PubMed]
95. Jeselsohn, R.; De Angelis, C.; Brown, M.; Schiff, R. The Evolving Role of the Estrogen Receptor Mutations in Endocrine Therapy-Resistant Breast Cancer. *Curr. Oncol. Rep.* **2017**, *19*, 35. [CrossRef] [PubMed]
96. Turner, N.C.; Kingston, B.; Kilburn, L.S.; Kernaghan, S.; Wardley, A.M.; Macpherson, I.R.; Baird, R.D.; Roylance, R.; Stephens, P.; Oikonomidou, O.; et al. Circulating Tumour DNA Analysis to Direct Therapy in Advanced Breast Cancer (PlasmaMATCH): A Multicentre, Multicohort, Phase 2a, Platform Trial. *Lancet Oncol.* **2020**, *21*, 1296–1308. [CrossRef]
97. Gilles, A.; Frechin, L.; Natchiar, K.; Biondani, G.; Loeffelholz, O.; von Holvec, S.; Malaval, J.-L.; Winum, J.-Y.; Klaholz, B.P.; Peyron, J.-F. Targeting the Human 80S Ribosome in Cancer: From Structure to Function and Drug Design for Innovative Adjuvant Therapeutic Strategies. *Cells* **2020**, *9*, 629. [CrossRef]
98. Reis, S.D.; Pang, Y.; Vishnu, N.; Voisset, C.; Galons, H.; Blondel, M.; Sanyal, S. Mode of Action of the Antiprion Drugs 6AP and GA on Ribosome Assisted Protein Folding. *Biochimie* **2011**, *93*, 1047–1054. [CrossRef]
99. Hunt, R.C.; Simhadri, V.L.; Iandoli, M.; Sauna, Z.E.; Kimchi-Sarfaty, C. Exposing Synonymous Mutations. *Trends Genet.* **2014**, *30*, 308–321. [CrossRef]

UNIVERSITÉ DE RENNES 1

Titre : Mécanismes d'échappement des cellules tumorales mammaires au contrôle hormonal : étude de l'impact du biais d'usage des codons

Mots clés : cancer du sein, ERα, résistance hormonale, usage des codons, traduction

Résumé : Le cancer du sein est le cancer le plus diagnostiqué dans le monde, et le plus meurtrier chez la femme. 70% des cas sont caractérisés par l'expression du récepteur aux œstrogènes ERα, dont l'activation stimule la prolifération cellulaire. Une modification de l'expression ou fonction d'ERα est donc fréquemment impliquée dans la tumorigenèse mammaire. En cas de cancer du sein ERα-positif, le traitement de choix est alors l'hormonothérapie, mais une résistance au traitement se déclare chez 30% des patientes. Une nouvelle hypothèse pouvant expliquer les modifications de fonction d'ERα menant à cet échappement au contrôle hormonal est que la conformation du récepteur serait modifiée au cours de sa traduction en cellules tumorales. Afin de mieux comprendre ce mécanisme, les travaux réalisés au cours de cette thèse ont alors exploré le rôle de l'usage des codons dans la régulation de la traduction d'ERα. Il est apparu que l'utilisation de codons synonymes avait en effet un impact sur l'expression et la fonction du récepteur, en fonction du niveau de différentiation des cellules ou lors du cycle cellulaire. Plus précisément, nous avons identifié un rôle joué par la nature de la troisième base des codons dans la régulation de l'expression des gènes en protéines au cours du cycle cellulaire. Enfin, une dynamique de traduction préférentielle par des monosomes ou polysomes a été mise en évidence en fonction de la composition en codons des messagers et de l'état cellulaire. Ces travaux lèvent alors le voile sur un nouveau niveau de régulation de la traduction, pouvant jouer un rôle dans l'expression et la fonctionnalité des protéines en fonction de l'état cellulaire, et ainsi être impliqué dans l'échappement des cellules tumorales mammaires au contrôle hormonal.

Title: Hormonal escape mechanisms of breast cancer cells: study of the impact of codon usage bias

Keywords: breast cancer, ERα, endocrine resistance, codon usage, translation

Abstract: Breast cancer is the most common cancer and the deadliest among women worldwide. 70% of the cases express the estrogen receptor ERα, whose activation stimulates cell proliferation. The modification of ERα expression or function leads then often to breast cancer development. In case of ERα-positive breast cancer, endocrine therapy is available but a resistance occurs for 30% of the patients. A new hypothesis to explain the modifications of ERα functions leading to this hormonal escape is that the conformation of the receptor could be modified during its translation in cancerous cells. To better understand this mechanism, experiments were performed to investigate the role played by codon usage in the regulation of ERα translation. It appeared that using synonymous codons indeed impacted ERα expression and function, depending on the differentiation status of cells or cell cycle phase. More precisely, we identified a role played by the third base of codons in the regulation of gene to protein expression during cell cycle. Finally, a dynamics of preferential translation by monosomes or polysomes was observed depending on the codons composing messengers and cell state. This project highlights then a new level in the regulation of translation that could play a role in the expression and functional properties of proteins depending on cell state, which could be involved in the hormonal escape of breast cancer cells.