



HAL
open science

Etude structurale du complexe CCR4-NOT

Jérôme Basquin

► **To cite this version:**

Jérôme Basquin. Etude structurale du complexe CCR4-NOT. Biologie structurale [q-bio.BM]. Université de Strasbourg, 2015. Français. NNT : 2015STRAJ084 . tel-01394684

HAL Id: tel-01394684

<https://theses.hal.science/tel-01394684>

Submitted on 9 Nov 2016

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 STRASBOURG

ÉCOLE DOCTORALE Sciences de la vie et de la santé
UPR 9002 – Architecture et réactivité de l'ARN

THÈSE

présentée par :

Jérôme BASQUIN

soutenue le : **21 décembre 2015**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : **Aspects moléculaires et cellulaires de la biologie**

Etude structurale du complexe CCR4-NOT

Thèse dirigée par :

Mr Claude SAUTER

Dr, CNRS, IBMC, Strasbourg

Président du Jury :

Mr Emmanuel CAILLAUD

Professeur, université de Strasbourg

Rapporteurs :

Mme Pascale ROMBY

Dr, CNRS, IBMC, Strasbourg

Mr Christophe MULLER

Dr, EMBL, Heidelberg

Mr Herve LE HIR

Dr, CNRS, ENS, Paris

Mr Patrick SCHULTZ

Dr, CNRS, IGBMC, Strasbourg

Mr Vincent MIKOL

Dr, SANOFI, Vitry sur seine

Superviseur VAE :

Mme Danielle HAUG

Directrice service VAE, université de Strasbourg



Service Validation des Acquis de l'Expérience

Année Universitaire 2013 / 2014

DOSSIER DE DEMANDE DE VALIDATION DES ACQUIS DE L'EXPÉRIENCE

Loi n° 2002-72 du 17 janvier 2002 de modernisation sociale et Décret n° 2013-756 du 19 août 2013 relatif aux dispositions réglementaires des livres VI et VII du code de l'éducation.

CANDIDAT(E)

NOM d'usage : BASQUIN..... Prénom : Jérôme

Diplôme postulé: Doctorat mention Sciences de la Vie et de la Santé spécialité Aspects moléculaires et cellulaires de la biologie.....

Déclaration sur l'honneur

Je déclare sur l'honneur que toutes les informations fournies dans ce dossier sont exactes.

J'ai été rendu attentif au fait qu'agrémenter mon travail de citations en omettant d'en citer les sources représente un acte de plagiat. Le plagiat est une atteinte au droit d'auteur et à la propriété intellectuelle.


La présente candidature à la validation des acquis de l'expérience en vue de l'obtention diplôme postulé, précité, constitue l'unique demande pour ce diplôme pour la même année civile.

Je m'engage également à ne pas présenter plus de trois candidatures à la validation des acquis de l'expérience pour des diplômes, certificats ou titres différents durant la présente année civile.

Fait à Martinsried _____

le 15.09.2015 _____

Signature du candidat



La loi punit quiconque se rend coupable de fausses déclarations.

« Constitue un faux toute altération frauduleuse de la vérité, de nature à causer un préjudice et accomplie par quelque moyen que ce soit, dans un écrit ou tout autre support d'expression de la pensée qui a pour objet ou qui peut avoir pour effet d'établir la preuve d'un droit ou d'un fait ayant des conséquences juridiques. Le faux et l'usage de faux sont punis de trois ans d'emprisonnement et de 45000 euros d'amende (code pénal, art. 441-1)

Le fait de se faire délivrer indûment par une administration publique ou par un organisme chargé d'une mission de service public, par quelque moyen frauduleux que ce soit, un document destiné à constater un droit, une identité ou une qualité ou à accorder une autorisation, est puni de deux ans d'emprisonnement et de 30000 euros d'amende » (code pénal art. 441-6)

Remerciements

Je tiens tout d'abord à remercier l'ensemble des membres du jury d'avoir accepté d'évaluer ce travail: XXXX. Je souhaite remercier Mme Danielle HAUG pour m'avoir guidé dans ma démarche de Validation des Acquis de l'Expérience pour l'obtention d'un doctorat ainsi que pour son accompagnement tout au long de la rédaction de ce manuscrit.

La démarche de l'obtention du titre de doctorat a toujours représenté à mes yeux une quête du Graal. Ma démarche a été motivée par des rencontres tout au long de ma carrière et cette envie a pu être finalement concrétisée par ce manuscrit. Les compétences que j'ai pu compiler au fur et à mesure des années, je les dois à de nombreuses personnes que je souhaite remercier.

Je dois beaucoup à Jean-Yves Crenne, qui m'a accueilli dans son laboratoire au début de ma carrière. Grace à toi, j'ai pu accueillir une solide expérience dans l'expression et la purification des protéines recombinantes. Tu as su m'insuffler la curiosité et la rigueur. Tu m'as toujours accompagné et encouragé dans les évolutions successives de ma carrière. De plus, j'ai beaucoup appris sur les méthodes managériales d'un groupe de recherche en t'observant. Mon plan de carrière était de te rejoindre à un moment où un autre, malheureusement le destin en a décidé autrement. Je ne t'oublierai jamais !

L'EMBL est un endroit propice aux rencontres formatrices. J'ai eu la chance de croiser la route de nombreux talents dans cet environnement international. Je souhaite remercier Renaud Vincentelli pour tout ce qu'il m'a apporté durant ma période EMBL (organisation, automatisation du laboratoire) ainsi que les aspects de communication en interne. Il a su trouver les mots au moment où je devais gérer une situation personnelle délicate. J'ai toujours apprécié son franc parler et son bagou. Par ailleurs Renaud a écrit sa thèse comme moi après de nombreuses années d'expérience et j'ai été quelque peu inspiré par sa démarche. Renaud, le Schumli et le Schoko croissant de 9H00 de l'EMBL à tes cotés est un manque que je n'ai toujours pas réussi à combler.

Si je devais remercier Claude Sauter à sa juste contribution dans mon parcours, je devrais consacrer un chapitre à cette thèse. Je connais Claude depuis 16 ans. Nous sommes arrivés à l'EMBL à la même période dans le même groupe de recherche où la situation nous permis de travailler ensemble de façon indépendante. Ces années sont précieuses à mes yeux. J'ai appris énormément avec toi. Maintenant je suis aussi Maître du Crystal grâce à ton enseignement. Nous avons passé des nuits de désillusion au synchrotron forcément cela crée des liens. J'ai fait mes

premiers pas de cristallographe à tes côtés. Je tiens également à te remercier pour ton accompagnement scientifique dans ma démarche de Validation des Acquis de l'Expérience pour l'obtention d'un doctorat. Mon manuscrit a pris de la valeur grâce à tes conseils. Je te serais éternellement reconnaissant pour ton aide et le partage de ton expertise.

J'ai aussi eu la chance de rencontrer Eric Ennifar à l'EMBL. J'ai beaucoup appris à ses côtés dans la résolution de structures macromoléculaires. Nous avons beaucoup relativisé ensemble face à la « why don't you try » attitude. Eric prépare toi, je veux fêter la concrétisation de mon projet en courant un trail avec toi !

C'est aussi à l'EMBL que j'ai rencontré Elena Conti. Cette rencontre a été déterminante pour moi. Elena est maintenant devenue un véritable mentor pour moi. Elle m'a fait confiance, m'a offert la possibilité de créer une plateforme technologique en me laissant carte blanche. D'autre part, elle m'a encouragé à développer des activités de recherches en parallèle. Grâce à sa confiance, je gère maintenant un projet de recherche. Elena m'a aussi soutenu dans ma démarche de VAE.

Je souhaite remercier Vincent Oliéric avec qui j'ai pu me perfectionner dans la collecte de données cristallographiques. A ton contact, j'ai eu l'occasion de peaufiner ma stratégie de collecte et je suis maintenant capable de tirer toute la quintessence de mes cristaux sur les lignes du SLS. C'est toujours un plaisir de te voir quand je viens au SLS. De plus, notre passion commune pour l'œnologie (devrais je dire la dégustation) ne manquera pas de nous réunir à nouveau.

Je tiens à remercier mon équipe plateforme : Karina, Sabine et Ariane qui contribuent grandement au succès de l'activité, tout en adoptant une attitude ouverte et agréable. Je tiens à remercier Walter avec qui je démonte et répare régulièrement de nombreux instruments scientifiques afin d'éviter des coûts de maintenance prohibitifs.

Je tiens à remercier Varun Bhaskar et Sevim Özgür avec lesquels j'ai eu le privilège de travailler sur le projet CCR4-NOT. Malgré une compétition ardue avec un autre laboratoire, notre groupe a su rester valeureux et focalisé afin de produire les résultats scientifiques dans des temps courts tout en ayant à cœur de maintenir un esprit convivial.

Esben Lorentzen a aussi été déterminant dans ma carrière et formation. J'ai eu l'occasion de travailler à ses côtés sur le projet Rrp44. J'ai beaucoup appris à cette occasion sur le phasage et le traçage de modèle dans des cartes de densité à basse résolution. Il m'a conseillé à de nombreuses occasions pour des cas de phasage difficile. J'ai su apprécier au fil des années son humour décalé et décapant. J'observe avec admiration la façon dont il gère son groupe de recherche. Je lui

souhaite beaucoup de succès dans la poursuite de sa carrière.

Finalement, je remercie de tout mon cœur Claire, avec qui je serpente les sentiers escarpés de la vie. Tu as toujours été là pour me soutenir dans cet effort et m'a encouragé dans les moments de doute au cours de la rédaction de ce manuscrit. Merci aussi pour les corrections et l'aide pour la mise en page. Merci infiniment pour tout le reste.

A Jean Yves

A mon fils Arno

A Claire

A mes parents

TABLE DES MATIERES

Déclaration sur l'honneur	2
REMERCIEMENTS	3
ABREVIATIONS	8
ETUDE STRUCTURALE DU COMPLEXE.....	9
CCR4-NOT	9
A. INTRODUCTION GENERALE	10
B. DECOUVERTE DU COMPLEXE CCR4-NOT	11
C. LES FONCTIONS BIOLOGIQUES DU COMPLEXE CCR4-NOT	12
C.1. Le complexe CCR4-NOT est responsable de la déadénylation des ARN	12
C.2. Les activités biochimiques principales du complexe CCR4-NOT	12
C.3. Le complexe CCR4-NOT et la régulation de la transcription	15
C.4. Le complexe CCR4-NOT régule l'équilibre entre synthèse et dégradation des ARNs messagers.....	17
C.5. Le complexe CCR4-NOT est un régulateur de la synthèse des ARN	18
C.6. Le complexe CCR4-NOT a un rôle capital dans la relation entre traduction et dégradation des ARN messagers	19
C.7. Rôle dans la répression de la traduction	19
C.8. Rôle dans la fidélité de la traduction	20
C.9. CCR4-NOT et réponse immunitaire innée	21
C.9.1. Implication du complexe CCR4-NOT dans la régulation du signal en réponse aux interférons.....	21
C.9.2. Le complexe CCR4-NOT affecte le complexe majeur d'histocompatibilité	22
C.10. Mécanismes de sélection de l'ARN messager	22
D. ARCHITECTURE ET ANALYSE STRUCTURALE DU COMPLEXE CCR4-NOT ET DES SOUS UNITES QUI LE COMPOSENT 24	
D.1. Les déadénylases du complexe CCR4-NOT	24
D.1.1. CAF1 : une déadénylase de type DEDD.....	24
D.1.2. CCR4 : une déadénylase de type EEP	26
D.2. Le module de déadénylation du complexe CCR4-NOT	27
D.2.1. Le domaine NOT1 (754-1000) adopte un repliement MIF4G.....	27
D.2.2. L'interaction du domaine NOT1 (754-1000) avec CAF1 implique des résidus conservés	28
D.2.3. L'interaction de CAF1 avec le domaine LRR de CCR4 forme un cœur hydrophobe	29
D.2.4. Le domaine nucléase de CCR4 est adjacent au domaine LRR	30
D.2.5. Mutations entraînant la déstabilisation de NOT1-CAF1 et CAF1-CCR4.....	32
D.3. Le domaine NOT1 (154-754) est une plateforme étendue de répétitions HEAT.....	33
D.4. Le module NOT	34
D.5. Structure CNOT9 et CNOT1	39
D.6. Le domaine central de CNOT1 interagit directement avec la protéine DDX6 à boîte DEAD	40
D.6.1. La structure du complexe CNOT1 MIF4G DDX6 explique le mode d'interaction spécifique et conservé	41
D.7. Structure de TTP en complexe avec CNOT1	42
D.8. Structure de NOT4.....	43
D.9. Structure de NOT1 en complexe avec NOT4.....	45

D.9.1.	NOT4c entoure les répétitions HEAT N-terminales de NOT1c	46
E.	CONTRIBUTION DE MON TRAVAIL DE THESE	47
E.1.	Etudes fonctionnelles et structurales du module nucléase du complexe CCR4-NOT	49
E.1.1.	Contexte et objectifs du projet	49
E.1.2.	Approches expérimentales	49
E.1.3.	Résumé de l'étude	53
E.2.	Détermination de la structure du module NOT1-NOT2-NOT5 du complexe Ccr4-Not de levure	87
E.2.1.	Contexte et objectifs du projet	87
E.2.2.	Approches expérimentales	87
E.2.3.	Résumé de l'étude	89
E.3.	Caractérisation biochimique et structurale du rôle du complexe CCR4-NOT et ATPase DDX6 dans la répression des microARN.	112
E.3.1.	Contexte et objectif de l'étude	112
E.3.2.	Approches expérimentales	112
E.3.3.	Résumé de l'étude	114
E.4.	Architecture du module d'ubiquitination du complexe CCR4-NOT	161
E.4.1.	Contexte et objectif de l'étude	161
E.4.2.	Approches expérimentales	161
E.4.3.	Résumé de l'étude	162
F.	CONCLUSIONS ET PERSPECTIVES	176
G.	REFERENCES	179

Abreviations

ADN : Acide désoxyribonucléique
AFMB : Architecture et Fonction des Macromolécules Biologiques
ALA : Association for Lab Automation
AMP : Adenosine 5' monophosphate
ARE: AU-rich éléments)
ARM: Armadillo-repeat protein
ARN: Acide ribonucléique
ARN: Polymérase II (RNAPII)
ARNi: ARN interférents
ARNm: ARN messagers
CDE: Constitutive decay element
CNRS: Centre National de la Recherche Scientifique
EMBL: European Molecular Biology Laboratory
EMBO European Molecular Biology Organisation
GLP: Good Laboratory Practice
GMP Good Manufacturing Practice
IBMC: International Conference of Crystallisation of Biological Macromolecules
IFN: Interférons
IGBMC: Institut de génétique et de biologie moléculaire et cellulaire
ITC: Isothermal titration calorimetry
LRR: Leucine Rich Repeat
MAD: Multiwavelength Anomalous Dispersion
MIF4G: Middle domain of eukaryotic initiator factor 4G
MNHN: Muséum National d'Histoire Naturelle
MR: Molecular Replacement
mRNP: Messenger ribonucleoprotein
NAC: Nascent-polypeptide-associated peptide
NMD: Non sense Mediated Decay
NOT: Negative On TATA-less
PB: P-bodies
PAPB: protéine se liant à la queue poly(A)
PDB: Protein Data Bank
PTC : Premature Termination Codon
RAMC: Recent Advances of Macromolecular Crystallisation
SAD: Single-wavelength Anomalous Diffraction
SAXS Small-angle X-ray scattering
SLS Swiss Light Source
TAF: TBP associated factors
TBP: TATA-binding protein
TTP: protéine tristétraproline
URA: unité de recherche associée
VAE : Validation des Acquis de l'expérience

A. Introduction générale

Les acides ribonucléiques (ARN) sont des molécules biologiques trouvées dans pratiquement tous les organismes vivants. Ils se présentent comme un enchaînement linéaire de nucléotides, chaque nucléotide étant constitué d'un groupement phosphate, d'un sucre (ribose) et d'une base azotée (adénine, guanine, cytosine et uracile). Les cellules vivantes les utilisent comme un support intermédiaire des gènes pour fabriquer les protéines dont elles ont besoin. Les ARN sont produits par transcription à partir de l'ADN support de l'information génétique et situé dans le noyau des cellules eucaryotes. Chaque ARN est une copie de l'un des brins d'une région d'ADN. Les ARN ainsi produits assurent trois grands types de fonctions, 1) ils peuvent être un support de l'information génétique d'un ou plusieurs gènes codant des protéines (ARN messagers), 2) ils peuvent adopter une structure secondaire et tertiaire stable pour accomplir des fonctions catalytiques et l'on parle alors de ribozymes (par exemple la ribonucléase P ou l'ARN ribosomique), 3) ils peuvent enfin intervenir dans la régulation de l'expression génique, soit en s'appariant directement à un ARNm cible dans le cas des petits ARN régulateurs et ARN antisens, soit en agissant après une étape de maturation par des facteurs multi-protéiques dans le cas des petits ARN interférents et des micro ARN (Berg et al. 2015). Des études sur la production cellulaire d'ARN ont montré que les cellules fabriquaient plus de molécules d'ARN qu'elles n'en accumulaient, suggérant l'existence de processus de dégradation (Houseley and Tollervey 2009). En général, ils sont détruits dès la fin de leur utilisation par la cellule. Les molécules d'ARN comportant des défauts d'épissage ou des problèmes de repliement seront détruites par la machinerie de surveillance. Comme la dégradation des ARN est un processus clé de la vie cellulaire, ce mécanisme doit être finement contrôlé afin d'identifier les ARN cibles. La voie de dégradation est largement contrôlée par l'action coordonnée et successive de complexes macromoléculaires multi-protéiques.

Le complexe CCR4–NOT est une machine macromoléculaire multifonctionnelle extrêmement conservée et impliquée dans divers processus cellulaires, en particulier dans le contrôle du métabolisme des ARN messagers (Collart and Panasenko 2012). De nombreuses études menées, dans un premier temps chez la levure et plus récemment chez les eucaryotes supérieurs, ont mis en évidence que le complexe régule l'expression des gènes à de multiples niveaux ce qui influe sur l'initiation de la transcription, le contrôle de la taille de l'ARN messager par déadénylation qui provoque la dégradation de l'ARN messager, la traduction et même la dégradation des protéines.

Les composants du complexe CCR4–NOT sont par ailleurs associés à diverses fonctions dans différentes espèces, soit dans le noyau, soit dans le cytoplasme. Ceci inclut en particulier la réparation de l'ADN et la méthylation des histones chez la levure, la régulation de la longueur des microtubules chez *Caenorhabditis elegans* et la spermatogénèse chez la souris (Chapat and Corbo 2014).

B. Découverte du complexe CCR4-NOT

Le complexe CCR4-NOT a été découvert et étudié chez la levure. La majorité des composants de ce complexe a été identifié par une sélection génétique dont le but était de mettre en évidence des activateurs de la transcription. Il a été dans un premier temps présumé que le complexe était composé de deux sous-complexes indépendants : CCR4 et NOT. Le module CCR4 (Carbon Catabolite Repressor) est composé des protéines CCR4 et CAF1. CCR4 a été identifié comme un gène régulant les enzymes répressibles au glucose. CCR4 interagit avec la protéine CAF1 (CCR4-associated factor 1). Les protéines NOT (Negative On TATA-less) ont été identifiées grâce à un criblage de gènes réprimant les promoteurs sans boîte TATA. En 1998, Clyde Denis et ses collègues ont publié une communication mettant en évidence que le cœur du complexe était composé des modules CCR4 et NOT. Cette étude était principalement basée sur des expériences de co-immunoprécipitation et de spectrométrie de masse révélant un complexe chez la levure d'une taille allant de 0.9 à 1.2 MDa (Liu et al. 1998b). Au total, le complexe de levure est constitué de cinq protéines NOT (NOT1 à NOT5) et deux sous-unités catalytiques, CCR4 et CAF1.

Le cœur du complexe CCR4–NOT a été extrêmement conservé au cours de l'évolution chez les organismes eucaryotes (Draper et al. 1994; Albert et al. 2000). Chez les mammifères, les protéines homologues forment aussi un complexe similaire qui a un rôle important dans la régulation de plusieurs processus cellulaires (Mahadevan and Struhl 1990; Collart et al. 1993; 1994; Oberholzer et al. 1998; Badarinarayana et al. 2000; Albert et al. 2002; Deluen et al. 2002; Lenssen et al. 2002; Maillet et al. 2002; Viswanathan et al. 2004). D'autres sous-unités périphériques ont été identifiées par la suite et se composent des protéines CNOT9, CAF130, NOT10 (CNOT10), C2orf29 (CNOT11), and TAB182. Certaines de ces sous-unités sont espèces-spécifiques : par exemple CAF130 est spécifique à la levure (Collart and Panasenko 2012), alors que les protéines NOT10, C2orf29, et TAB182 se trouvent dans le complexe humain et sont absentes chez la levure. Par ailleurs, NOT10 and C2orf29 sont conservées chez les métazoaires (Bawankar et al. 2013; Mauxion et al. 2013).

C. Les fonctions biologiques du complexe CCR4-NOT

C.1. Le complexe CCR4-NOT est responsable de la déadénylation des ARN

Ce complexe est responsable du processus de déadénylation, qui consiste à catalyser la dégradation de la queue poly(A) des ARN. Cet événement a pour conséquences de réprimer la traduction et de marquer l'ARN messager pour le processus de dégradation (Collart 2003). La déadénylation constitue un processus crucial notamment dans des événements impliquant l'expression des gènes et qui a un rôle dans des processus biologiques variés tel que le développement embryonnaire. Cette réaction de déadénylation est une étape biphasique (Yamashita et al. 2005; Wahle and Winkler 2013). La première phase est une réaction synchronisée par étape qui consiste à raccourcir la queue poly(A) de l'ARN messager d'une longueur avoisinant 110 nucléotides. Dans une seconde phase, déterminante dans le déclenchement de la dégradation de l'ARN, la queue devient plus hétérogène en longueur, pouvant varier d'une centaine à une vingtaine de ribonucléotides. Plusieurs travaux ont mis en évidence ce processus biphasique et l'hypothèse consensuelle avancée à ce jour propose que ce processus est le résultat d'une action séquentielle du complexe de déadénylation PAN2–PAN3 qui domine dans la première phase, alors que le complexe CCR4–NOT serait prédominant dans la seconde (Figure 1) (Yamashita et al. 2005; Zheng et al. 2008; Wahle and Winkler 2013).

C.2. Les activités biochimiques principales du complexe CCR4-NOT

Deux activités biochimiques majeures sont associées au complexe CCR4-NOT, la déadénylation et l'activité E3 ubiquitine ligase. La première limite la vitesse de dégradation d'un ARN messager normal par les voies 5'→3' ou 3'→5'. Des études chez la levure ont montré que le complexe CCR4-NOT était principalement responsable de cette activité *in vivo* (Tucker et al. 2002). Les protéines responsables de l'activité de déadénylation sont les protéines CCR4 et CAF1. CCR4 appartient à la classe des exonucléases de type III qui possèdent une activité exoribonucléasique 3'-5' pour les ARNs poly(A) *in vitro*. CAF1 contient un domaine RNase D/DEDD et possède une activité 3' exonucléasique (Arraiano et al. 2013). Bien que CCR4 et CAF1 soient requises pour atteindre des taux de déadénylation normaux chez la levure, il a été montré que CCR4 était le composé majoritairement actif du complexe (Chen et al. 2002; Tucker et al. 2002). En effet chez *Saccharomyces cerevisiae*, la protéine CAF1 a des acides aminés substitués en son site actif (Berg

et al. 2015; Thore et al. 2003) qui sont indispensables à son activité, mais qui contribuent au recrutement de la protéine CCR4 sur le complexe. D'autres études plus récentes ont mis en évidence que chez les métazoaires les protéines CAF1 ont aussi conservé les acides aminés de la version de la levure. Les protéines CAF1 et CCR4 sont actives sous forme isolée *in vitro* (Houseley and Tollervey 2009; Chen et al. 2002; Thore et al. 2003; Bianchin et al. 2005; Clark et al. 2004) mais *in vivo* la déadénylation semble se produire dans le cadre du complexe où la protéine d'échafaudage Not1 est indispensable. L'identification d'homologues de ces protéines chez les métazoaires a démontré que la voie de dégradation des ARNs est un processus conservé au cours de l'évolution (Collart and Panasenko 2012; Bianchin et al. 2005; Bartlam and Yamamoto 2010; Denis et al. 2003; Dupressoir et al. 2001; Temme et al. 2004; Viswanathan et al. 2004).

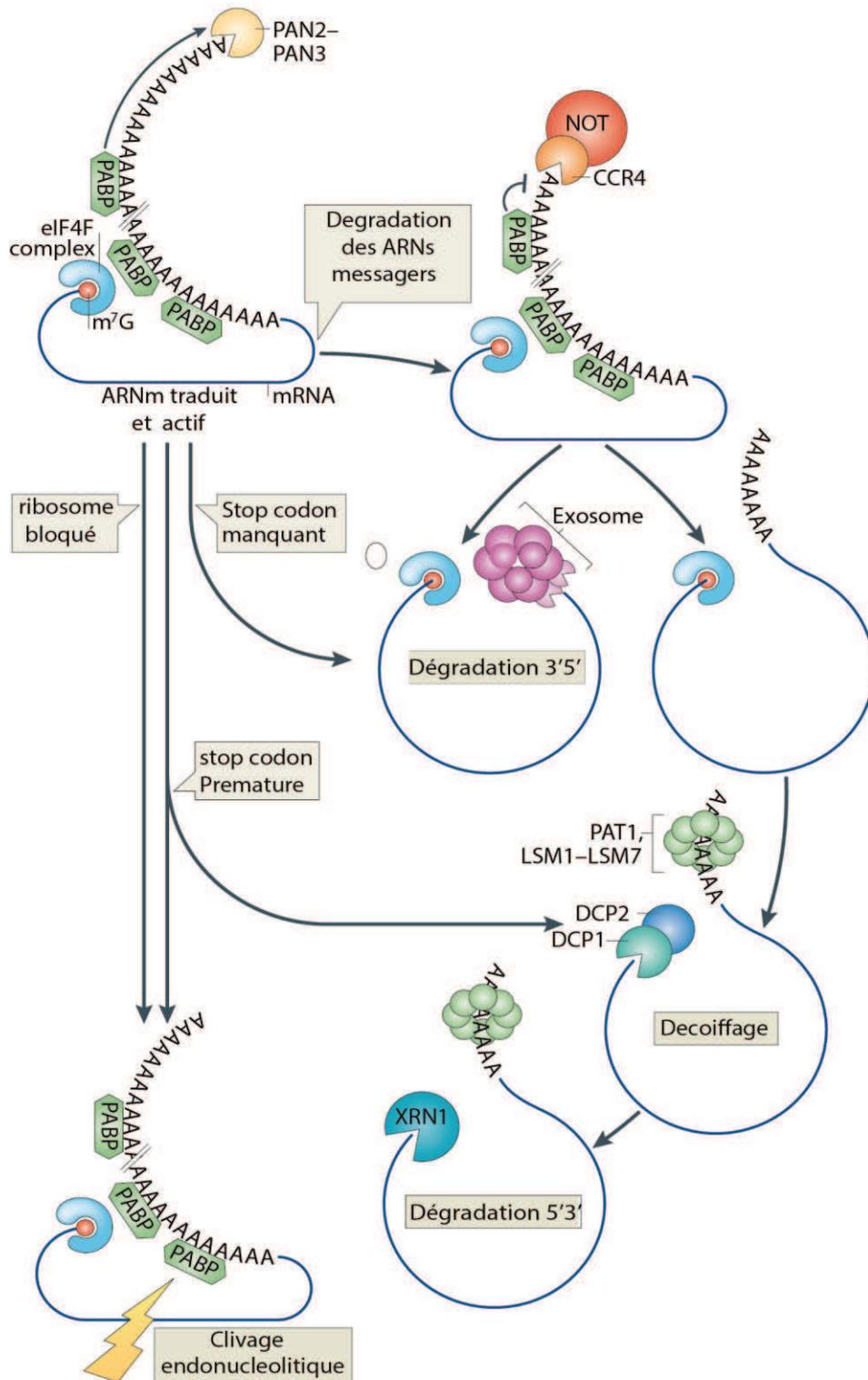


Figure 1 : Déadénylation d'un ARNm cytoplasmique, dégradation et contrôle qualité
Adapté de Norbury, C. J. (2013).

Un ARNm est représenté dans son état translationnel et polyadénylé. Par souci de clarté, les facteurs d'initiation de la traduction autres que le complexe eIF4F ne sont pas représentés. Un raccourcissement de la queue poly(A) est d'abord effectué par le complexe PAN2-PAN3 qui est stimulé par PABP. La dégradation s'effectue ensuite par une déadénylation par le complexe CCR4-NOT. Cette étape de déadénylation est soit suivie par la dégradation 3'-5' par le complexe Exosome ou alors plus fréquemment par le recrutement du complexe PAT1-LSM1-LSM7 qui stimule le décoiffage par les protéines (DCP1)-DCP2 et qui induit l'attaque exonucléasique de

XRN1. L'absence de codon stop court-circuite cette voie de dégradation en stimulant la voie 3'-5' de dégradation par l'exosome (non stop decay) ou les ribosomes bloqués déclenchent le clivage endonucléotidique des ARNm (no-go-decay). La présence d'un codon stop prématuré conduit aussi au clivage endonucléotidique et à la déadénylation et au décoiffage (NMD:nonsense-mediated decay).

Cytoplasmic RNA: a case of the tail wagging the dog. *Nature Reviews. Molecular Cell Biology*, 14(10), 643–653.

La seconde activité enzymatique liée au complexe décrite *in vitro* est l'ubiquitination. L'ubiquitination est une modification post-traductionnelle biochimique nécessitant plusieurs étapes de modifications successives par des ligases de type E pour aboutir à la fixation covalente d'une ou de plusieurs protéines d'ubiquitine (8 kDa) sur une ou plusieurs lysines acceptrices de la protéine substrat. Ces modifications biochimiques ont plusieurs fonctions dont la plus connue est la dégradation de la protéine ubiquitinée par le protéosome. La protéine NOT4 possède une activité ligase ubiquitine RING E3 conservée de la levure aux métazoaires (Draper et al. 1994; Albert et al. 2002; 2000; Mersman et al. 2009; Mulder et al. 2007). Les enzymes E2 qui fonctionnent avec la protéine NOT4, sont Ubc4 et Ubc5 (Mahadevan and Struhl 1990; Panasenکو et al. 2012; Collart et al. 1993; 1994; Oberholzer et al. 1998; Badarinarayana et al. 2000; Albert et al. 2002; Deluen et al. 2002; Lenssen et al. 2002; Maillet et al. 2002; Viswanathan et al. 2004), deux enzymes de conjugaison de l'ubiquitine qui sont impliquées dans des processus d'ubiquitination. Peu de substrats de l'enzyme E4 ont été décrits, on retrouve parmi eux la protéine ribosomal Rps7A47 et le complexe protéine chaperonne NAC (Nascent-polypeptide associated complex) dont l'ubiquitination régule l'association au ribosome et au protéosome (Collart and Panasenکو 2012; Panasenکو et al. 2009). D'autres cibles d'ubiquitination ont pu être identifiées chez la levure en particulier une enzyme de déméthylation des histones, la protéine Jhd2. L'ubiquitination de cette enzyme de modification de la chromatine conduit à sa dégradation par le protéosome. Bien que l'homologue humain de Jhd2 (JARID1C) soit ubiquitiné *in vitro* par la protéine CNOT4, il reste à confirmer que ces cibles d'ubiquitination sont conservées chez les métazoaires.

C.3. Le complexe CCR4-NOT et la régulation de la transcription

Le complexe CCR4-NOT a été décrit dans un premier temps comme étant un régulateur global de la transcription (Bawankar et al. 2013; Denis et al. 2003; Mauxion et al. 2013; Dupressoir et al.

2001; Temme et al. 2004; Viswanathan et al. 2004; Albert et al. 2002; Mersman et al. 2009; Mulder et al. 2009; Collart 2003) où CCR4 influait sur l'activation de la transcription, bien que les protéines NOT semblaient être des facteurs répresseurs des promoteurs sans boîte TATA. Des expériences d'interactions génétiques, confirmées par des mesures d'interactions physiques ont pu mettre en évidence que les sous-unités du complexe CCR4-NOT interagissaient avec des facteurs régulant la transcription tels que TBP (TATA-binding protein) (Collart 2003; Badarinarayana et al. 2000) et ses cofacteurs TAF (TBP associated factors) (Yamashita et al. 2005; Lemaire et al. 2000; Wahle and Winkler 2013; Deluen et al. 2002; Sanders et al. 2002), le complexe SAGA agissant comme une acétyltransférase d'histone et enfin le complexe SRB/médiateur (Yamashita et al. 2005; Benson et al. 1998; Zheng et al. 2008; Wahle and Winkler 2013). Toutes ces données ont permis d'élaborer un modèle où le complexe CCR4-NOT régule l'initiation de la transcription en modulant la fonction des complexes TBP/TFIID. La participation des sous-unités du complexe CCR4-NOT dans la régulation de la transcription a fait l'objet de revues détaillées (Tucker et al. 2002; Collart and Panasenko 2012; Miller et al. 2012). Chez les métazoaires, l'interaction du complexe CCR4-NOT dans la modulation de la transcription a été mise en évidence principalement par des études focalisées sur la sous-unité CNOT3. En effet de part sa proximité avec le promoteur et son rôle dans la transcription, CNOT3 semble être un régulateur très important de fonctions biologiques telles que l'hémostase rétinienne, la physiologie du cœur ou le renouvellement des cellules souches (Arraiano et al. 2013; Neely et al. 2010; Venturini et al. 2012). De plus, il semblerait que le complexe CCR4-NOT par l'intermédiaire des récepteurs nucléaires hormonaux chez les mammifères influence la régulation de la transcription. En effet les sous-unités CNOT6, CNOT9 et CNOT7 accroissent significativement l'activation de la transcription de nombreux récepteurs nucléaires, tels que RAR α , RXR β , TR α et Er α (Chen et al. 2002; Morel et al. 2003; Tucker et al. 2002; Prévôt et al. 2001; Garapaty et al. 2008). Par ailleurs la sous-unité CNOT1 réprime l'activation de la transcription du récepteur nucléaire Er α , en interagissant directement avec ce dernier et en recrutant les autres sous-unités du complexe sur les promoteurs régulant les œstrogènes (Bardwell et al. 2006).

Les protéines CNOT2 et CNOT3 possèdent également un domaine appelé NOT-box connu pour être un répresseur de la transcription. Le complexe de co-répression SMRT/NCoR-HDAC3 est suspecté d'agir sur la répression de la transcription en interagissant avec la sous-unité CNOT2. Le contrôle nucléaire de la transcription par le complexe CCR4-NOT a des répercussions indirectes au sein du cytoplasme. La délétion des protéines (NOT2, NOT5 et NOT4) chez la levure provoque l'accumulation de protéines agrégées dans la cellule (Berg et al. 2015; Halter et al. 2014). La perte de la protéine NOT4 qui exhibe une activité ubiquitine ligase E3 induit une accumulation

significative des protéines nouvellement synthétisées ainsi que des protéines poly-ubiquitinées (Houseley and Tollervey 2009; Panasenko et al. 2012; Badarinarayana et al. 2000; Lemaire et al. 2000; Deluen et al. 2002; Sanders et al. 2002; Benson et al. 1998; Neely et al. 2010; Venturini et al. 2012; Prévôt et al. 2001; Garapaty et al. 2008). Il n'est pas prouvé que la machinerie de transcription soit directement affectée chez les mutants dépourvus des sous-unités NOT, mais le phénotype observé conforte l'idée que l'influence significative de l'action du complexe CCR4-NOT au sein du cytoplasme puisse jouer un rôle sur la transcription dans le noyau. Cette idée est renforcée par le fait que le complexe CCR4-NOT contrôle la transcription d'un pool de gène en ciblant le facteur de transcription STAT1, qui joue un rôle important dans la transmission du signal interféron de type I (α/β) et de type II (γ) (Collart and Panasenko 2012; Bromberg and Darnell 2000) dans le cytoplasme (Draper et al. 1994; Chapat et al. 2013; Albert et al. 2000). La sous-unité humaine CAF1 interagit directement avec la fraction cytoplasmique de STAT1. Chez un mutant où a été supprimé hCAF1, on observe une activation de la transcription des gènes régulés par STAT1 et une décondensation locale de la structure de la chromatine sur leurs promoteurs et sites d'accrochages (Mahadevan and Struhl 1990; Chapat et al. 2013; Collart et al. 1993; 1994; Oberholzer et al. 1998; Badarinarayana et al. 2000; Albert et al. 2002; Deluen et al. 2002; Lenssen et al. 2002; Maillet et al. 2002; Viswanathan et al. 2004).

C.4. Le complexe CCR4-NOT régule l'équilibre entre synthèse et dégradation des ARNs messagers

Des études récentes ont démontré que la dégradation des ARN messagers dans le cytoplasme était coordonnée avec la transcription dans le noyau (Collart and Panasenko 2012; Dahan and Choder 2013). Le taux de synthèse et de dégradation des ARN messagers a évolué de manière inverse chez différentes espèces de levure (Bawankar et al. 2013; Dori-Bachash et al. 2011; Mauxion et al. 2013), comme le montre une étude où ont été mutés Rpb4 mais aussi certaines unités du complexe CCR4-NOT (NOT5, CCR4 et CAF1). L'existence de ce phénomène a été renforcée par une autre étude où la délétion des sous-unités responsables de la déadénylation augmente les taux de dégradation mais aussi décroît le taux de synthèse de l'ARN messager (Collart 2003; Sun et al. 2012). Ces données semblent indiquer que la synthèse et dégradation sont interconnectées par l'action de Rpb4 (sous unité de l'ARN polymérase II) et du complexe CCR4-NOT, qui régule l'équilibre entre ces deux processus. En cohérence avec cette observation une autre étude a mis en évidence que Rpb4 en complexe avec Rpb7 formaient un module de dissociation de l'ARN polymérase II (RNAPII) qui pourrait s'associer à l'ARNm nouvellement

synthétisé contribuant à sa traduction et sa dégradation dans le cytoplasme (Yamashita et al. 2005; Harel-Sharvit et al. 2010; Wahle and Winkler 2013; Lotan et al. 2005; Goler-Baron et al. 2008). Il est donc probable que les modules Rpb4/7 et le complexe CCR4-NOT agissent de manière concertée. Il a été établi que le complexe CCR4-NOT est un facteur qui stimule l'élongation chez la levure en influant sur les complexes RNA polymérase II (RNAPII). Cette étude montre également que l'association de RNAPII et du complexe CCR4-NOT ne nécessite pas la présence de Rpb4b (Yamashita et al. 2005; Kruk et al. 2011; Zheng et al. 2008; Wahle and Winkler 2013). Dans ce cas, il est possible que la régulation de la synthèse et de la dégradation des ARN messagers fasse intervenir le complexe CCR4-NOT dans le transfert de Rpb4/7 des complexes d'élongation sur les ARN messagers.

C.5. Le complexe CCR4-NOT est un régulateur de la synthèse des ARN

Parallèlement à l'activité de régulation de la transcription, le complexe semble être aussi impliqué dans des processus liés à la formation nucléaire ainsi qu'au contrôle qualité des particules ribonucléoprotéiques liées à l'ARN messager (mRNP). Une étude a ainsi montré que des mutations du complexe CCR4-NOT pouvaient mener à une accumulation de petits ARN nucléaires polyadénylés (Tucker et al. 2002; Azzouz et al. 2009). Ces transcrits se retrouvent normalement dégrader par les mécanismes de surveillance des ARN dont les acteurs principaux sont l'exosome et le complexe de polyadénylation (TRAMP) composé des protéines Mtr4/Trf4/Air2.

Plus récemment, il a été démontré que le complexe CCR4-NOT jouait aussi un rôle dans la voie de transformation des ARN messagers nucléaires. Une approche protéomique dans un système utilisant des cellules HeLa exprimant des sous-unités CNOT a révélé l'existence d'interaction entre ces CNOT et des facteurs liés à la transcription et à la stabilité des ARN, mais aussi avec des facteurs impliqués dans la machinerie de traitement des ARN, tels que des facteurs d'épissage ainsi que des protéines du complexe du pore nucléaire (Arraiano et al. 2013; Kerr et al. 2011). Des liens physiques et fonctionnels entre le complexe CCR4-NOT et le complexe de pore nucléaire ont pu être confirmés par une autre étude chez la levure (Chen et al. 2002; Kerr et al. 2011; Tucker et al. 2002). Cependant les mutants CCR4-NOT n'ayant pas d'effets observables sur le transport des ARN, ce lien devra être confirmé.

C.6. Le complexe CCR4-NOT a un rôle capital dans la relation entre traduction et dégradation des ARN messagers

Il est clairement établi que la protéine PAPB (poly A binding protein) se fixe sur la queue poly(A) des ARN messagers lors de leur transport dans le cytoplasme. PAPB a pour rôle de stabiliser les ARNm et de faciliter leur traduction. PAPB se fixe au facteur d'initiation de la traduction eIF4G, qui lui-même se fixe sur eIF4E (complexe se fixant à la coiffe). La circularisation des messagers stimule la traduction et l'initiation (Thore et al. 2003; Kahvejian et al. 2001). De plus les ARN possédant des queues poly(A) plus courtes quittent les polysomes et se retrouvent ségrégués dans les P-Bodies (Decker and Parker 2012). Il est généralement accepté que la queue poly(A) doit contenir une taille de 250 à 300 adénosines afin de permettre la traduction (Chen et al. 2002; Weill et al. 2012; Thore et al. 2003; Bianchin et al. 2005; Clark et al. 2004). Il n'a pas encore été établi un lien strict entre la déadénylation de la queue poly(A) et le contrôle de la traduction, mais plusieurs indices suggèrent une implication du complexe CCR4-NOT dans la traduction sans intervention de sa fonction de déadénylation.

C.7. Rôle dans la répression de la traduction

Il est maintenant fermement établi que le recrutement du complexe CCR4-NOT sur l'ARN messager provoque la déadénylation et la répression. Des tests de « tethering » ont mis en lumière une activité inattendue du complexe, l'impliquant dans la répression de la traduction. En effet dans un test où le complexe CCR4-NOT est recruté sur un ARN rapporteur sans queue poly(A), on observe un effet répressif. Ce phénomène a été décrit la première fois chez des oocytes de *Xenopus laevis* où la déadénylase CAF1 couplée artificiellement à un rapporteur d'ARN messager réprime la traduction (Temme et al. 2004). Cet événement est indépendant de l'activité de déadénylation, puisque l'inactivation de CAF1 n'empêche pas la répression de la traduction (Bianchin et al. 2005; COOK et al. 2012; Bartlam and Yamamoto 2010; Denis et al. 2003; Dupressoir et al. 2001; Viswanathan et al. 2004). Ces observations ont été complétées par la suite par d'autres études observant un effet similaire sur des micro ARN (Albert et al. 2002; Braun et al. 2011; Mersman et al. 2009; Chekulaeva et al. 2011; Mulder et al. 2007).

Bien que les observations convergent vers l'implication du complexe dans la répression de la traduction, le mode d'action reste à élucider. Une hypothèse est qu'un domaine de la protéine NOT1, domaine appelé MIF4G (middle domain of eukaryotic initiator factor 4G) et situé au centre de la protéine, pourrait jouer un rôle dans cette fonction de répression de la traduction. Ce

domaine de 30kDa est retrouvé communément dans d'autres protéines impliquées dans la régulation de la traduction tels que eIF4G, CPP80 ou même PAIP1 (Panassenko et al. 2012; Lei et al. 2011; Marcotrigiano et al. 2001; Mazza et al. 2001). Dans le cas d'eIF4G, le domaine MIF4G est impliqué dans l'interaction avec l'hélicase DEAD d'eIF4A. Dans ce contexte, notre laboratoire a mis en évidence l'association du domaine MIF4G de CNOT1 avec DDX6, une hélicase qui est vraisemblablement impliquée dans l'activation des facteurs de décoiffage (Panassenko et al. 2009; Mathys et al. 2014). De plus nous avons montré que l'activité ATPasique de DDX6 était nécessaire à l'inhibition traductionnelle (silencing) par les micros ARN, et que par ailleurs cette activité était stimulée par CNOT1 (Denis et al. 2003; Mathys et al. 2014; Dupressoir et al. 2001; Temme et al. 2004; Viswanathan et al. 2004; Albert et al. 2002; Mersman et al. 2009; Mulder et al. 2009; Collart 2003).

Une autre étude a montré que le recrutement du complexe CCR4-NOT par la protéine GW182 induisait une dissociation de PABP de la queue poly(A) des ARNm en absence de déadénylation, ce qui provoquait leur circularisation et leur traduction (Badarinarayana et al. 2000; Zekri et al. 2013). Dans chaque cas le recrutement du complexe CCR4-NOT induit un remodelage des complexes protéines – ARN messenger qui de ce fait, facilite la répression de leur traduction.

C.8. Rôle dans la fidélité de la traduction

Dans le cadre d'expérience de fractionnement des polysomes, il a été montré que les sous-unités du complexe CCR4-NOT pouvaient se retrouver dans des fractions de polysomes engagés dans la traduction. De plus une délétion des sous-unités NOT2, NOT4 et NOT5 induit d'importants changements sur le profil de fractionnement des polysomes, conduisant à de l'agrégation protéique (Lemaire et al. 2000; Panassenko et al. 2012; Deluen et al. 2002; Sanders et al. 2002). La contribution du complexe CCR4-NOT à la fonction ribosomale, principalement par son association à la chaperonne ribosomale NAC (Nascent-polypeptide-associated peptide), a également été démontrée (Benson et al. 1998; Panassenko and Collart 2011; Mulder et al. 2007; Hiraishi et al. 2009). L'activité ligase E3 de NOT4 a été décrite comme pouvant renforcer l'interaction de la chaperonne NAC avec le ribosome et le protéosome (Collart and Panassenko 2012; Panassenko et al. 2009; Miller et al. 2012). La proximité du complexe CCR4-NOT avec les ribosomes semble nécessaire au bon contrôle des événements de traduction. En effet, dans le cas où un ARN messenger ne possède pas de codon de terminaison pouvant causer une traduction délétère de la queue poly(A) qui conduit à l'arrêt de la traduction, il a été montré que NOT4 était nécessaire pour

effectuer la dégradation du peptide nouvellement synthétisé chez la levure (Neely et al. 2010; Dimitrova et al. 2009; Venturini et al. 2012).

Des expériences de co-purification ont mis en évidence que le complexe CCR4-NOT pouvait se trouver associé avec la protéine SMG7, un facteur pouvant être recruté par le complexe de surveillance composé des protéines UPF1, UPF2 et UPF3 sur les ARN anormaux. Ce complexe s'associe autour d'un codon de terminaison prématuré (PTC) et active la voie de dégradation de type NMD (Non sense Mediated Decay). SMG7 en s'associant à la sous-unité CAF1 et, de par sa fonction de déadénylation, augmente la dégradation de l'ARN contenant un codon de terminaison prématuré (Morel et al. 2003; Loh et al. 2013; Prévôt et al. 2001; Garapaty et al. 2008).

C.9. CCR4-NOT et réponse immunitaire innée

C.9.1. Implication du complexe CCR4-NOT dans la régulation du signal en réponse aux interférons

Comme indiqué plus haut, des études récentes ont mis en lumière l'implication du complexe CCR4-NOT dans la régulation négative du signal des interférons de type I et II, en influant sur la fonction de STAT1, un facteur clé de la réponse immunitaire innée (Bardwell et al. 2006; Chapat et al. 2013). Les interférons (IFN) identifiés pour leurs activités antivirales, exercent aussi des activités immunomodulatrices anti-prolifératives et apoptotiques. Les IFN sont utilisés dans le traitement de certains cancers, de l'hépatite C et de la sclérose en plaque. Ils sont composés de trois types : I (IFN alpha, IFN bêta, IFN oméga, IFN kappa), II (IFN gamma) et III (IFN lambda). Les IFN se fixent sur des récepteurs spécifiques et activent deux voies de signalisation dépendantes des protéines JAK/STAT, aboutissant à l'homo ou l'hétérodimérisation des STAT, qui déclenche leur translocation vers le noyau et provoque l'induction de gènes cibles. Les produits de ces gènes sont les médiateurs des effets biologiques des IFN. STAT1 est activée en réponse à une stimulation du IFN-gamma par la voie de transduction du signal JAK-STAT, au cours de laquelle les protéines kinases JAK et STAT sont activées par phosphorylation sur les résidus tyrosines par les kinases de la famille JAK. Cette phosphorylation entraîne leur dimérisation et permet leur entrée dans le noyau où elles jouent un rôle de facteur de transcription. Il a été montré que la protéine CAF1 pouvait s'associer avec une forme dormante de STAT1 dans le cytoplasme de cellules latentes. Une stimulation de ces cellules par IFN gamma induit une dissociation de STAT1 et de CAF1, suggérant que cette dernière joue un rôle sur le contrôle du trafic de STAT1. De plus CAF1

joue un rôle de par son activité déadénylase car elle participe à l'accélération de dégradation des ARNm codant pour STAT1, permettant de moduler la réponse immunitaire. Le rôle de CAF1 dans la signalisation des interférons a été confirmé dans des cellules n'exprimant pas CAF1, qui exhibent un meilleur niveau de résistance à l'infection virale (Chapat et al. 2013).

C.9.2. Le complexe CCR4-NOT affecte le complexe majeur d'histocompatibilité

D'autre part, un vaste criblage d'ARN interférents (ARNi) a mis en évidence que les sous-unités du complexe CCR4-NOT affectaient le complexe majeur d'histocompatibilité (CMH), un système de reconnaissance du soi présent chez la plupart des vertébrés (Dorf 1981). Les molécules du CMH sont à la surface des cellules et assurent la présentation des antigènes aux lymphocytes T afin de les activer. Cette étude a montré que la répression de CNOT1 augmentait l'expression des gènes CMHC-II. Ces gènes sont eux-mêmes contrôlés par le signal des interférons. Ces résultats suggèrent que le complexe joue un rôle dans la présentation de l'antigène. De plus une autre étude a identifié la protéine CNOT4 comme interagissant et régulant la cascade JAK/STAT chez la drosophile et l'homme (Grönholm et al. 2012).

C.10. Mécanismes de sélection de l'ARN messenger

La vitesse de dégradation d'un ARN messenger est principalement déterminée par les éléments de séquence capable de recruter des protéines responsables de la régulation de sa dégradation. Certaines de ces protéines ont été trouvées comme pouvant interagir avec les sous-unités du complexe CCR4-NOT. Par exemple, son recrutement sur l'ARN messenger est conduit par l'action concertée de la protéine se liant à la queue poly(A) (PAPB) et des protéines TOB1&TOB2, en interaction avec la sous-unité CAF1 (Ezzeddine et al. 2007; Okochi et al. 2005). Par ailleurs d'autres facteurs peuvent s'associer directement de manière spécifique aux protéines pouvant se lier à l'ARN. C'est le cas de la protéine Pumilio fem3 qui s'associe au complexe CCR4-NOT et induit la déadénylation et la répression de la traduction (Goldstrohm et al. 2007; 2006). D'autre part, la protéine tristétraproline (TTP) qui s'associe aux facteurs AREs (AU-rich éléments), peut interagir directement avec la sous-unité CNOT1 du complexe CCR4-NOT pour entraîner une augmentation rapide de la dégradation (Fabian et al. 2013; Sandler et al. 2011). De la même manière le complexe est recruté sur des ARNm spécifiques par l'action de facteurs que l'on retrouve aussi dans la répression des micros ARNs (Huntzinger and Izaurralde 2011). En effet les éléments CDE (constitutive decay element) adoptent une structure en tige/boucle et constituent un motif très

conservé responsable de la dégradation des ARN. Ainsi il a été démontré qu'un de ces CDE, stimulant la dégradation à travers son association avec la protéine Roquin, pouvait recruter le complexe CCR4-NOT (Leppek et al. 2013).

La localisation des ARN au sein de la cellule influe grandement sur leur propension à être dégradés plus ou moins rapidement. Les protéines mRNP qui s'associent avec les facteurs responsables de la dégradation et du décoiffage peuvent être séquestrées au sein des particules P-bodies (PB) (Eulalio et al. 2007; Parker and Sheth 2007). Dans ce contexte, il est intéressant de mentionner que les sous-unités CCR4 et CAF1 peuvent elles aussi y être retrouvées et que l'activité de déadénylation est indispensable à la formation de ces particules (Zheng et al. 2008). La séquestration des ARN messagers dans les PB, ainsi que des facteurs de dégradation associés constitue un moyen de les extraire de façon temporaire de la fraction disponible pour la traduction, ou de les conserver pour une utilisation ultérieure.

Le complexe CCR4-NOT est organisé autour d'une protéine d'échafaudage de 240 kDa. Cette protéine NOT1 représente une surface étendue sur laquelle viennent s'ancrer les autres composants du complexe. Une étude de microscopie électronique a permis de révéler l'architecture globale du complexe (Nasertorabi et al. 2011). Les auteurs de cette étude ont placé les différentes sous unités du complexe dans l'enveloppe en fonction de leur taille respective mais sans validation expérimentale. Le complexe arbore une architecture en L. Notre laboratoire a contribué depuis 2012 à enrichir les connaissances sur l'architecture de ce complexe. Nous avons pu notamment apporter des détails sur la structure et la fonction des différents modules qui le composent (Module de Déadénylation, Module NOT, Module impliqué dans la répression de la traduction et Module d'ubiquitination). J'ai choisi dans cette partie de mon manuscrit de détailler les différentes découvertes des structures des composants du complexe CCR4-NOT et de les exposer dans l'ordre chronologique de leurs découvertes.

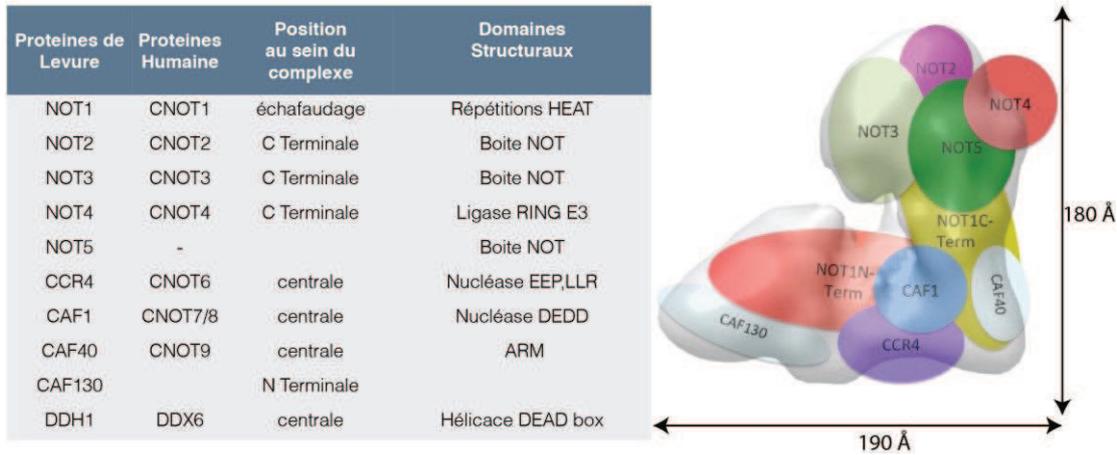


Figure 2 : Composants du complexe CCR4-NOT chez l'humain et la levure. Dans le tableau de gauche sont détaillées les différentes protéines qui composent le complexe CCR4-NOT, leurs positions respectives sur la protéine d'échafaudage au sein du complexe, ainsi que les domaines structuraux qui les caractérisent. Le schéma de droite donne une représentation de l'organisation du complexe.

D. Architecture et analyse structurale du complexe CCR4-NOT et des sous unités qui le composent

D.1. Les déadénylases du complexe CCR4-NOT

Le complexe CCR4-NOT comporte deux enzymes de déadénylation qui appartiennent à deux types distincts. La protéine CAF1 appartient au type de nucléase DEDD, ainsi nommé à cause des acides aspartiques et glutamines conservés au sein du site actif. Par ailleurs la protéine CCR4 appartient à la classe de l'exonucléase-endonucléase-phosphatase de type EEP. Le complexe de levure comporte ces deux enzymes qui sont toutes deux impliquées dans la dégradation des ARN messagers, bien que seule la nucléase CCR4 soit indispensable à cette fonction. Le complexe humain présente une composition similaire mais possède la particularité d'avoir deux orthologues de chaque enzyme de déadénylation. Dans le cas de CCR4 on retrouve les orthologues humains CNOT6 et CNOT6L, qui sont impliqués dans des fonctions distinctes. D'autre part, CNOT7 et CNOT8 sont les orthologues humains de CAF1. Ils partagent une très forte homologie de séquence et jouent un rôle dans la prolifération cellulaire. Ces orthologues sont retrouvés de manière exclusive au sein du complexe, suggérant qu'ils partagent le même site de fixation.

D.1.1. CAF1 : une déadénylase de type DEDD

La première structure cristallographique de CAF1 a été déterminée chez la levure *Saccharomyces cerevisiae* en 2003 (Thore et al. 2003). Deux études similaires chez *Schizosaccharomyces Pombe* en 2007, puis en 2008 (Jonstrup et al. 2007; Andersen et al. 2009) ont permis de compléter les informations structurales disponibles sur cette déadénylase. CAF1 adopte une structure en forme de croissant composée de 13 hélices α et de 6 feuillets β (Figure 3). La comparaison de la structure de CAF1 avec les structures disponibles dans les bases de données a permis de la classer comme faisant parti de la famille des nucléases de type DEDD qui adopte les repliements classiques de type DnaQ. La famille DEDD a pour caractéristique le fait d'arborer un site actif conservé qui comporte les acides aminés Asp-Glu-Asp-Asp. Bien que l'homologie de séquence soit très faible, CAF1 partage des caractéristiques structurales très proches de la sous-unité ϵ - de l'ADN polymérase III et du domaine exonucléasique de l'ADN polymérase I. La détermination de la structure de CAF1 chez *S. Pombe* a apporté des détails qui contribuent à expliquer la sélectivité de ce type de déadénylase. Dans cette structure, deux ions divalents ont été localisés dans le site actif (appelés sites A et B), et des expériences de biochimie ont montré qu'ils étaient essentiels à son activité. Par ailleurs cette étude a montré que le site A était favorablement occupé par du zinc (Zn^{2+}) et que le site B quant à lui favorisait l'occupation du manganèse (Mn^{2+}). La composition de ces ions dans le site actif a une influence sur les cinétiques de déadénylation. Par exemple en présence d'ions Mg^{2+} , Mn^{2+} et Zn^{2+} , la déadénylation est très lente et non spécifique. En absence de Zn^{2+} , CAF1 exhibe une activité très rapide et spécifique de substrat arborant une queue poly(A). Ce résultat conduit à l'hypothèse que les niveaux ioniques cellulaires contribuent à réguler le taux de dégradation des ARNm. Chez *S. cerevisiae*, le motif DEDD de CAF1 est substitué par un motif SEDQ. La perte de deux acides aminés catalytiques conservés explique que chez la levure CAF1 ait perdu sa fonction catalytique et que son rôle soit essentiellement structural. En effet, CAF1 contribue au recrutement de CCR4 sur le complexe CCR4-NOT. Il a été montré par des essais de dégradation *in vitro* que CAF1 était capable de dégrader des substrats ARN, poly(U), poly(C) et poly(A), avec une légère préférence pour ces derniers. Cette faible spécificité peut être expliquée par le fait que les interactions entre ces substrats ARN et l'enzyme sont principalement des interactions de type Van der Waals. La structure de CNOT7, l'orthologue humain de CAF1, a été résolue en présence de la protéine TOB, un facteur antiprolifératif (Horiuchi et al. 2009). La structure révèle un cœur catalytique caractéristique de la famille des ARNase D arborant le motif DEDD, comme dans le cas des protéines de levure. La Figure 3 montre une représentation du complexe ainsi qu'une vue détaillée du site actif.

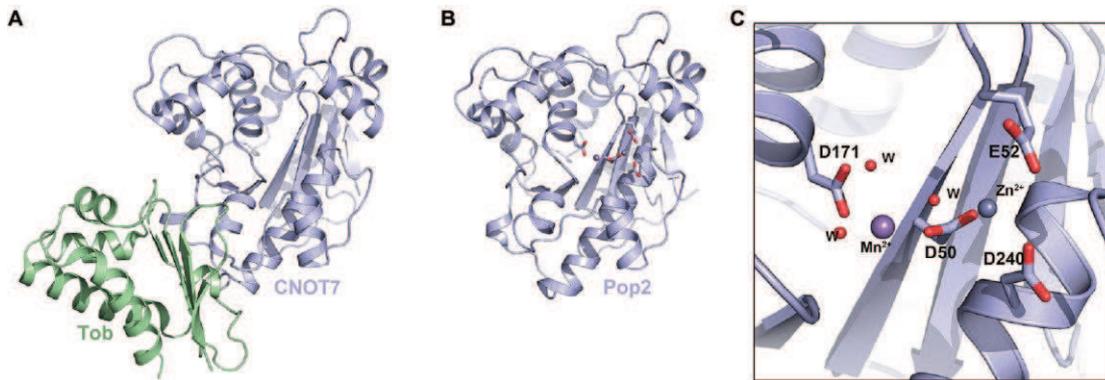


Figure 3 : Les déadénylases DEDD

(A) Structure cristalline du Complexe humain CNOT7-Tob.

(B) Structure cristalline de Caf1 de *S.Pombe*.

(C) Zoom du site actif montrant le motif DEDD en représentation bâton avec les ions Mn²⁺ et Zn²⁺ représentés par des sphères.

D.1.2. CCR4 : une déadénylase de type EEP

CCR4 a été identifiée par plusieurs études comme principale responsable de l'activité de déadénylation chez la levure. La structure du domaine nucléase de la protéine humaine CNOT6, orthologue de CCR4 a été déterminée en 2010 (Wang et al. 2010). Le repliement de CNOT6 est caractérisé par un sandwich α - β dans lequel deux feuillets β sont pris en sandwich entre deux couches d'hélices α . Le site actif se situe au sommet de la molécule où se trouvent deux ions magnésium (Mg²⁺). Le Glutamate 240 est responsable de la coordination d'un des ions. La mutation de cet acide aminé ou l'absence de magnésium, abolit l'activité, suggérant que l'action de CNOT6 est dépendante du magnésium. L'autre ion magnésium est coordonné par l'action combinée des acides aminés Asp410, Asn412 et His529. Des essais de déadénylation ont montré que CNOT6 possédait une spécificité stricte pour les substrats poly(A). Deux structures de CNOT6 en complexe soit avec un ADN poly(A) ou avec une molécule de AMP (adenosine 5' monophosphate) permettent d'expliquer la spécificité de cette déadénylase. Dans le site actif de la structure avec l'ADN poly(A) on peut observer trois nucléotides. L'interaction est induite par des liaisons hydrogènes de l'ion magnésium et de l'ASN412 sur le phosphate de la deuxième base. L'adénine de la base A2 se cale entre la Phe484 et la Pro365. Les nucléotides A1 et A3 ne forment pas d'interaction particulière avec les résidus du site actif. Une seconde structure avec une molécule d'AMP confirme l'importance de ces résidus dans la reconnaissance du substrat. Une mutation des acides aminés Asn412 et Phe484 abolit l'activité. Les autres bases, de part leur spécificité chimique, ne pourraient pas être accommodées à l'intérieur du site actif.

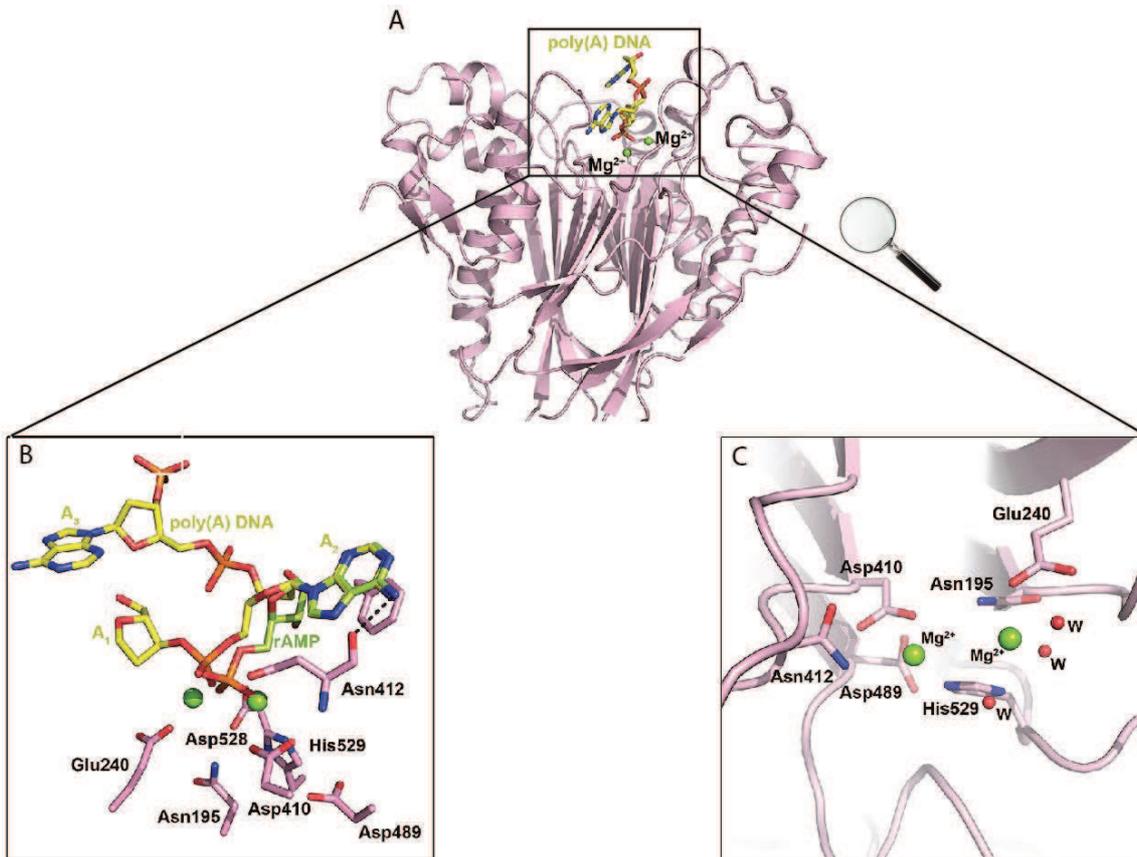


Figure 4 : Les déadénylases EEP

(A) Structure cristalline du domaine nucléase humain de CCR4 (CNOT6L) en complexe avec un poly(A) ADN. Le ligand est représenté en bâton jaune et les ions Mg^{2+} sont représentés en sphères.

(B) Substrat dans le site actif montrant le poly(A) ADN en bâtons jaunes et le rAMP en bâton vert pâle. Les résidus de CNOT6L sont représentés en bâton rose et les ions Mg^{2+} en sphères vertes.

(C) Vue détaillée du site actif de CNOT6L montrant les résidus en bâton, les ions Mg^{2+} en sphères et les molécules d'eau en sphères rouges.

D.2. Le module de déadénylation du complexe CCR4-NOT

L'étude structurale que j'ai menée a permis d'élucider l'architecture du sous-complexe qui comprend les protéines CAF1/CCR4 et NOT1 de *S. cerevisiae* (Basquin et al. 2012). Ce complexe ternaire est situé à la fin de la première moitié de la protéine d'échafaudage NOT1.

D.2.1. Le domaine NOT1 (754-1000) adopte un repliement MIF4G

Le domaine NOT1 (754-1000) contient dix hélices antiparallèles. Cet arrangement est caractéristique de ce qui peut être observé pour les répétitions HEAT. Ces motifs sont formés par deux hélices antiparallèles (nommées hélice A et B) qui s'empilent l'une sur l'autre (Andrade et al.

2001). Une recherche dans les bases de données structurales avec le logiciel DALI (Holm and Rosenström 2010) a permis d'identifier un domaine MIF4G (middle domain of eukaryotic initiator factor 4G) composé de répétitions HEAT. Les cinq répétitions HEAT de NOT1 (754-1000) peuvent être superposées avec les domaines MIF4G de eIF4G (Schütz et al. 2008) et UPF2 (Kadlec et al. 2004) avec un écart quadratique moyen inférieur à 2.8 Angströms, sur soixante dix pour cent des atomes de carbone. De plus, NOT1 (754-1000) possède deux segments étendus aux extrémités N et C-terminale qui font partie intégrante du domaine MIF4G et s'associent aux répétitions HEAT par des interactions hydrophobes. Ce domaine est similaire dans les structures déterminées en l'absence ou en présence de CAF1 (écart quadratique moyen inférieur à 0,5 Å, sur l'ensemble des atomes de carbone). L'analyse structurale du domaine NOT1 (754-1000) met en évidence une surface concave composée d'une large région comprenant des résidus conservés sur les hélices B des répétitions HEAT 1-5. Une plus petite région de résidus conservés est aussi présente sur l'autre surface entre les boucles connectant les répétitions HEAT 3-5 et HEAT 4-5 (Figure 5).

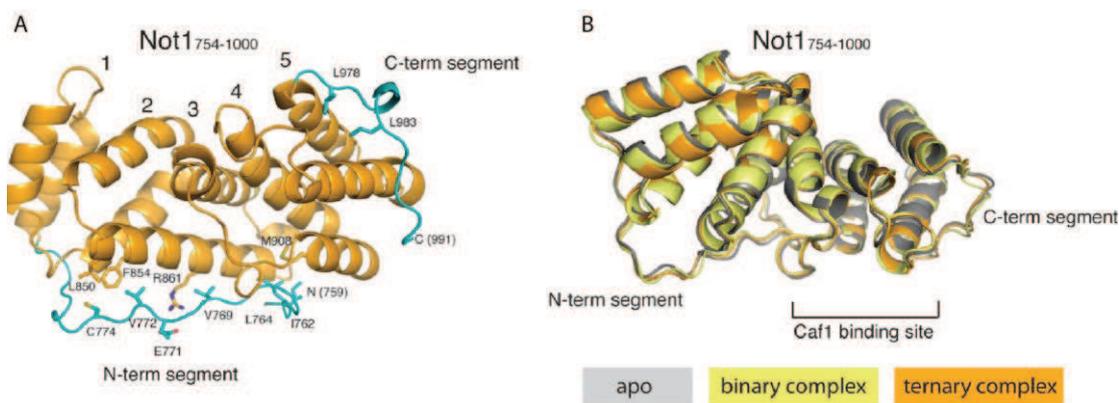


Figure 5 : Le domaine MIF4G de NOT1

Sur le panneau de gauche sont représentées les interactions des segments N et C- terminales de NOT1 (754-1000) en cyan sur les hélices des répétitions HEAT. Les résidus conservés impliqués dans les interactions polaires et hydrophobes sont indiqués.

Sur le panneau de droite sont représentés les différentes structures du domaine MIF4G de NOT1. En gris la structure isolée, en jaune en complexe avec CAF1 et en orange en complexe avec CAF1 et CCR4 (Basquin et al. 2012).

D.2.2. L'interaction du domaine NOT1 (754-1000) avec CAF1 implique des résidus conservés

CAF1 est reconnue par une petite surface composée par des résidus conservés situés entre les boucles connectant les répétitions HEAT 3-5 et HEAT 4-5. Comme indiqué plus haut, CAF1 adopte un repliement contenant un feuillet α - β entouré d'hélices alpha, à l'exception d'une petite surface qui forme la cavité du site actif. Le site actif de CAF1 de *S. cerevisiae* possède les résidus

Ser188-Glu190-Asp310-Gln394 (SEDDQ) qui sont sub-optimaux comparé à la séquence canonique (DEDD) chez *S. Pombe* ou chez les orthologues humain. Dans le complexe CCR4-NOT, CAF1 interagit dans une région diamétralement opposée au site actif. L'interaction entre CAF1 et NOT1 (754-1000) est conduite par des résidus conservés, en particulier, par Met290, Met296 et Trp333 chez CAF1 et Pro897, Phr938, Pro943 et Trp944 chez NOT1 (Figure 6). Au total, cette interface comprend seulement 8,4% de la surface de solvant accessible (230 Å² pour CAF1 et 420 Å² pour NOT1), laissant une large surface de résidus conservés sur la surface concave opposée libre et exposée au solvant.

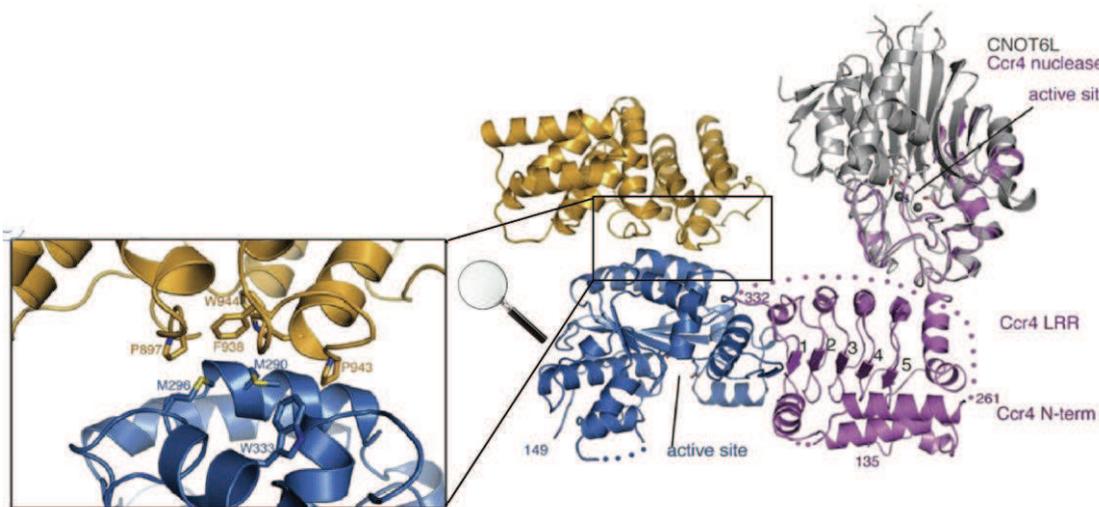


Figure 6 : Interface NOT1-CAF1

Zoom montrant les résidus en bâtons à l'interface NOT1-CAF1

D.2.3. L'interaction de CAF1 avec le domaine LRR de CCR4 forme un cœur hydrophobe

Le site actif de CAF1 reste accessible dans le complexe car la fixation de CCR4 est conduite par le domaine LRR sur une surface adjacente. L'interaction de CAF1 avec le domaine LRR de CCR4 forme un cœur hydrophobe continu. Le domaine LRR de CCR4 contient cinq répétitions riches en leucine (Leucine Rich Repeat) qui forment une structure en croissant. La surface convexe extérieure est formée par des hélices, alors que la surface concave est formée par des brins β parallèles. La surface concave est souvent liée à des interactions intermoléculaires (Bella et al. 2008). Chez CCR4 de levure, la surface concave du domaine LRR interagit intramoléculairement avec les hélices de la région N-terminale. La région de 70 acides aminés reliant la partie N-terminale et le LRR est désordonnée. La partie ordonnée de la région N-terminale recouvre la surface hydrophobe du feuillet β du domaine LRR (entouré par les résidus Tyr363, Val384, Tyr407, Tyr409, Phe411 et Phe430). L'analyse structurale suggère que la région N terminale et le domaine LRR de CCR4 forme une unité structurale unique. Chez les vertébrés, CCR4 ne possède pas la

région N-terminale et la surface concave du domaine LRR est polaire. Le domaine LRR de CCR4 inclut deux hélices qui recouvrent les résidus leucines de la cinquième répétition et conduit directement au domaine nucléase (Figure 7). Le cœur hydrophobe de la première répétition du domaine LRR est recouvert par CAF1. CAF1 interagit avec CCR4 via une surface formée par une longue boucle (résidus 193-206 - Figure 7) et une hélice (résidus 207-219). Cette région de CAF1 subit un changement de conformation, comparé à la structure de CAF1 isolée. Les résidus de CAF1 197-199 forme un brin β qui étend le feuillet β de CCR4. Les résidus hydrophobes de la première répétition du domaine LRR contactent des chaînes latérales apolaires de CAF1 (Ala215 et Phe219). Les résidus responsables de l'interaction CAF1 - CCR4 sont conservés des champignons aux animaux, et la surface d'interaction a été confirmée par mutagenèse (Clark et al. 2004; Ohn et al. 2007).

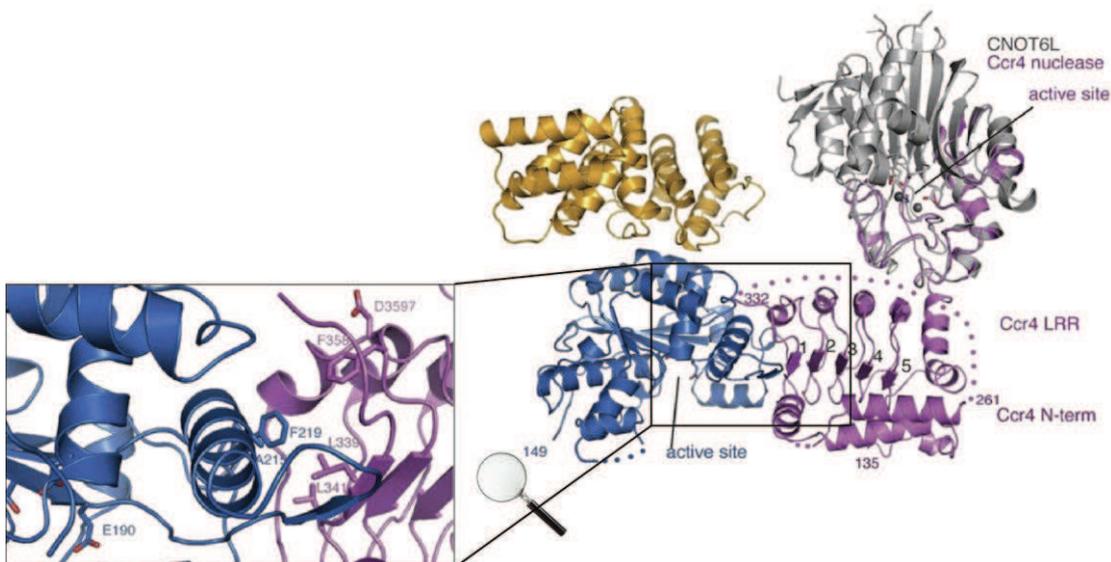


Figure 7 : Interface CCR4-CAF1

Zoom montrant les résidus en bâtons à l'interface CCR4-CAF1.

D.2.4. Le domaine nucléase de CCR4 est adjacent au domaine LRR

Le domaine nucléase de CCR4 est adjacent au domaine LRR. Une étude structurale précédente a montré que le domaine catalytique humain CNOT6, orthologue de CCR4, peut être classé parmi les membres de la famille de l'endonucléase-exonucléase-phosphatase (EEP) (Wang et al. 2010). Le repliement de CNOT6 est caractérisé par un sandwich α - β constitué de deux feuillets β pris en sandwich entre deux couches d'hélices, formant un site actif proéminent. Dans deux molécules de l'unité asymétrique des cristaux du complexe NOT1 (754-1000), CAF1 et CCR4, les hélices de la moitié du domaine LRR interagissent avec le domaine nucléase, quand l'autre

moitié est désordonnée. La partie ordonnée du domaine nucléase correspondant à l'orthologue humain de CNOT6 adopte la même conformation et peut être **superposée** avec un écart quadratique moyen inférieur à 1,1 Å, sur l'ensemble des atomes de carbone de la chaîne principale. L'interaction entre les domaines nucléase et LRR recouvre une surface d'environ 400 Å². L'interface présente des contacts hydrophobes (entre Phe403 de la surface du feuillet β du domaine LRR et la Pro773 du domaine nucléase) et des contacts polaires (entre Asp460 de l'hélice C terminale du domaine LRR et la Ser528 de la nucléase). Bien que les interactions ioniques intramoléculaires de CCR4 soient de nature faibles et stabilisées par des contacts au niveau du réseau cristallin, elles sont conservées et forment une région continue de résidus conservés exposés au solvant (Tyr457 et Asn461 du domaine LRR et Trp529, Asp534 et Tyr535 du domaine nucléase). De plus ces interactions ont été précédemment suggérées par un criblage d'interaction double hybride chez la levure (Clark et al. 2004). L'arrangement relatif des deux domaines en forme de L est semblable à la région en L décrite par la structure du complexe CCR4 NOT chez *S.cerevisiae* déterminée par microscopie électronique en coloration négative (Nasertorabi et al. 2011).

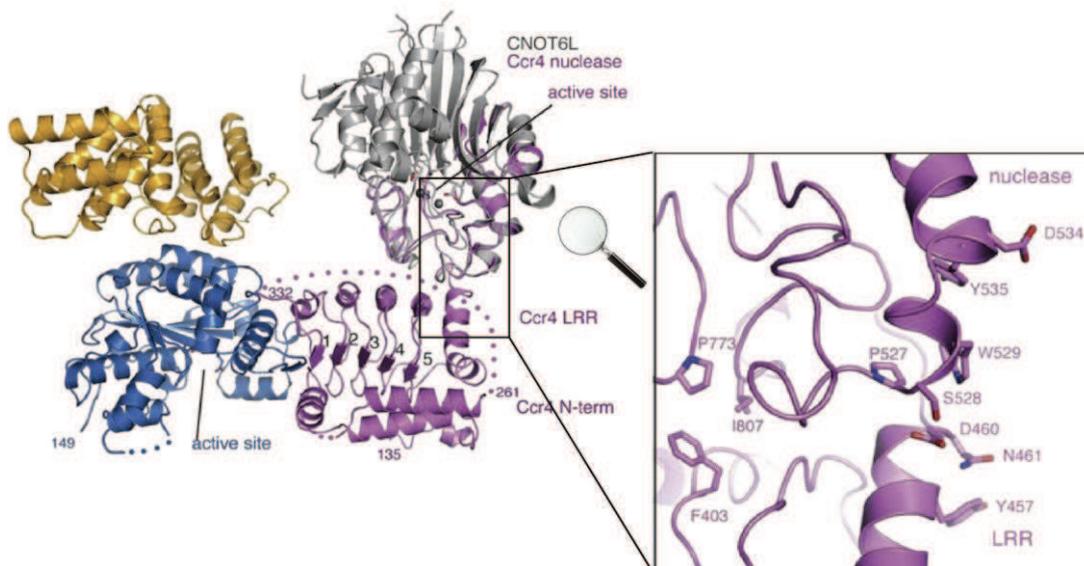


Figure 8 : Interface LRR-Nucléase de CCR4

Zoom montrant les résidus en bâtons à l'interface LRR-Nucléase de CCR4.

Les résidus hydrophobes à l'interface sont indiqués ainsi que les résidus conservés sur la surface.

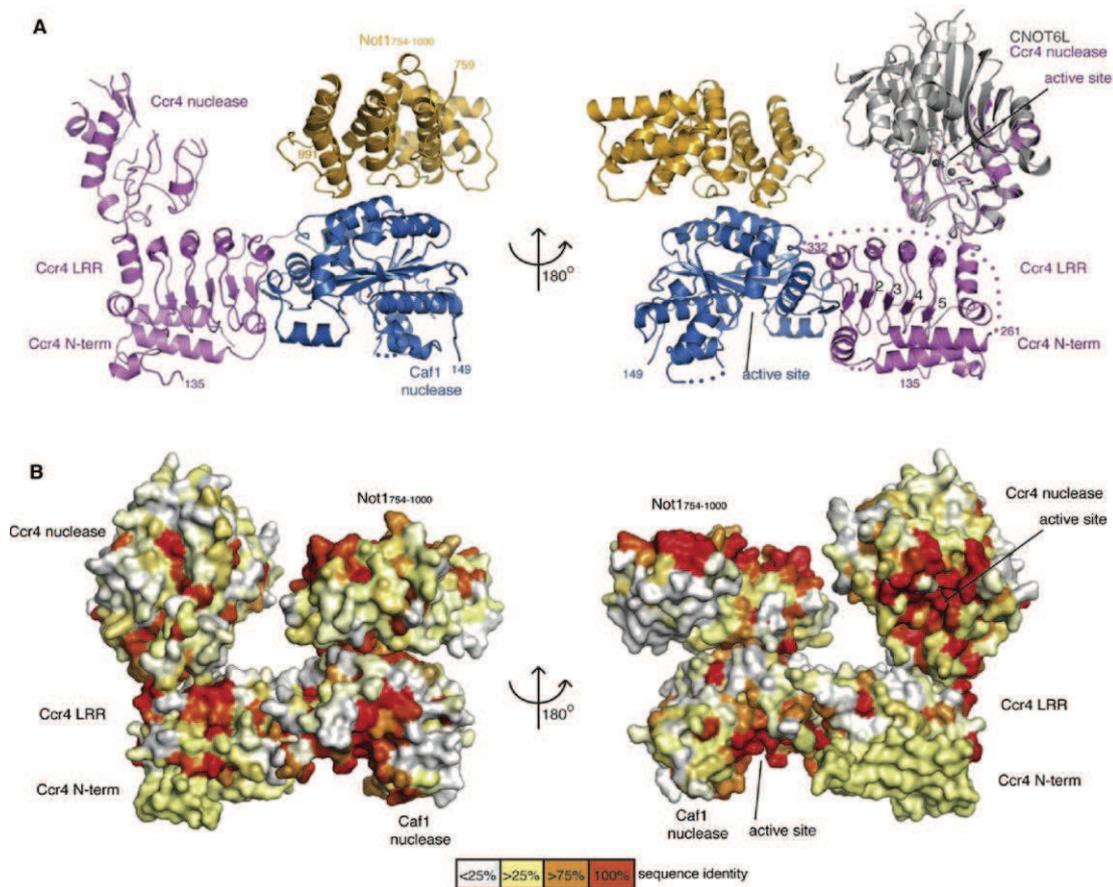


Figure 9 : Structure de complexe NOT1 (754-1000) CAF1-CCR4.

Panneau A : Le complexe est représenté dans deux orientations tournées de 180° autour de l'axe vertical. NOT1 est en orange, CAF1 en bleu et CCR4 en rose. Sur la droite le complexe est représenté superposé à l'orthologue humain CNOT6L, en gris.

Panneau B : Représentation de la surface du complexe dans une orientation identique au panneau A. La surface est colorée en fonction du niveau de conservation des résidus.

D.2.5. Mutations entraînant la déstabilisation de NOT1-CAF1 et CAF1-CCR4

Des mutations basées sur l'analyse structurale ont été introduites pour désorganiser les interactions des protéines NOT1-CAF1 et CAF1-CCR4. Afin de déstabiliser l'interface NOT1 (754-1000) et CAF1 nous avons conçu un double mutant de CAF1 (M290K, M296K) et un double mutant de NOT1 (F938E, P943Y). Des expériences d'immunoprécipitation utilisant des protéines étiquetées ont montré que le mutant de CAF1 perdait l'interaction avec NOT1 tout en préservant celle avec CCR4. De même le mutant de NOT1 affecte l'interaction avec CAF1.

Afin de déstabiliser l'interface entre CAF1 et CCR4, nous avons conçu un double mutant de CAF1 (A215E et F219E) et un double mutant de CCR4 (L339E et L341E). Le mutant de CAF1 abolit l'interaction avec CCR4 dans des expériences d'immuno-précipitation et réduit légèrement l'interaction avec NOT1. Inversement le mutant de CCR4 abolit l'interaction avec CAF1 mais n'impacte pas l'interaction avec NOT1. Il est important de souligner que dans tous les cas les

protéines mutantes sont produites à des niveaux normaux (Basquin et al. 2012).

Bien que les déadénylases soient actives *in vitro*, les résultats des expériences *in vivo* confirment que l'incorporation de CCR4 et CAF1 au complexe CCR4-NOT est cruciale pour l'activité de déadénylation. En effet lorsque l'on perturbe l'interaction entre CAF1 et NOT1, les levures exhibent un phénotype de croissance perturbé à haute température. Par ailleurs, on observe l'accumulation d'intermédiaires de dégradation des ARNm. Ces observations ont été confortées par une étude où un mutant de la protéine CAF1 chez la drosophile qui empêche l'interaction avec NOT1 avait un large effet sur la dégradation des ARN messagers (Petit et al. 2012).

Il est à noter que dans le même temps, une autre étude a confirmé cette structure chez l'humain avec un complexe binaire des protéines CNOT1 et CNOT7, qui sont les orthologues de NOT1 et CAF1 (Petit et al. 2012).

D.3. Le domaine NOT1 (154-754) est une plateforme étendue de répétitions HEAT

La structure du domaine de NOT1 (154-753) est entièrement composée d'hélices antiparallèles arrangées côte à côte qui forment une molécule allongée de 120 Å. L'analyse de la structure révèle 13 épingles à cheveux hélicoïdales semblable à des répétitions HEAT. Les répétitions HEAT 1-9 présentent un arrangement antiparallèle à l'exception d'une rotation inhabituelle entre les répétitions HEAT 4 et 5. Les répétitions HEAT 10-13 sont aussi organisées parallèlement mais cette région est tournée de 90 degrés par rapport aux répétitions HEAT 1-9. Une recherche de similarité structurale avec le logiciel DALI (Holm and Rosenström 2010) a identifié une similarité entre les répétitions HEAT 10-13 (résidus 571-746) et quatre répétitions HEAT de MIF4G1. Les répétitions HEAT 10-13 représentent la partie la plus conservée du domaine NOT1 (154-753). En effet les résidus situés sur les hélices A des répétitions HEAT 10-11 forment une région conservée au cours de l'évolution sur la surface extérieure de ce domaine. Ces surfaces conservées des domaines NOT1 (154-753) et NOT1 (754-1000) constituent des candidats de choix pour développer des interactions protéines-protéines soit dans le cadre du cœur du complexe CCR4-NOT, soit pour le recrutement d'autres facteurs.

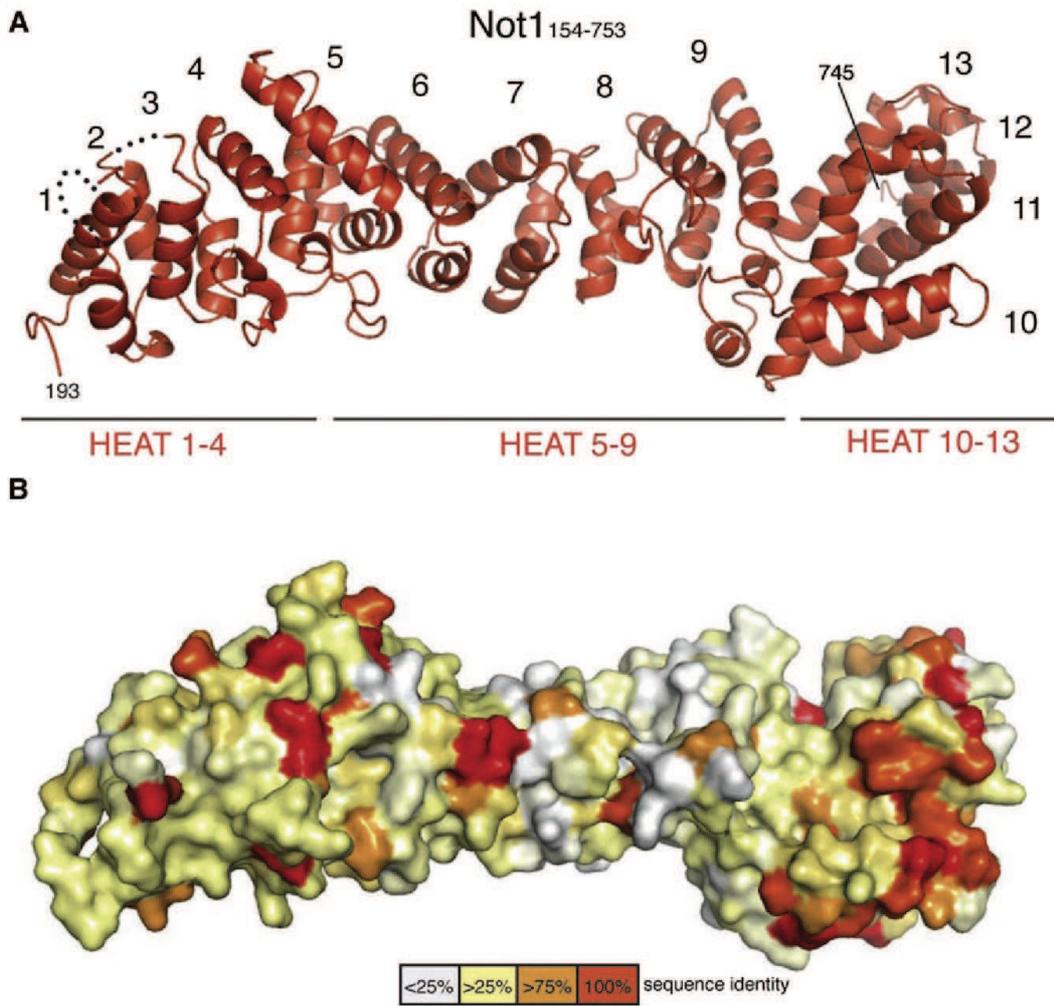


Figure 10 : NOT1 (154-753) est une plateforme de répétitions HEAT.

Panneau A : Structure cristalline de NOT1 (154-753) représenté en ruban et coloré en rouge. Les répétitions HEAT sont numérotées.

Panneau B : Représentation de la surface de NOT1 (154-753) dans une orientation identique au panneau A. La surface est colorée en fonction du niveau de conservation des résidus.

D.4. Le module NOT

Le module NOT est formé par l'association des protéines CNOT1, CNOT2 et CNOT3 chez l'humain et des protéines NOT1, NOT4 et, NOT5 ou NOT3, chez la levure. Dans le complexe de la levure, NOT2 est un orthologue de CNOT2, les protéines NOT5 et NOT3 sont orthologues de CNOT3. Les protéines NOT5 et NOT3 sont par ailleurs paralogues probablement issues d'un événement de duplication de gène. Ces deux protéines partagent la même partie N-terminale, mais se distinguent par leur partie C-terminale (Figure 11). Il a été démontré que NOT5 était indispensable pour la viabilité chez de la levure, tandis que NOT3 présentait un phénotype de croissance très léger (Oberholzer et al. 1998). La fonction du module NOT n'est pas très bien caractérisée, mais il semble qu'il soit impliqué dans la régulation de l'activité du complexe CCR4-NOT et de son

recrutement sur des ARNm spécifiques. Ce module est par ailleurs impliqué dans diverses fonctions, par exemple la transcription.

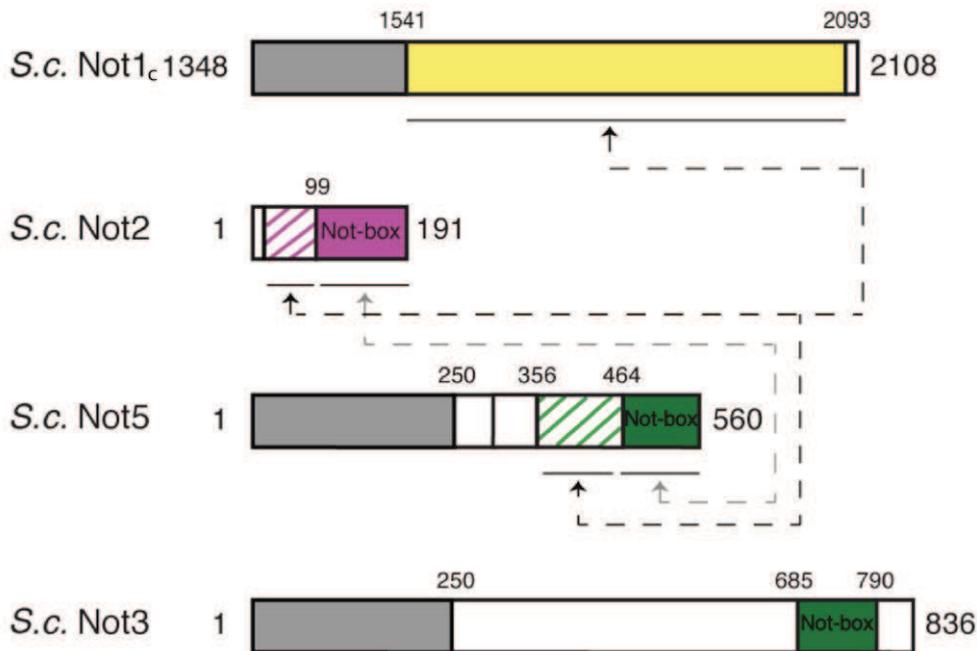


Figure 11 : Organisation et carte d'interaction des protéines NOT de levure.

Les rectangles de couleur montrent les domaines structuraux connus. Les rectangles hachurés représentent des régions de basse complexité impliquées dans des interactions protéine-protéine. Les rectangles gris et blanc représentent des régions pour lesquelles il n'existe pas d'information structurale.

Le complexe NOT1c (1567-2058), NOT2 et NOT5c (partie C terminale de NOT5) est organisé autour de la protéine NOT1. Celle-ci est composée entièrement d'hélices α arrangées antiparallèlement, formant ainsi une molécule allongée de 85 Å. La topologie de CNOT1c est typique de ce qui peut être observé dans les répétitions de types HEAT (décrites plus haut). Un enchaînement de plusieurs répétitions HEAT forme une structure superhéliçoïdale qui présente une surface convexe formée par les hélices de type A et une surface concave formée par les hélices B (Figure 12). NOT1c comporte dix motifs HEAT qui peuvent être divisés en deux groupes. Le premier groupe comprend les répétitions HEAT de 1 à 6 (résidus 1567-1849) et constitue une structure régulière qui se différencie des motifs habituellement retrouvés. Le second groupe formé par les répétitions HEAT de 7 à 10 (résidus 1888-2058) présente aussi une structure régulière à l'exception d'une longue structure beta en épingle à cheveux connectant les répétitions HEAT 7 à 8 et une hélice C-terminale additionnelle (résidus 2059-2079). Par ailleurs, ce second groupe possède quatre des cinq répétitions HEAT qui sont caractéristiques d'un domaine MIF4G. Un segment de 40 résidus connecte les répétitions HEAT 6 à 7 et s'enroule autour afin de former une large région hydrophobe. Les deux groupes de répétitions HEAT se fixent l'un sur l'autre perpendiculairement afin de former une structure en T. Comme dans la structure du domaine N-terminal de NOT1

(décrite plus haut), les domaines N et C-terminaux sont construits suivant des principes extrêmement similaires, bien que l'orientation de chacun de ces domaines diffère dans le détail.

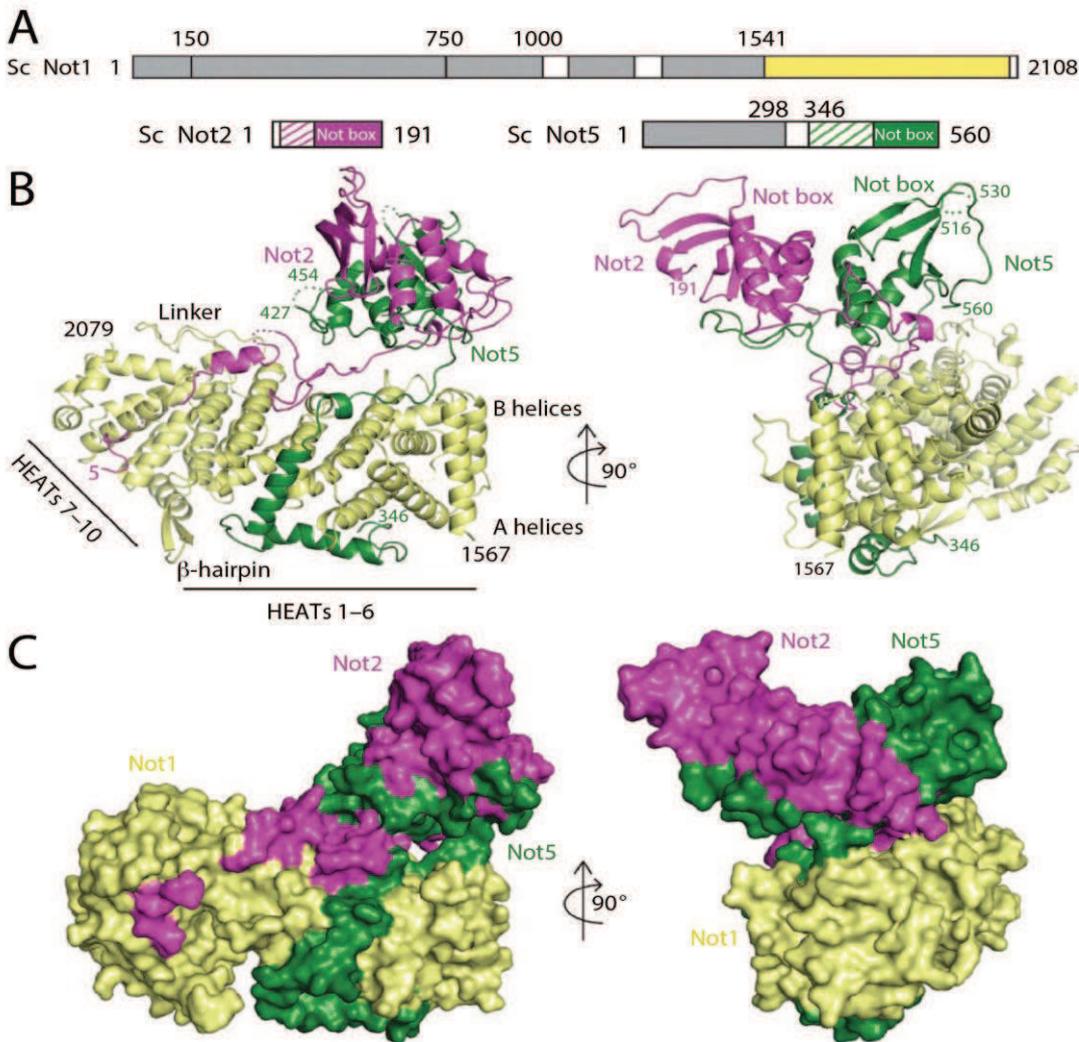


Figure 12 : Structure du complexe de levure NOT1_c-NOT2-NOT5

Panneau A : Représentation schématique de l'organisation des domaines des protéines *S. cerevisiae* NOT1_c, NOT2 et NOT5. Les rectangles de couleur indiquent les domaines globulaires présents dans la structure (Jaune, NOT1 ; Magenta, NOT2 ; vert, NOT5).

Panneau B : Structure du complexe NOT1_c -NOT2-NOT5 représenté en ruban dans deux orientations.

Le domaine de NOT2 se fixant à NOT1_c peut être découpé en trois segments. Le premier segment (NOT2 résidus 5-13) se fixe au domaine MIF4G de NOT1_c, principalement sur ces hélices A des répétitions HEAT 9 et 10. A cet endroit se trouve une poche hydrophobe dans laquelle s'enfonce la Leucine 9, un résidu conservé dont on a montré qu'il était important fonctionnellement *in vivo* (Russell *et al.* 2002). Le second segment (NOT2 résidus 31-64) se fixe à NOT1_c sur une répétition HEAT adjacente, zigzagant entre les hélices B des répétitions HEAT 4-6. Ce segment de NOT2 forme une hélice courte et une épingle à cheveux. L'hélice se fixe à des résidus hydrophobes de la surface conservée de la répétition HEAT 5, centrée sur les résidus Arg1811 et Leu1814. L'épingle à cheveux plonge dans un creux hydrophobe de résidus conservés des répétitions HEAT4-5 (Phe1751 à Ile1812). Le troisième segment (NOT2 résidus 65-75) s'étend au-dessus de l'hélice B vers la répétition HEAT 3. Le domaine de NOT5 (298-560) se fixant à NOT1_c peut être également découpé en trois segments. Le premier segment (NOT5 résidus 346-373) contient une hélice α et se fixe à l'hélice A de NOT1_c au travers d'interactions hydrophobes. Le second segment (NOT5

résidus 374-391) contient également une hélice alpha et se fixe sur le flanc de NOT1c formé par la connexion des répétitions HEAT 3-5 au travers d'interactions hydrophobes. Le troisième segment (NOT5 résidus 392-404) s'étend au-dessus des hélices B de NOT1c entre les répétitions HEAT 1-3 formant des contacts polaires et apolaires. il interagit directement avec le troisième segment de NOT2 par le biais d'un pont salin (entre Asp393 et Arg65).

Panneau C : Représentation de la surface du complexe dans les mêmes orientations que le panneau B.

NOT2 et NOT5(298-560) contiennent des domaines globulaires similaires appelés NOT box qui sont précédés par des extensions N-terminales dont les résidus 5-75 de NOT2 et les résidus 346-404 de NOT5(298-560) sont responsables de l'interaction avec NOT1. Ces régions N-terminales s'étendent sur de plus de 100 Å et couvrent une surface de 3700 Å². Chacune de ces extensions peut être découpée en trois segments interagissant avec NOT1c principalement pour des interactions hydrophobes (Figure 13). L'analyse de la structure suggère que NOT2 et NOT5(298-560) se fixent à NOT1c de façon synergique. Cette hypothèse a été vérifiée par des expériences de co-précipitation avec NOT1c ou des protéines NOT2 et NOT5 amputées de leur extrémité N-terminale ne pouvait s'associer au sein du complexe.

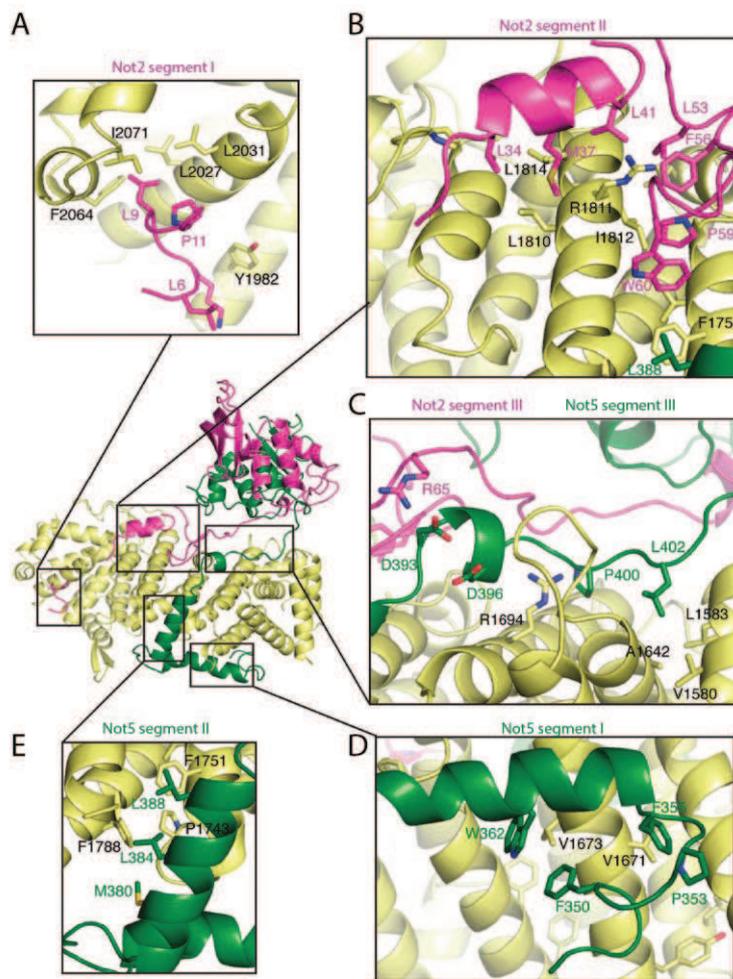


Figure 13 : NOT1 interagit avec les extensions de NOT2 et NOT5.

Panneau A-E : Zoom des régions d'interactions de NOT1 avec NOT2 et NOT5 montrant les trois segments (I,II,III) des extensions N-terminale de NOT2 et NOT5 qui forment les interactions avec NOT1. Les résidus conservés responsables des interactions sont indiqués.

Les domaines globulaires de NOT2 et NOT5 sont positionnés au-dessus des hélices B des répétitions HEAT 1-4 de NOT1, qui prennent en sandwich en partie les extensions N-terminales de NOT2 et NOT5. Ces domaines globulaires contiennent chacun une boîte NOT (NOT box). Celle de NOT2 (résidus 99-191) se compose de trois hélices et d'un feuillet beta formé par quatre brins beta antiparallèles. Le feuillet beta est extrêmement courbé, tandis que les brins β 3 et β 4 sont longs et incurvés avec un résidu glycine conservé (gly166) au point de pliure/courbure du brin β 3. Une courte extension C terminale se fixe contre le brin β 1 pour créer un petit tonneau beta.

La boîte NOT de NOT5 (résidus 464-560) est très similaire structurellement à celle de NOT2. En effet, ces deux domaines peuvent être superposés avec un écart quadratique moyen inférieur à 1,3 Å, sur l'ensemble des atomes de la chaîne principale (Figure 14). La différence principale réside dans le fait que les brins β de NOT5 sont plus courts, formant un feuillet β plus plat.

Une recherche d'homologie de structure des boîtes NOT avec le serveur DALI a identifié les domaines Sm comme structures présentant la plus forte similarité. Les protéines Sm, constitutants du coeur du spliceosome, sont au nombre de sept et s'associent pour former un anneau heptamérique entourant un segment du snARN (Salgado-Garrido et al. 1999). Toutefois les boîtes NOT diffèrent des domaines canoniques Sm par plusieurs aspects. Tout d'abord, il manque les motifs Sm1 et Sm2 dans la séquence d'acides aminés. D'un point de vue structural, les boîtes NOT ne possèdent pas le cinquième brin beta responsable de la dimérisation des domaines et a fortiori de l'oligomérisation en structure en anneau (Törö et al. 2002; Khusial et al. 2005; Leung et al. 2011). Les boîtes NOT utilisent un autre mode d'interaction, leur dimérisation s'opérant via leurs hélices α N-terminale. L'hélice α 1 de la boîte NOT de NOT2 se fixe contre la base du feuillet β de NOT5 et vice et versa. Au centre de l'interaction, les hélices α 2 de NOT2 et de NOT5 se fixent l'une à l'autre (Figure 14). L'interface de dimérisation est régie par des interactions multiples centrées sur les résidus conservés Phe114 et Leu115 de NOT2 et leurs correspondants chez NOT5 (Phe479 et Ile480). Enfin les extensions N-terminales des domaines globulaires, les résidus 67-93 de NOT2 s'enroulent autour de la boîte NOT de NOT5, et les résidus 408-427 de NOT5 s'enroulent autour de celle de NOT2 (Figure 14). Les extensions N-terminales quant à elles ancrent les boîtes NOT à NOT1, comme décrit plus haut.

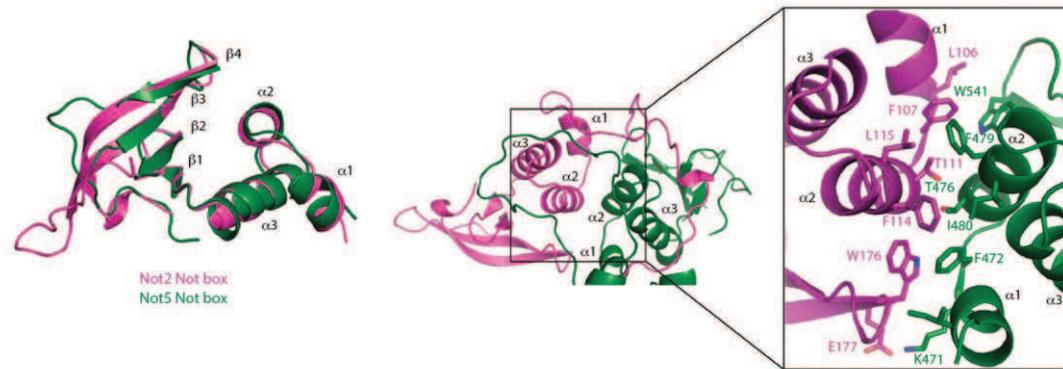


Figure 14 : Superposition et surface d'interaction des boîtes NOT de NOT2 et NOT5. Zoom des surfaces d'interaction entre les boîtes Not de NOT2 et NOT5. Les résidus conservés impliqués dans l'interaction sont indiqués.

D.5. Structure CNOT9 et CNOT1

Il a été montré chez l'homme que la sous-unité CNOT9 (homologue CAF40 chez la levure) du complexe CCR4-NOT interagissait avec la protéine CNOT1 par son domaine DUF3819 (Domain of Unknown Function), situé après (Bawankar et al. 2013). La structure de CNOT9 a été déterminée en 2007 (Garces et al. 2007). Le repliement de CNOT9 en complexe avec CNOT1 est identique à celui de la protéine seule (Figure 15). CNOT9 se compose d'un tandem de six répétitions de type armadillo (ARM). Ces répétitions ARM sont chacune formées par trois hélices (H1, H2 et H3) qui ensemble forment une structure en forme de croissant, avec une surface convexe délimitée par les hélices H1 et une surface concave délimitée par les hélices H2 et H3. Dans la structure de la protéine isolée, la surface convexe des répétitions ARM 2 et 3 fournit une interface pour la dimérisation. A contrario, dans le cadre du complexe CNOT9/CNOT1 (Mathys et al. 2014) cette surface est utilisée pour l'interaction avec CNOT1, qui se présente comme une molécule allongée articulée autour de trois longues hélices qui s'entrelacent. Cet ensemble est complété par deux segments avec des éléments hélicoïdaux. Dans la partie N-terminale se trouve un segment de 35 acides aminés formant une région hydrophobe à l'extrémité de la molécule. Dans la partie C-terminale se trouve un segment de 50 acides aminés qui s'enroule autour de l'autre extrémité de CNOT1 et de la première répétition ARM de CNOT9.

D'autre part il a été démontré que la protéine GW182, protéine d'échafaudage pouvant s'associer aux ARN messagers et stockée dans les P bodies (Braun et al. 2011; Zielezinski and Karlowski 2015), au travers de son domaine CED, interagissait avec la partie CNOT9 au travers des résidus tryptophanes (Braun et al. 2011; Chekulaeva et al. 2011; Fabian et al. 2011). Une expérience de trempage des cristaux du complexe CNOT9-CNOT1 dans une solution de

tryptophane a permis de résoudre la structure du complexe en présence de tryptophane. En effet, la molécule de tryptophane a pu être modélisée dans la surface concave entre les répétitions ARM5 et ARM6. Le groupement indole du tryptophane s'enfonce dans une poche hydrophobe formée par les résidus Tyr203, Phe206 et Ala248. La chaîne principale du tryptophane est exposée au solvant dans une position compatible avec une géométrie de la chaîne polypeptidique étendue de GW182. Une molécule de MES, issue du tampon de cristallisation, a pu être modélisée dans une poche située entre les répétitions ARM5 et ARM6. Cette seconde poche est également compatible avec l'interaction d'une molécule de tryptophane. Des mutants conçus pour perturber les poches d'interactions des tryptophanes bloquent totalement l'interaction de CNOT9 avec le domaine CED de GW182, prouvant ainsi que les deux poches étaient occupées synergiquement par les résidus tryptophanes de GW182 (Mathys et al. 2014)

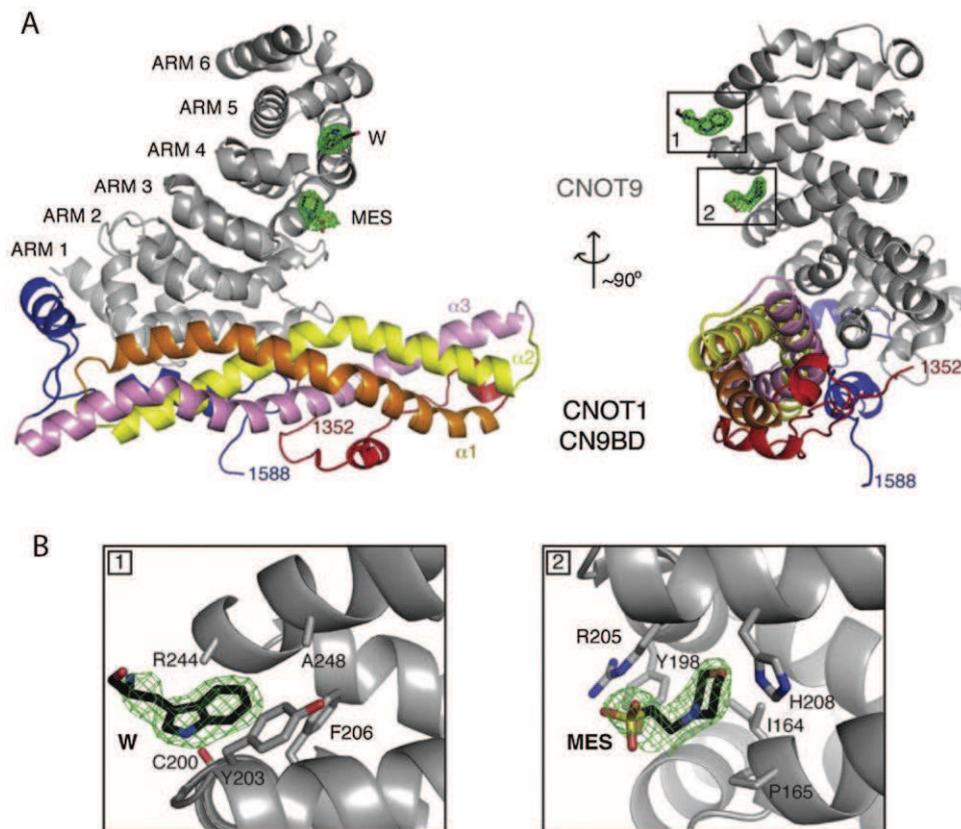


Figure 15 : Structure du complexe CNOT1 et CNOT9.

Panneau A : Structure de CNOT9 (en gris avec les 6 répétitions ARM numérotées) et CNOT1 (les différents éléments sont colorés de rouge à bleu). La structure est représentée selon deux orientations, avec la densité électronique pour les petites molécules (Trp et MES).

Panneau B : Zoom des deux poches d'interaction contenant les petites molécules (Trp et MES).

D.6. Le domaine central de CNOT1 interagit directement avec la protéine DDX6 à boîte DEAD

Le domaine central MIF4G (se liant à CAF1) de CNOT1 interagit via une petite région de résidus

conservés dans sa surface convexe (Basquin et al. 2012; Petit et al. 2012), bien que la surface la plus conservée se trouve sur la surface concave de CNOT1. D'autres études structurales ont pu mettre en évidence que les domaines MIF4G de plusieurs protéines interagissaient avec des hélicases/ATPases ARN dépendantes de la famille à boîte DEAD via leur surface concave, tels que eIF4A–eIF4G (Schütz et al. 2008), Gle1-Dbp5 (Montpetit et al. 2011), et CWC22-eIF4AIII (Buchwald et al. 2013). Nous avons émis l'hypothèse que CNOT1 pouvait interagir de façon similaire avec une protéine à boîte DEAD. Afin de confirmer cette hypothèse, nous avons exprimé une construction de la protéine humaine GST-NOT1 dans des cellules HEK293, effectué une purification d'affinité et analysé les interactants par spectrométrie de masse. Cette expérience a identifié DDX6, une ATPase/hélicase membre de la famille à boîte DEEA. DDX6 est un inhibiteur de la traduction et un activateur du décoiffage des ARNs et est aussi impliqué dans la répression des micro ARNs (Chu et Rana 2006, Su et al 2011), comme le complexe CCR4-NOT chez la levure (Maillet et al. 2002). Nous avons pu sur-exprimer les protéines CNOT1 et DDX6 et formé un complexe stable en chromatographie d'exclusion de taille. Il a été ensuite possible d'obtenir des cristaux qui ont abouti à la résolution de la structure tridimensionnelle (Mathys et al. 2014).

D.6.1. La structure du complexe CNOT1 MIF4G DDX6 explique le mode d'interaction spécifique et conservé

DDX6 possède un repliement caractéristique des protéines à boîte DEAD, avec deux domaines RecA contenant les résidus interagissant avec l'ARN et l'ATP (Henn et al. 2012). Nous avons dans premier temps résolu la structure de DDX6 sous forme isolée où les deux domaines se calent l'un contre l'autre adoptant une conformation très compacte impliquant de nombreux contacts intramoléculaires. Il est à noter que cette conformation est différente de celle observée chez l'orthologue de levure DDH1 (Cheng et al. 2005). Nous avons ensuite déterminé la structure de DDX6 en complexe avec le domaine MIF4G de CNOT1. Au sein du complexe les domaines RecA adoptent une configuration différente et contactent le domaine MIF4G de CNOT 1 par sa surface concave. L'extension N-terminale du domaine MIF4G contribue à l'interaction avec le domaine RecA2. Les répétitions HEAT 1 et 2 du domaine MIF4G interagissent avec le domaine RecA2 de DDX6, et la répétition HEAT 5 interagit avec le domaine RecA1. Nous avons pu identifier trois régions conservées d'interactions (Figure 16). La première est centrée autour des résidus Phe1101, Asn1105 et Lys 1114 de CNOT1, la deuxième autour des résidus Glu1142, Asn1144 et Phe1145 de CNOT1. Ces deux régions se situent à l'interface d'interaction avec le domaine RecA2. La troisième région contactant le domaine RecA1 est formé par les résidus Phe1281, Glu1284 et Lys1276 de

CNOT1. Ces interactions impliquent les résidus Arg105 et Phe112 du domaine RecA1 de DDX6 et les résidus Arg331, Gln333 et Asn335 du domaine RecA2.

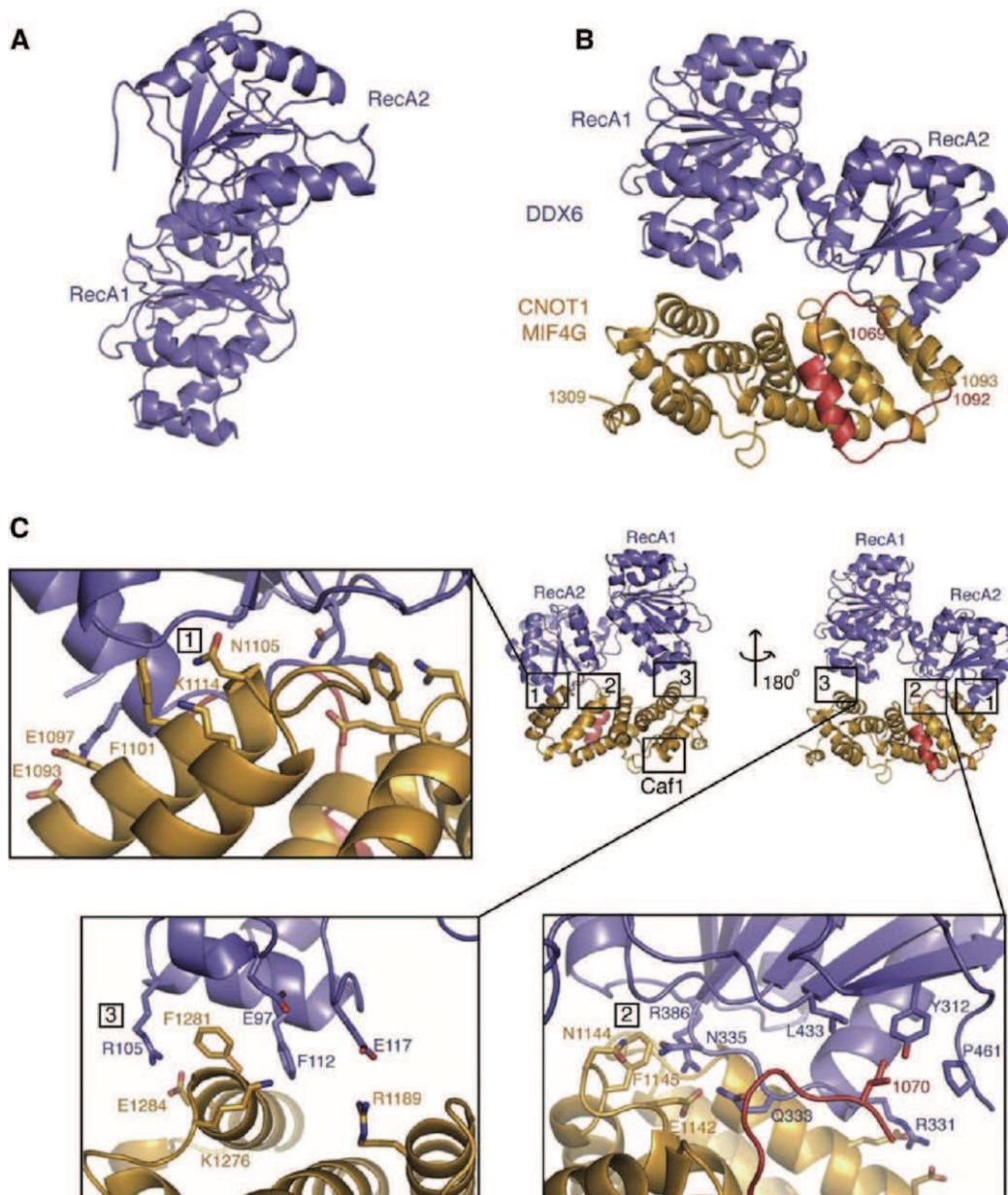


Figure 16 : Structure de DDX6 en complexe avec le domaine MIF4G de CNOT1(1069-1309).

Panneau A : Structure de DDX6 sous forme isolée avec ses deux domaines RecA.

Panneau B : Structure de DDX6 en bleu interagissant avec le domaine MIF4G de CNOT1. Les cinq répétitions HEAT sont colorées en orange et la partie N-terminale en rouge.

Panneau C : Régions d'interactions entre les deux protéines.

D.7. Structure de TTP en complexe avec CNOT1

La tristetraproline (TTP), protéine à répétitions de proline, peut se lier à l'ARN et interagir avec le complexe CCR4-NOT afin de réprimer l'expression des gènes. TTP est capable d'interagir avec les éléments riches en AU (AREs) se trouvant dans la région 3' de la séquence non codante des ARN messagers et d'induire ainsi leur déadénylation. Une étude structurale a identifié un motif

conservé dans la partie C terminale de TTP pouvant se lier à la protéine CNOT1 (Fabian et al. 2013). La structure révèle qu'un peptide de 14 acides aminés de TTP interagit avec une cavité hydrophobe conservée des hélices $\alpha 1$ et $\alpha 3$ du domaine MIF4G de CNOT1. La partie centrale du peptide forme deux tours d'hélice α qui interagissent avec la cavité, et sa partie C-terminale interagit électrostatiquement avec les résidus chargés de CNOT1. Des essais de déadénylation *in vitro* ont, par ailleurs, montré qu'un mutant de TTP, abolissant l'interaction avec CNOT1, affectait de façon significative l'activité déadénylase.

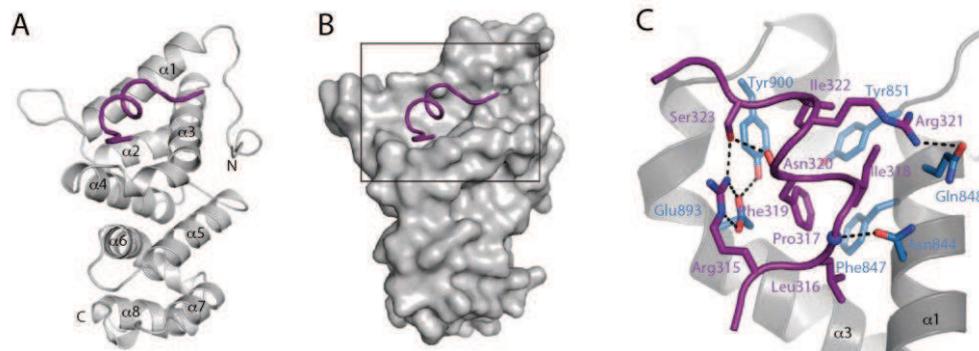


Figure 17 : Structure cristalline du complexe TTP-CNOT1 (adapté de (Fabian et al. 2013)).
Panneau A : Représentation de CNOT1 (820-999) en gris avec un peptide TTP en violet.
Panneau B : Représentation de la surface de CNOT1 avec le peptide TTP.
Panneau C : Interface d'interaction entre CNOT1 et le peptide de TTP.

D.8. Structure de NOT4

NOT4 a été identifiée par un criblage génétique comme étant un régulateur de la transcription (Liu et al. 1998; Collart et al. 1994). La structure du domaine RING de CNOT4 (orthologue humain de NOT4) a été déterminée par résonance magnétique nucléaire (RMN) (Hanzawa et al. 2001). Cette structure a révélé un nouveau motif C4C4 qui coordonne deux ions zincs de manière diagonale. Il a été montré que le domaine RING de CNOT4 possédait une activité ligase ubiquitine E3 avec UbcH5. Les résidus à l'interface de CNOT4-UbcH5 ont été identifiés par une approche combinée de RMN, de mutagenèse et un modèle théorique a été proposé (Albert et al. 2002; Winkler et al. 2004). Une seconde étude a mis évidence par purification d'affinité en tandem (TAP-tag) que Ubc4/Ubc5 (homologue de UbcH5 chez la levure) interagissaient avec le complexe CCR4-NOT (Mulder et al. 2007).

Le domaine RING de NOT4 régule de nombreuses voies cellulaires en ubiquitinant des substrats variés tels que le facteur de transcription Yap1, la Cycline C, la déméthylase d'histone Jhd2 et des facteurs s'associant aux ribosomes comme les facteurs de transcription NAC (Gulshan et al. 2012;

Laribee et al. 2007; Mulder et al. 2007; Cooper et al. 2012; Mersman et al. 2009). On pense que le domaine RING de NOT4 joue un rôle essentiel dans le maintien du niveau d'ubiquitine libre dans la cellule pour faciliter l'assemblage du protéosome (Panassenko and Collart 2011). D'autre part, une autre étude a révélé que NOT4 pouvait stimuler la voie de transduction du signal JAK/STAT (Grönholm et al. 2012).

Le domaine RING de NOT4 possède une hélice centrale qui est entourée de chaque côté par deux boucles coordonnées par deux ions zinc. Le domaine RING de NOT4 est très similaire à celui de CNOT4 (écart quadratique moyen inférieur à 1 Å). De même le repliement de Ubc4 en complexe avec le domaine RING de NOT4 est très similaire à la structure de Ubc4 isolé, avec une hélice N terminale suivie d'un feuillet β central, composé de 4 brins β , et de trois hélices C-terminales (COOK et al. 1993).

L'interface du complexe NOT4 (RING)-Ubc4 est principalement formée par les interactions des deux ions zinc coordonnant les boucles de NOT4 avec l'hélice 1, ainsi que deux régions de Ubc4 formant une boucle connectant le troisième et quatrième brin β et une boucle connectant le quatrième brin β à l'hélice 2 (Figure 18). La structure du complexe NOT4 (RING)-Ubc4 est très similaire au modèle théorique précédemment proposé (Albert et al. 2002; Mulder et al. 2007), à l'exception de différences subtiles qui incluent un mouvement modéré de l'hélice alpha de NOT4 et de l'hélice 1 de Ubc4.

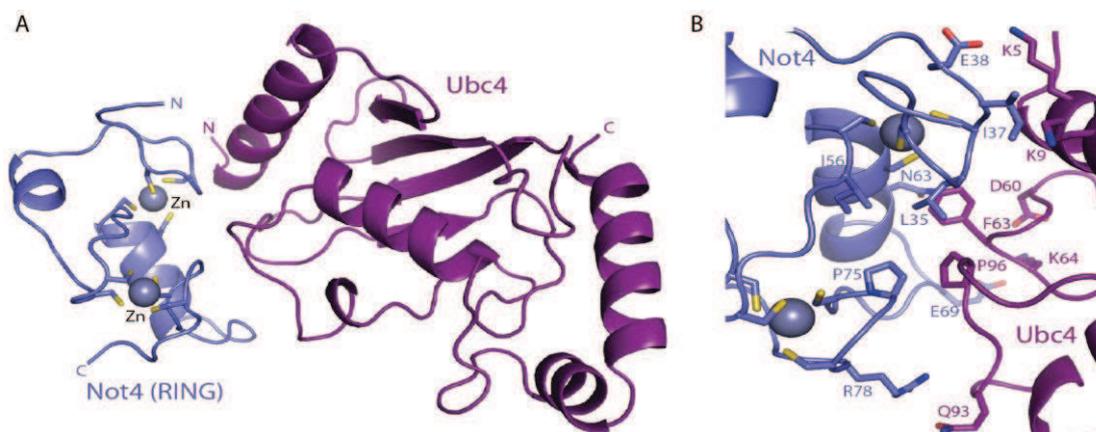


Figure 18 : Structure cristalline du complexe NOT4 (RING)-Ubc4.

Panneau A : Représentation du complexe NOT4 (RING)-Ubc4. NOT4 est coloré en bleu pâle et Ubc4 est coloré en violet.

Panneau B : Zoom des différents segments de NOT4 interagissant avec Ubc4.

Le centre de l'interaction est formé par l'insertion de la Phe63 de Ubc4 dans une poche hydrophobe formée par Leu35, Iso37, Iso56, Cys60, Asn63, Ile64, Pro75 de NOT4 et Asp60, Pro62 et Pro95 de Ubc4. Cette région hydrophobe est entourée de chaque côté par deux interactions de type pont cristallin. Le premier pont salin est formé par la Lys5 de Ubc4 et la Glu38 de NOT4. Le deuxième pont salin implique la lys64 de Ubc4 et le Glu69 de NOT4. D'autres liaisons hydrogènes impliquant Arg78 de NOT4 et Gln93 de Ubc4 sont présentes à la périphérie de l'interface.

L'analyse de la structure du complexe NOT4 (RING)-Ubc4 et sa comparaison avec les autres structures E2-E3 indique que la spécificité est principalement conférée par les trois interactions électrostatiques situées à la périphérie de la poche hydrophobe conservée. Tous les résidus impliqués dans cette interaction sont conservés uniquement chez les homologues de Ubc4/Ubc5, indiquant que Ubc4/Ubc5 serait la seule protéine qui agirait comme enzyme de type E2 pour NOT4.

D.9. Structure de NOT1 en complexe avec NOT4

L'architecture de NOT1c(1541-2093) en complexe avec la partie C-terminale de NOT4 ou NOT4c(418-477) est très similaire à la structure de NOT1c-NOT2-NOT5c (Bhaskar et al. 2013). Le premier résidu de NOT1 visible dans cette structure est la Val1567. La plupart des résidus de NOT1c sont bien représentés dans la carte de densité électronique, à l'exception des régions entre les acides aminés 1791 à 1800, et, 2065 à 2071. Une différence notable par rapport à la structure NOT1c-NOT2-NOT5c est un changement de conformation dans la région entre les résidus 1917 à 1930, qui est prédite comme n'étant pas structurée et qui connecte les répétitions HEAT 7 et 8. Dans la structure précédente, cette région adopte une conformation en épingle à cheveux β , qui implique un contact cristallin, mais dans la structure NOT1c-NOT4c cette région est déstructurée et contacte la molécule adjacente dans le cristal.

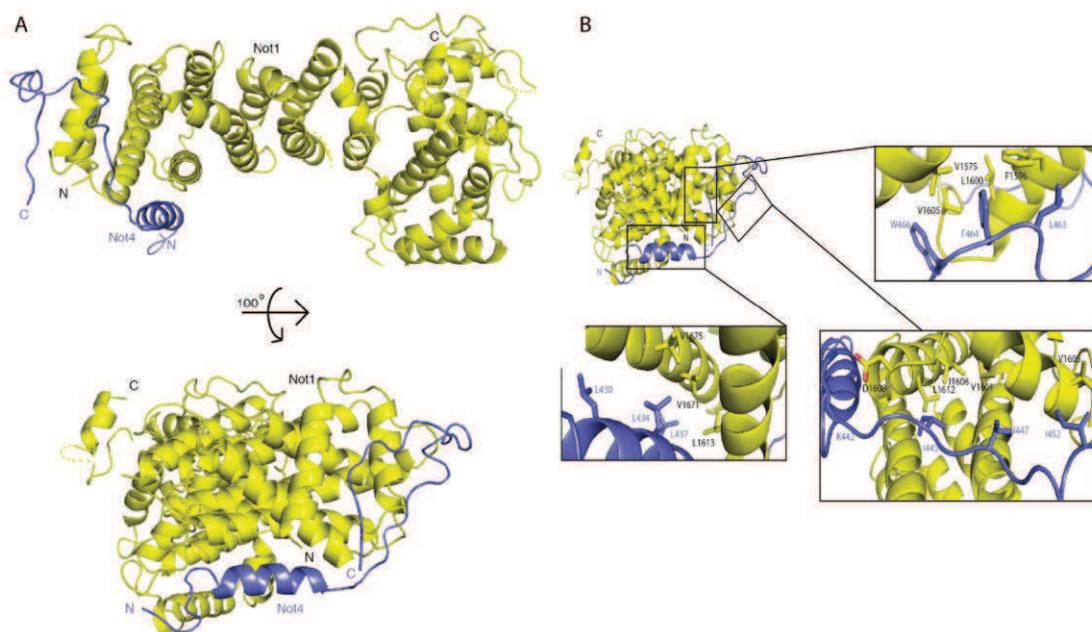


Figure 19 : Structure cristalline du complexe NOT1c(1541-2093)-NOT4c(418-477).

Panneau A : Représentation du complexe NOT1c-NOT4c. NOT4c est coloré en bleu pale et NOT1c est coloré en jaune.

Panneau B : Zoom des différents segments de NOT4c interagissant avec NOT1c.

D.9.1. NOT4c entoure les répétitions HEAT N-terminales de NOT1c

NOT4c entoure les trois premières répétitions HEAT de NOT1c (partie C-terminale de NOT1), s'étirant sur une longueur de presque 100 Å et couvrant une surface de 3020 Å². Ces importants contacts entre NOT4c et NOT1c peuvent être divisés en trois segments. Le premier segment implique une région composée d'une hélice et d'une boucle où l'hélice α de NOT4c (résidus 428-439) se cale contre les hélices α de la seconde et de la troisième répétition de NOT1c. Cette interface est principalement caractérisée par des interactions hydrophobes impliquant les résidus Leu430, Leu434 et Leu437 de NOT4c et les résidus Leu1613, Leu1671, Leu1674 et Val1675 de NOT1c. Le second segment se compose des régions entre les résidus 442-452 de NOT4, qui interagissent avec deux boucles de NOT1c, la première connectant les répétitions HEAT 1 et 2 et la seconde connectant les répétitions HEAT 3 et 4. Une partie de ce segment interagit aussi avec l'hélice B de la première répétition HEAT de NOT1c. Cette interface est principalement dominée par des interactions polaires et implique également un nombre réduit de contacts hydrophobes. Le troisième segment de NOT4c (résidus 462-469) est très structuré et se cale entre les hélices A et B de la première répétition HEAT de NOT1c. Cette région de NOT4c se compose de résidus hydrophobes et aromatiques tels que Val1575, Phe1596, Leu1600 et Val1605 de NOT1c. La structure suggère que NOT4c contacte NOT1c à travers deux régions d'interactions majeures, la région N-terminale de NOT4c (résidus 428-439) et la région C-terminale de NOT4c (résidus 462-469) formant un motif décomposé en deux parties. Afin de tester l'implication des différentes interfaces nous avons généré deux mutants de troncation GST-NOT4442-477 (GST-NOT4c- Δ N) et GST-NOT4418-462 (GST-NOT4c- Δ C) et testé leur capacité à se lier à NOT1c dans des tests d'interaction. La GST a été utilisé comme témoin négatif et GST-NOT4c comme contrôle positif. Alors que GST-NOT4c- Δ N peut interagir de la même manière que GST-NOT4c, GST-NOT4c- Δ C n'est pas capable d'interagir avec NOT1c. Puisque seule la région C-terminale de NOT4 abolit l'interaction avec NOT1c, nous concluons que cette région est indispensable pour l'association avec NOT1 et les deux autres segments sont des surfaces d'interactions additionnelles.

E. Contribution de mon travail de thèse

Les différents travaux sur lesquels j'ai eu l'occasion de travailler depuis 2010 ont conduit à la résolution des structures des différents modules du complexe CCR4-NOT. Mon travail a contribué à élucider l'organisation générale du complexe ainsi que de révéler des détails structuraux qui ont permis de contribuer à comprendre et confirmer certaines des fonctions biologiques du complexe. D'autre part, mon travail a aussi permis de mettre en lumière la connexion du complexe CCR4-NOT avec la répression de la traduction. Ces travaux apportent des données précieuses pour les chercheurs dont l'intérêt est focalisé sur la dégradation des ARN et plus généralement sur la régulation de l'expression des gènes. Les travaux qui sont présentés dans ce manuscrit, en dépit de leurs relatives nouveautés, ont déjà été cités à de nombreuses reprises.

Etudes fonctionnelles et structurales du module nucléase du complexe CCR4-NOT

E.1. Etudes fonctionnelles et structurales du module nucléase du complexe CCR4-NOT

E.1.1. Contexte et objectifs du projet

Le complexe CCR4-NOT est un des complexes majeurs de la dégradation des ARN. Ce complexe est étudié depuis de nombreuses années, mais très peu d'informations sur sa structure étaient connues lors de l'initiation de ce projet par notre laboratoire. Comme ce complexe, de par sa taille ainsi que par son organisation multi-modulaire, représente un véritable défi pour une approche par cristallographie aux rayons X, nous avons décidé de mettre en place une stratégie de résolution des structures des différents modules de ce complexe.

E.1.2. Approches expérimentales

Identification de la région de NOT1 s'associant à la déadénylase CCR4-NOT

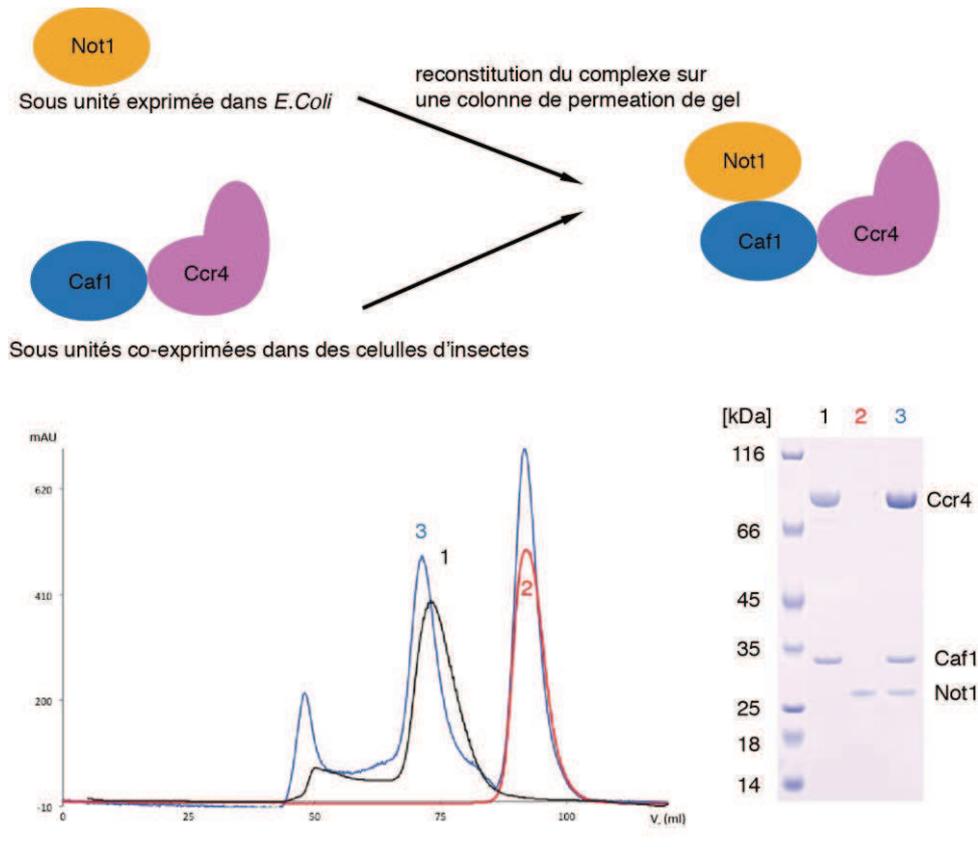
La protéine NOT1 de levure (*S.cerevisiae*) se compose de 2108 acides aminés. Des expériences de double hybride ont montré que la région de NOT1 interagissant avec les protéines CCR4 et CAF1 se situait entre les résidus 667 à 1152. Afin d'identifier de manière plus précise cette région de NOT1, j'ai utilisé une méthode combinant les approches bioinformatique et expérimentale. L'analyse de séquences a mis en évidence une large région structurée se situant dans les premiers 1000 résidus de NOT1, ce qui m'a conduit à sur-exprimer et purifier une construction de NOT1(1-1000). Afin d'identifier des domaines, j'ai utilisé la protéolyse ménagée qui a mis en évidence l'existence deux larges fragments : NOT1(154-753) et NOT1(754-1000). Ces fragments ont ensuite été sur-exprimés et purifiés individuellement.

Résolution par étape des structures cristallines du module de déadénylation du complexe CCR4-NOT

Pour les études structurales, il est nécessaire de surproduire et de purifier la protéine d'intérêt avec un haut degré de pureté (>95%). Les différentes constructions de NOT1 ont été sous-clonées dans un vecteur plasmidique de type pET codant pour une étiquette His-Sumo clivable positionnée à l'extrémité N-terminale, ce qui a permis de sur-exprimer les protéines NOT1 dans *E.coli* et de les purifier de façon optimale et efficace par chromatographie d'affinité. Le complexe CCR4-(résidus 110-837)-CAF1(résidus 146-433) a été cloné dans un vecteur pFastBac DUAL vector (Système Baculovirus Invitrogen) qui a été utilisé pour générer des bacmides et des virus pour l'infection de

cellules d'insectes de type SF9.

Des expériences de chromatographie d'exclusion de taille m'ont permis d'identifier le fragment de NOT1 pouvant s'associer de façon stable à la protéine CAF1. Finalement, j'ai pu reconstituer un complexe ternaire stœchiométrique tout d'abord en mixant les protéines CAF1, CCR4 (exprimées dans les cellules d'insectes) et le fragment 754–1000 de la protéine NOT1 (exprimée dans *E.coli*) en léger excès, puis en purifiant le complexe sur une colonne de chromatographie d'exclusion de taille.



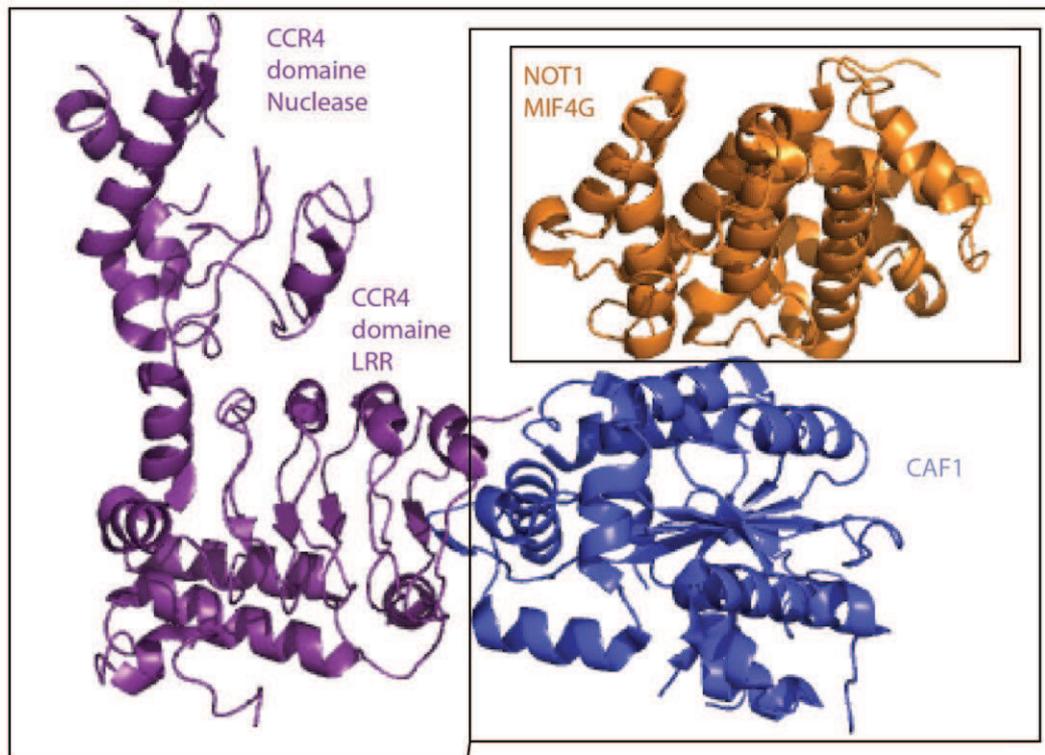
Afin de caractériser ces interactions de manière précise, j'ai entrepris de cristalliser les différents composés du complexe et de résoudre leurs structures tridimensionnelles par cristallographie aux rayons X.

Pour la recherche des conditions de cristallisation de ce projet, j'ai utilisé de manière systématique quatre matrices clairsemées : *Magic I suite (inhouse)*, *Magic II suite (inhouse)*, *Index (Hampton Research)*, *PEGs suite (Qiagen)*. Pour ces expériences, j'ai utilisé la technique de diffusion de vapeur en gouttes assises, sur des plaques 96 puits contenant 1 réservoir. Cette approche m'a permis de tester 384 conditions de cristallisation et a été réalisée grâce à un robot « *Phoenix* », sur la plateforme de cristallisation de l'institut Max-Planck de biochimie. J'ai isolé plusieurs conditions

de cristallisation, et travaillé en optimisation manuelle pour obtenir des cristaux de taille compatible ($100 \mu\text{m}^3$) avec les expériences de diffraction des rayons X.

Il a été possible de cristalliser et de résoudre les structures cristallines des deux fragments N-terminaux de la protéine NOT1 (NOT1154–753 et NOT1754–1000). Ces deux structures ont été déterminées après l'obtention de phases expérimentales. Pour le domaine NOT1754–1000, le signal anomal de l'iode contenu dans la solution de cristallisation a été utilisé. Les cristaux du fragment 154–753 ont été trempés dans une solution de $\text{KAu}(\text{CN})_2$. J'ai ensuite réussi à cristalliser un complexe binaire composé du fragment NOT1754–1000 et de la protéine CAF1. J'ai finalement pu reconstituer un complexe ternaire composé des protéines CCR4, CAF1 et du domaine NOT1754–1000. Ce complexe a été cristallisé et sa structure a été résolue par remplacement moléculaire.

Structure cristalline du domaine MIF4G de NOT1
Résolution: 1,7 Å
Phasage expérimental avec de l'iode
pdb: 4B89



Structure cristalline du complexe ternaire CCR4-CAF1-NOT1
Résolution: 3,4 Å
Remplacement moléculaire
pdb: 4B8C

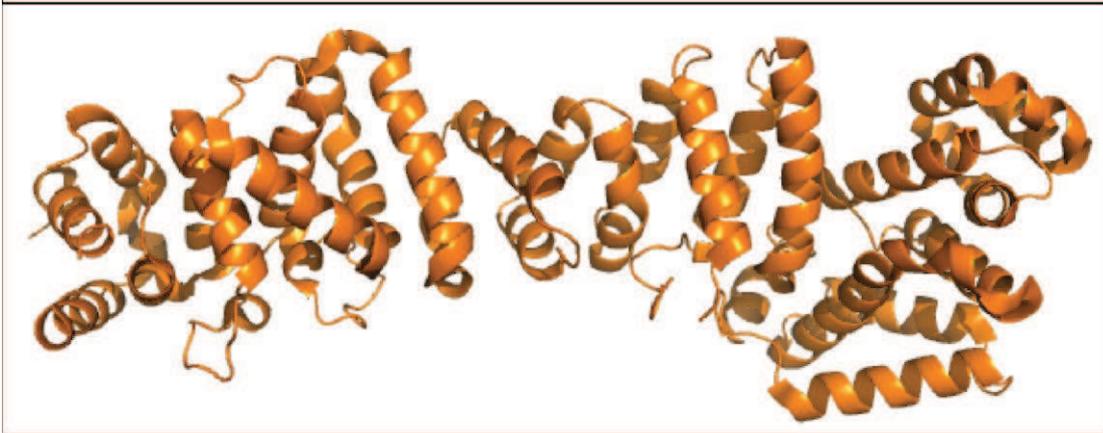
Structure cristalline du complexe binaire CAF1-NOT1
Résolution: 2,6 Å
Remplacement moléculaire
pdb: 4B8A

Structure cristalline du domaine N-terminal de NOT1(153-753)

Résolution: 2,6 Å

Phasage expérimental avec un composé contenant l'or (KAu(CN)₂)

pdb: 4B8B



E.1.3. Résumé de l'étude

Cette étude structurale a permis de mettre en évidence que la partie N terminale de NOT1 était composée de répétitions HEAT qui sont caractéristiques de domaines adoptant le repliement MIF4G. Un domaine MIF4G est positionné dans la partie centrale de NOT1 qui est responsable de l'interaction avec la protéine CAF1, qui à son tour, contacte le domaine LRR de la protéine CCR4 qui recrute le domaine nucléase. Les interactions qui forment le module de déadénylation du complexe sont conservées à travers l'évolution. La perturbation de ces interactions agit sur la viabilité cellulaire et sur la déadénylation des ARN messagers *in vivo* chez la levure. La partie N-terminale de NOT1 forme une plateforme étendue caractéristique des protéines d'échafaudage et place le module de déadénylation dans une position pivot au centre de complexe CCR4-NOT1.

Architecture of the Nuclease Module of the Yeast Ccr4-Not Complex: the Not1-Caf1-Ccr4 Interaction

Jérôme Basquin,^{1,3} Vladimir V. Roudko,^{2,3} Michaela Rode,¹ Claire Basquin,¹ Bertrand Séraphin,² and Elena Conti^{1,*}

¹Department of Structural Cell Biology, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

²Equipe Labellisée La Ligue, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM) U964/Centre National de Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 67404 Illkirch, France

³These authors contributed equally to this work

*Correspondence: conti@biochem.mpg.de

<http://dx.doi.org/10.1016/j.molcel.2012.08.014>

SUMMARY

Shortening eukaryotic poly(A) tails represses mRNA translation and induces mRNA turnover. The major cytoplasmic deadenylase, the Ccr4-Not complex, is a conserved multisubunit assembly. Ccr4-Not is organized around Not1, a large scaffold protein that recruits two 3'-5' exoribonucleases, Caf1 and Ccr4. We report structural studies showing that the N-terminal arm of yeast Not1 has a HEAT-repeat structure with domains related to the MIF4G fold. A MIF4G domain positioned centrally within the Not1 protein recognizes Caf1, which in turn binds the LRR domain of Ccr4 and tethers the Ccr4 nuclease domain. The interactions that form the nuclease core of the Ccr4-Not complex are evolutionarily conserved. Their specific disruption affects cell growth and mRNA deadenylation and decay *in vivo* in yeast. Thus, the N-terminal arm of Not1 forms an extended platform reminiscent of scaffolding proteins like eIF4G and CBP80, and places the two nucleases in a pivotal position within the Ccr4-Not complex.

INTRODUCTION

Since the discovery of the poly(A) tail over 40 years ago, a wealth of functional information has accumulated on the major influence it exerts on the posttranscriptional regulation of eukaryotic gene expression (Edmonds et al., 1971; reviewed in Zhang et al., 2010; Eckmann et al., 2011). The poly(A) tail consists of a long stretch of adenosine nucleotides and is coated by multiple poly(A)-binding proteins in the nucleus and in the cytoplasm (Mangus et al., 2003). The resulting structure of the mature messenger ribonucleoprotein particle (mRNP) protects the 3' extremity of the messenger RNA (mRNA) from exoribonucleolytic degradation, thus increasing mRNA stability (reviewed in Zhang et al., 2010). It also increases translational efficiency by promoting the recruitment of translation initiation factors (Tarun et al., 1997; reviewed in Sonenberg and Dever, 2003). Conversely,

poly(A) tail shortening, a process known as deadenylation, is linked to cytoplasmic mRNA decay and to translational repression (Decker and Parker, 1993; reviewed in Goldstrohm and Wickens, 2008; Zhang et al., 2010).

In yeast and mammalian cells, deadenylation begins with the Pan2-Pan3 complex (Boeck et al., 1996; Brown and Sachs, 1998; Yamashita et al., 2005) and continues with the Ccr4-Not complex (Daugeron et al., 2001; Tucker et al., 2001, 2002; Chen et al., 2002). Ccr4-Not is thought to reduce the poly(A) tail to a short oligo(A) tract before the body of the mRNA is degraded by subsequent enzymatic activities (the exosome-Ski complex, the decapping complex, and Xrn1) (reviewed in Garneau et al., 2007). The core enzymes and regulators of this pathway are universally conserved throughout eukaryotes. Studies in yeast have shown that neither the Pan2-Pan3 nor the Ccr4-Not deadenylase activities are essential *in vivo* (Malvar et al., 1992; Boeck et al., 1996; Brown et al., 1996). Synthetic phenotypes observed in double mutants of the deadenylase complexes suggest that they act in a partially redundant manner (Tucker et al., 2001).

While the interplay between Pan2-Pan3 and Ccr4-Not is not fully understood, it is clear that their concerted action is often the rate-limiting step for mRNA decay (Cao and Parker, 2001) and a focal point for regulation. The direct recruitment of Ccr4-Not to specific regulators impacts for example ARE-mediated decay (Sandler et al., 2011), microRNA-mediated gene silencing (Braun et al., 2011; Chekulaeva et al., 2011; Fabian et al., 2011), and translational repression during developmental processes (de Moor et al., 2005). Finally, Ccr4-Not has been proposed to modulate mRNA levels by acting at the transcription level, via interactions with transcription factors (reviewed in Collart and Timmers, 2004) and with the elongating RNA polymerase II (Denis et al., 2001; Kruk et al., 2011). Recently, altered transcription in Ccr4-Not mutants was also shown to arise in compensation of reduced mRNA decay (Sun et al., 2012).

The core subunits of Ccr4-Not were first identified in the early nineties (reviewed in Collart and Timmers, 2004). Biochemical approaches later indicated that these gene products exist in the context of a multiprotein complex with an approximate mass of 1 MDa (Liu et al., 1998; Maillet et al., 2000; Chen et al., 2001). Yeast two-hybrid and coimmunoprecipitation studies suggest that Ccr4-Not is organized around Not1 (also

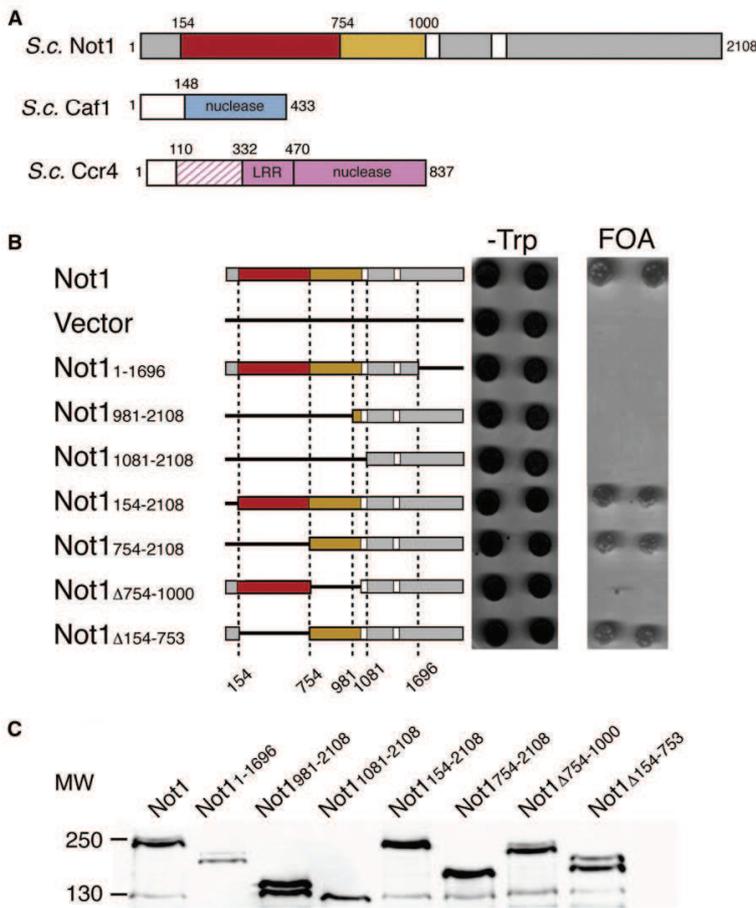


Figure 1. The Caf1-Ccr4-Binding Domain of Not1 Is Essential In Vivo

(A) Schematic representation of the domain arrangement of yeast Not1, Caf1, and Ccr4. The Not1₁₅₄₋₇₅₃ and Not1₇₅₄₋₁₀₀₀ domains were identified biochemically (Figure S1) and are shown in red and orange, respectively. Other predicted structured regions in Not1 are colored in gray. The nuclease domain of Caf1 is in blue, and the LRR and nuclease domains of Ccr4 in pink. The N-terminal region of yeast Ccr4 present in the crystal structure reported in this manuscript is shaded in pink. Unstructured regions are in white. (B) Analysis of the ability of Not1 truncation mutants to complement a Not1 deletion. *TRP1*-marked plasmids encoding different truncations of Not1 (fused to the TAP tag) are schematized on the left. The plasmids were introduced in strain T26N28 carrying a chromosomal deletion of *NOT1* complemented by a wild-type copy of the gene on a *URA3*-marked plasmid. Complementation was assayed by monitoring growth on FOA, growth on -Trp media serving as a control.

(C) Analysis of the levels of Not1 protein accumulating in the wild-type strain and mutants expressing truncated forms of the protein. Total proteins extracted from the various strains were fractionated by SDS-PAGE, transferred to a membrane, and detected through the TAP tag present at the C terminus of the proteins. See also Figure S1.

known as Cdc39), a 240 kDa protein believed to function as a structural scaffold (Collart and Timmers, 2004). The N-terminal half of Not1 recruits the two interacting exoribonuclease subunits Caf1 (also known as Pop2 in yeast) and Ccr4, forming the "nuclease module" of the complex (Dupressoir et al., 2001; Draper et al., 1995; Bai et al., 1999). The C-terminal half of Not1 interacts with Not2, Not3, Not5, and Not4, forming the so-called "Not module" (Bai et al., 1999; Deluen et al., 2002). The overall architecture of the Ccr4-Not complex is likely conserved in humans, *Drosophila*, and trypanosomes, although the complexity of the interactions and the interplay of the nuclease activities differ (Albert et al., 2000; Temme et al., 2004; Schwede et al., 2008; Lau et al., 2009). Perhaps the most striking difference lies in the balance of the nuclease activities: while Ccr4 is the primary deadenylase in yeast (Chen et al., 2002; Tucker et al., 2002), in mammalian cells Caf1 contributes significantly, if not predominantly, to poly(A) tail shortening (Mauxion et al., 2008; Sandler et al., 2011).

Despite the importance of the Ccr4-Not complex, limited structural information is currently available. A negative-stain

electron microscopy study of Ccr4-Not revealed an overall L-shape architecture (Nasertorabi et al., 2011) and the crystal structures of the individual nuclease domains of Caf1 and Ccr4 elucidated their different active sites (Thore et al., 2003; Jonstrup et al., 2007; Horiuchi et al., 2009; Wang et al., 2010). However, how the Not1-Caf1-Ccr4 network of interactions is established and how it positions the two exoribonucleases within the complex is unknown. To address these questions, we undertook a structural and functional analysis of the nuclease module of the Ccr4-Not complex.

RESULTS AND DISCUSSION

Domain Organization of the 114 kDa N-Terminal Region of Yeast Not1

S. cerevisiae Not1 is a 2,108 amino acid residue protein (Figure 1A). Yeast two-hybrid experiments previously mapped the interaction of Caf1 and Ccr4 to residues 667-1,152 of Not1 (Bai et al., 1999). To precisely identify the domains present in this portion of Not1, we used a combination of computational and experimental approaches. Sequence analysis and secondary structure predictions indicated the presence of a largely structured region in the first 1,000 residues of Not1, followed by a less-conserved, low-complexity segment. We expressed and purified a construct of yeast Not1 encompassing residues

1–1,000 (Not1_{1–1000}), subjected it to limited proteolysis, and identified the domain boundaries of the two largest fragments as corresponding to Not1_{154–753} and Not1_{754–1000} (Figure S1A available online). These proteolytically resistant fragments could be separately expressed and purified to homogeneity (Figure S1A) and did not coelute in size-exclusion chromatography (data not shown). These results indicated that Not1_{154–753} and Not1_{754–1000} are stable and folded as individual domains.

The Not1_{754–1000} Domain Recruits Caf1-Ccr4 and Is Essential for Function in Yeast

Next, we identified the minimal domains required for the formation of a core Not1-Caf1-Ccr4 complex by mixing individually purified components in size-exclusion chromatography experiments (Figure S1B). These experiments were guided by previous coimmunoprecipitation studies showing that Caf1 bridges the interaction between Not1 and Ccr4 (Bai et al., 1999). We found that Not1_{754–1000} was sufficient to form a ternary complex with the nuclease domain of Caf1 (hereafter referred to as Caf1) and with a construct of Ccr4 encompassing part of the N-terminal region, the leucine-rich repeat (LRR) domain and the nuclease domain (residues 110–837, hereafter referred to as Ccr4) (Figures 1A and S1B) (Simón and Séraphin, 2007). The interacting regions of Caf1 and Ccr4 include the functionally important residues that had been previously identified (Bai et al., 1999; Clark et al., 2004).

To delineate precisely the regions of the N-terminal half of Not1 important for function *in vivo*, we generated plasmids expressing Not1 full-length, Not1 truncation mutants (Not1_{154–2108}, Not1_{754–2108}, Not1_{981–2108}, Not1_{1081–2108}), and Not1 deletion mutants (Not1 Δ _{754–1000}, Not1 Δ _{154–753}) with a C-terminal tandem affinity purification (TAP) tag (Figure 1B). As a control, we generated a plasmid with a C-terminal truncation (Not1_{1–1896}) that eliminates an essential portion of Not1 involved in the interaction with Not2/3/5 (of unknown function) and Not4 (a ubiquitinating enzyme) (Maillet et al., 2000; Collart and Timmers, 2004). These plasmids were introduced into a *not1* Δ haploid strain rescued by a *URA3*-marked *NOT1* plasmid. Growth phenotypes were assayed after counterselection of the *URA3* plasmid on 5-FOA. Wild-type Not1 as well as the Not1_{154–2108} and Not1_{754–2108} truncations complemented the chromosomal *not1* deletion (Figure 1B). In contrast, Not1_{1081–2108} and Not1_{1–1896} did not rescue cell growth on synthetic media (Figure 1B), consistent with the very slow growth reported only in rich media for Not1_{1319–2108} and Not1_{1151–2108} in a different genetic background (Maillet et al., 2000). Strikingly, the larger Not1 Δ _{154–753} deletion was viable while the smaller Not1 Δ _{754–1000} deletion was not (Figure 1B). Western blot detection of the TAP-tag confirmed that these phenotypes did not result from protein instability as all mutants accumulated at wild-type levels or higher (Figure 1C). We concluded that, in addition to the C-terminal region of Not1, the Caf1-Ccr4-binding domain (Not1_{754–1000}) is essential for Not1 function.

Overall Structure of the Not1_{754–1000}-Caf1-Ccr4 Complex

To determine the structure of the *S. cerevisiae* Not1_{754–1000}-Caf1-Ccr4 core complex, we proceeded in a stepwise manner.

We first solved the structure of the Not1_{754–1000} domain by iodine-based single anomalous diffraction (SAD) and refined it against a native data set at 1.4 Å resolution to an *R*_{free} of 21.0% (Table 1). We then determined the crystal structure of the Not1_{754–1000}-Caf1 binary complex by molecular replacement (using the coordinates of Not1_{754–1000} and of the previously published yeast Caf1 structure [Thore et al., 2003] and refined it at 2.3 Å resolution to an *R*_{free} of 23.8% (Table 1). Finally, we determined the structure of the ternary Not1_{754–1000}-Caf1-Ccr4 complex by molecular replacement using the coordinates of Not1_{754–1000}-Caf1, followed by iterative rounds of model building and refinement.

The final model of Not1_{754–1000}-Caf1-Ccr4 was refined to 3.4 Å resolution with an *R*_{free} of 27.4% and good stereochemistry (Table 1). The crystals contain four independent copies of the complex in the asymmetric unit. For each copy, the model includes Not1 residues 759–991, Caf1 residues 149–428 (with the exception of a disordered loop between 357–369), and Ccr4 residues 135–187 and 242–261 in the N-terminal region and 332–470 in the LRR domain (Figure S2). In two copies of the asymmetric unit, ordered electron density is also present for about half of the Ccr4 nuclease domain (Supplemental Experimental Procedures).

The Not1_{754–1000}-Caf1-Ccr4 complex is built by the sequential interaction of a MIF4G (middle domain of initiation factor 4G), nuclease, LRR, and nuclease domains (Figure 2). The relative orientation of the domains within the ternary complex is the same in the different copies of the asymmetric unit (Figure S3A). Below we describe the Not1_{754–1000}-Caf1-Ccr4 structure of one of the two copies in the asymmetric unit where the nuclease domain is partially ordered (Figure 2).

The Not1_{754–1000} Domain Has a MIF4G Fold

The Not1_{754–1000} domain contains ten antiparallel α helices (Figure 3A, left panel). The arrangement is characteristic of that observed in HEAT-repeat proteins. HEAT motifs consist of two antiparallel α helices (known as A and B helices) and pack against each other in an almost parallel fashion (Andrade et al., 2001). Database searches for structural similarities to Not1_{754–1000} using the program Dali (Holm and Rosenström, 2010) indeed identified the HEAT-repeat fold of MIF4G (Figure 3A, left panel). The five HEAT repeats of Not1_{754–1000} (residues 782–976) superpose with MIF4G domains of eIF4G (Schütz et al., 2008) and UPF2 (Kadlec et al., 2004) with a root-mean-square deviation (rmsd) of 2.8 Å for more than 70% of all α carbon atoms. In addition, Not1_{754–1000} features two extended segments at the N and C termini (Figure 3A, left panel) that are anchored to the HEAT-repeat core by extensive hydrophobic interactions (Figure S3B), and thus form an integral part of the domain. The entire Not1_{754–1000} domain is essentially unaltered when comparing the structures determined in the absence and presence of Caf1 (rmsd of 0.53 Å over more than 98% of α carbon atoms) (Figure S3C). Analysis of the Not1_{754–1000} structure shows that the concave surface of the MIF4G domain is characterized by a large patch of conserved residues contributed by the B helices of HEAT 1 to 5 (Figure 3A, right panel). A small patch of conserved residues is also present on the opposite surface (Figure 3A, right panel), at the interrepeat loops connecting HEAT 3–4 and HEAT 4–5.

Table 1. Crystallographic Data Collection/Refinement Statistics

Data Set	Not1 ₇₅₄₋₁₀₀₀	Not1 ₇₅₄₋₁₀₀₀ (I SAD)	Not1 ₇₅₄₋₁₀₀₀ -Caf1	Not1 ₇₅₄₋₁₀₀₀ -Caf1-Ccr4	Not1 ₁₆₄₋₇₅₃ (Au SAD)
Data Collection					
Beamline	SLS PXII	Xcalibur Nova (CuK α)	SLS PXII	SLS PXII	SLS PXIII
Space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2 ₁	P1	P2 ₁
Unit cell parameters (Å)	a = b = 53.6, c = 187.2	a = b = 53.6, c = 183.5	a = 71.8, b = 76.56, c = 109.9	a = 122.6, b = 122.9, c = 126.4, α = 89.4, β = 89.7, γ = 64.2	a = 55.0, b = 164.5, c = 86.1, β = 100.1°
Wavelength (Å)	1.000	1.540	1.000	1.000	1.0396
Resolution range (Å) ^a	93.6–1.4 (1.48–1.41)	20.3–2.6 (2.74–2.60)	47.7–2.3 (2.42–2.30)	48.2–3.4 (3.6–3.4)	47.7–2.8 (2.95–2.8)
Unique reflections ^a	54,005 (7,437)	8,699 (1,265)	27,615 (3,835)	90,023 (13,150)	36,994 (5,266)
Multiplicity ^a	7.0 (6.9)	42.5 (23.9)	6.0 (5.9)	3.6 (3.7)	4.8 (4.8)
Completeness (%) ^a	99.8 (99.6)	97.3 (98.9)	99.5 (96.5)	98.5 (98.4)	99.5 (97.1)
Anomalous completeness (%) ^a		98.6 (99.1)			97.5 (93.2)
I/ σ (I) ^a	22.2 (2.7)	40.4 (12.6)	15.3 (4.4)	10.0 (2.9)	20.0 (2.6)
R _{sym} (%) ^a	3.7 (59.4)	10.2 (30.0)	8.6 (45.0)	11.2 (51.1)	5.2 (53.0)
Refinement					
R _{work} (%)	19.5		19.0	23.5	23.7
R _{free} (%)	21.0		23.8	27.4	27.3
Rmsd bond (Å)	0.007		0.007	0.014	0.009
Rmsd angle (°)	1.16		1.05	1.64	1.28
Ramachandran values: Favored (%)	98.3		97.3	93.5	96.7
Ramachandran values: Disallowed (%)	0.0		0.0	0.0	0.0

^aValues in parentheses correspond to the highest-resolution shell.

The Interaction of Not1₇₅₄₋₁₀₀₀ with Caf1 Involves Evolutionary Conserved Residues

Caf1 is recognized by the small conserved patch of the Not1₇₅₄₋₁₀₀₀ MIF4G domain formed by the interrepeat loops of HEAT 3–4 and HEAT 4–5 (Figures 3A and 3B). The structure of Caf1 has been previously described (Thore et al., 2003; Jonstrup et al., 2007; Horiuchi et al., 2009). In brief, Caf1 has an RNase D fold characteristic of the DEDD superfamily of exoribonucleases. The fold contains a β sheet to a large extent surrounded by α helices, with the exception of a small surface area that forms the active-site cavity (Figure 2A). The active site of *S. cerevisiae* Caf1 has a suboptimal Ser188-Glu190-Asp310-Gln394 (SEDQ) sequence as compared to the canonical DEDD sequence of the *S. pombe* or human orthologs (Thore et al., 2003; Jonstrup et al., 2007; Horiuchi et al., 2009). In the Ccr4-Not complex, Caf1 binds Not1₇₅₄₋₁₀₀₀ at the opposite side of the molecule with respect to the active-site pocket (Figure 2A).

The interaction between Caf1 and Not1₇₅₄₋₁₀₀₀ is mediated by conserved hydrophobic residues, in particular by Met290, Met296, and Trp333 in Caf1 and Pro897, Phe938, Pro943, and Trp944 in Not1 (Figure 3B). Overall, the Not1₇₅₄₋₁₀₀₀-Caf1 interface buries only 8.4% of the total accessible solvent area (230 Å² for Caf1 and 420 Å² for Not1), leaving the large conserved patch on the concave surface of Not1₇₅₄₋₁₀₀₀ exposed to solvent and accessible (Figure 2B, right panel). The active-site

pocket of Caf1 is also solvent exposed in the ternary complex, since the binding of the Ccr4 LRR domain occurs at an adjacent and nonoverlapping surface (Figure 2B, right panel, and Figure S3C).

The Interaction of Caf1 with the LRR Domain of Ccr4 Forms a Contiguous Hydrophobic Core

The LRR domain of Ccr4 contains five leucine-rich repeats (Figure 2A, right panel). The repeats assemble side by side to form a crescent-shape structure, with the outer convex surface formed by α helices and the concave surface formed by parallel β strands. The concave surface of LRR domains is often the site of intermolecular interactions (Bella et al., 2008). In yeast Ccr4, the concave surface of the LRR domain interacts intramolecularly with the helices of the N-terminal region (Figure 2A). The 70 amino acid linker (residues 262–331) that connects the N-terminal region to the LRR domain is disordered. The ordered part of the N-terminal region buries the hydrophobic β sheet surface of the LRR domain (lined by Tyr363, Val384, Tyr407, Tyr409, Phe411, and Phe430). The structural analysis thus suggests that the N-terminal region and the LRR domain in yeast Ccr4 form a single structural unit. In vertebrates, Ccr4 lacks the N-terminal region (Figure S2), and the concave surface of the LRR domain is instead polar in nature.

The LRR domain of Ccr4 includes two helices that shield the leucine residues of the fifth repeat and lead directly into the

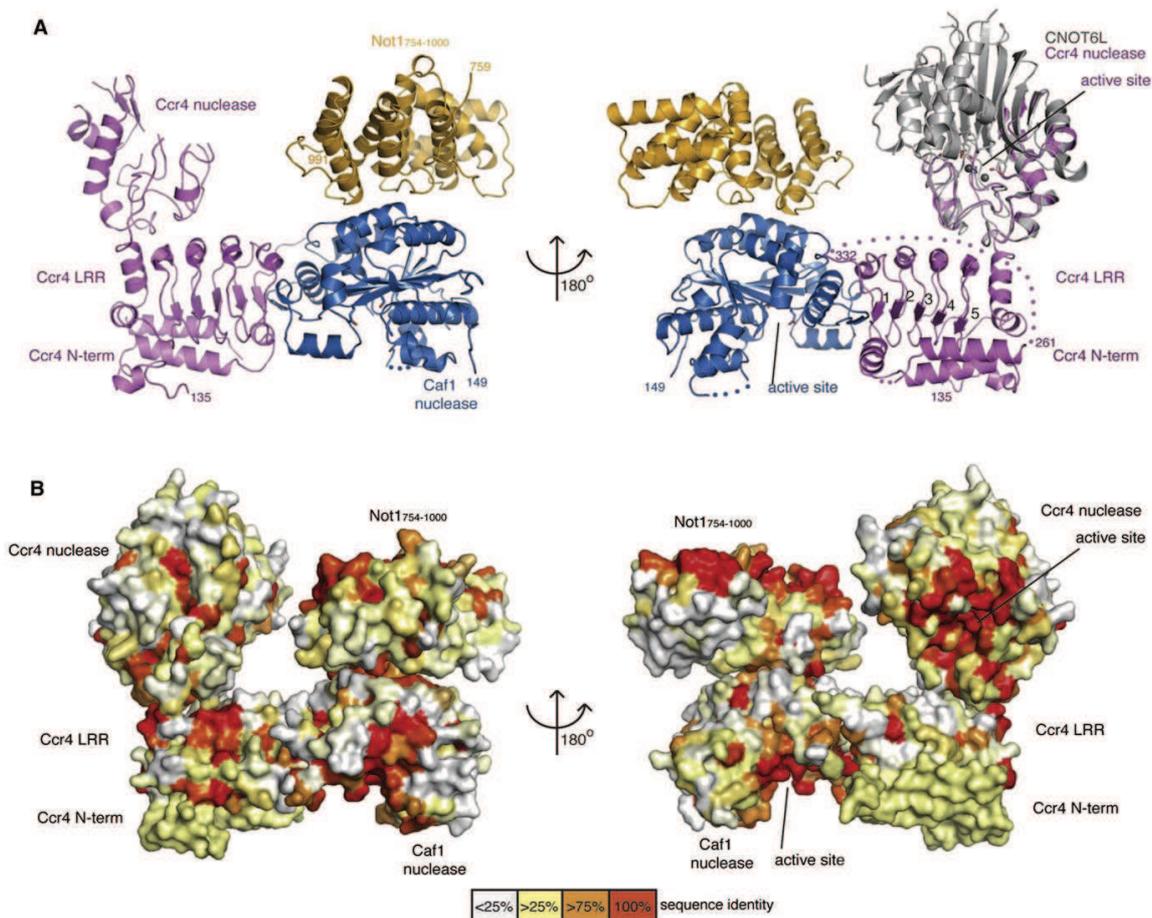


Figure 2. Structure of the Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 Complex

(A) The complex is shown in two different orientations related by a 180° rotation around a vertical axis. Not1 is in orange, Caf1 in blue, and Ccr4 in pink. On the right, the complex is shown superposed to the structure of the nuclease domain of the human Ccr4 ortholog CNOT6L (in gray, PDB 3NGQ). The five LRR motifs and the N and C termini are labeled. Disordered loops are indicated with a dotted line. Residues at the active sites of Caf1 and Ccr4 are shown in stick representation. All structure figures in this manuscript were generated with Pymol (<http://www.pymol.org/>).

(B) Surface representation of the complex in similar orientations as in (A) and including the superposed CNOT6L nuclease domain. The surface is colored according to sequence conservation, ranging from white (variable residues) to dark orange (invariant residues). The conservation was mapped with the program Consurf (Ashkenazy et al., 2010) based on the alignment in Figure S2. See also Figure S2.

Ccr4 nuclease domain (Figure 2A). It however lacks an analogous flanking region at the first repeat. The hydrophobic core of the first repeat is shielded intermolecularly by Caf1. Caf1 interacts with Ccr4 via a surface formed by a long loop (residues 193–206) and an α helix (residues 207–219) (Figure 3C). This region of Caf1 undergoes a localized conformational change as compared to the unbound Caf1 structure (Figure S3C). Caf1 residues 197–199 form a β strand that effectively extends the Ccr4 β sheet. The hydrophobic residues of the first LRR repeat (Leu339 and Leu341) contact nonpolar side chains of Caf1 (Ala215 and Phe219) (Figure 3C). The residues involved in the Caf1-Ccr4

interaction are conserved from fungi to animals (Figure S2), and the interaction surface is consistent with previous mutagenesis studies (Clark et al., 2004; Ohn et al., 2007).

The Nuclease Domain of Ccr4 Is Adjacent to the LRR Domain

Structural studies had shown that the catalytic domain of human CNOT6L, the ortholog of yeast Ccr4, belongs to the endonuclease-exonuclease-phosphatase (EEP) protein family (Wang et al., 2010). The fold is organized into two β sheets that face each other and are sandwiched between two outer layers of

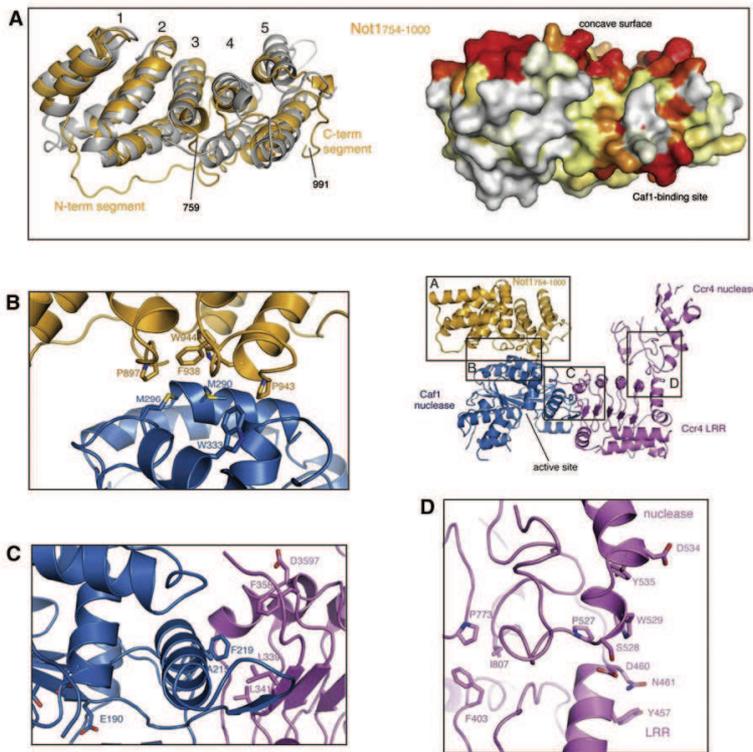


Figure 3. The Interactions in Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 Are Conserved

The figure shows four enlarged regions of the complex discussed in the text.

(A) On the left, the crystal structure of the Not1₇₅₄₋₁₀₀₀ domain is shown in orange superposed to a MIF4G domain of UPF2 (Kadlec et al., 2004, in gray). The five HEAT repeats are labeled and the N- and C-terminal segments are indicated. On the right, Not1₇₅₄₋₁₀₀₀ is shown in the same orientation as a surface representation colored according to sequence conservation, ranging from white (variable) to dark orange (conserved), as described for Figure 2B.

(B) Zoom-in showing a subset of residues in stick representation at the Not1-Caf1 interface. Residues mutated in the functional assays in Figure 4 are indicated.

(C) Zoom-in showing a subset of residues in stick representation at the Caf1-Ccr4 interface. Residues mutated in the functional assays in Figure 4 are labeled. Also shown are residues reported in previous mutagenesis studies (Asp357, Phe358) (Clark et al., 2004) and residues at the Caf1 active site (Glu190, etc.).

(D) Zoom-in of the LRR-nuclease interaction of Ccr4. Hydrophobic residues at the interface are indicated, as well as conserved residues exposed on the surface.

See also Figure S3.

helices, forming a prominent active-site pocket (Figure 2, right panels). In two molecules of the asymmetric unit of the Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 crystals, the helices in one half of the Ccr4 nuclease interact with the LRR domain, while the other half is disordered. The ordered part of the nuclease domain in the yeast structure has the same conformation as the corresponding part of human CNOT6L (rmsd of 1.1 Å upon superposition of 105 α carbon atoms) (Figure 2A, right panel).

The interaction between the LRR and nuclease domains of Ccr4 buries a surface of about 400 Å². The interface features hydrophobic contacts (between Phe403 from the β sheet surface of the LRR domain and Pro773 from the nuclease domain) and polar contacts (between Asp460 from the C-terminal helix of the LRR and Ser528 from the nuclease) (Figure 3D). Although the intramolecular interactions in Ccr4 appear to be weak and stabilized by lattice contacts, they are generally conserved (Figure S2) and create a continuous patch of evolutionary conserved solvent-exposed residues (Tyr457 and Asn461 from the LRR domain and Trp529, Asp534, and Tyr535 from the nuclease domain) (Figures 3D and S2). Moreover, an intramolecular contact between the nuclease domain and the LRR region had also been suggested by yeast two-hybrid studies (Clark et al., 2004). The relative arrangement of the two domains in Ccr4 results in an L-shape conformation, reminiscent of the central region in the L-shaped negative-stain EM structure of the Ccr4-Not complex (Nasertorabi et al., 2011).

Structure-Based Mutations Specifically Disrupt the Not1-Caf1 and the Caf1-Ccr4 Interfaces

Mutations based on the structural analysis were designed to specifically disrupt the interactions between Not1, Caf1, and Ccr4. To impair the Not1₇₅₄₋₁₀₀₀-Caf1 interface, we engineered a M290K, M296K double substitution in Caf1 (Caf1-290/296 mutant) and a F938E, P943Y substitution in Not1 (Not1-938/943 mutant). In coimmunoprecipitation experiments using tagged Caf1, Ccr4, and Not1 proteins, the Caf1-290/296 mutant disrupted the interaction with Not1 and did not affect the Caf1-Ccr4 interaction (Figure 4A). Conversely, the Not1-938/943 mutant impaired the interaction with Caf1 without interfering with Not1 stability or the assembly of the Caf1-Ccr4 heterodimer (Figure 4A). Similar results were obtained using yeast two-hybrid assays (Figure S4A).

To disrupt the Caf1-Ccr4 interface, we engineered an A215E, F219E double substitution in Caf1 (Caf1-215/219 mutant) and a L339E, L341E in Ccr4 (Ccr4-339/341 mutant). The Caf1-215/219 mutant abolished the interaction with Ccr4, as shown by coimmunoprecipitation experiments (Figure 4B) and in yeast two-hybrid assays (Figure S4A), and slightly reduced interaction with Not1 (Figure 4B). Conversely, the Ccr4-339-341 mutant abolished the interaction with Caf1 but did not impact on the Caf1-Not1 interaction (Figure 4C and S4A). Importantly, the mutant proteins accumulated at normal level in all cases.

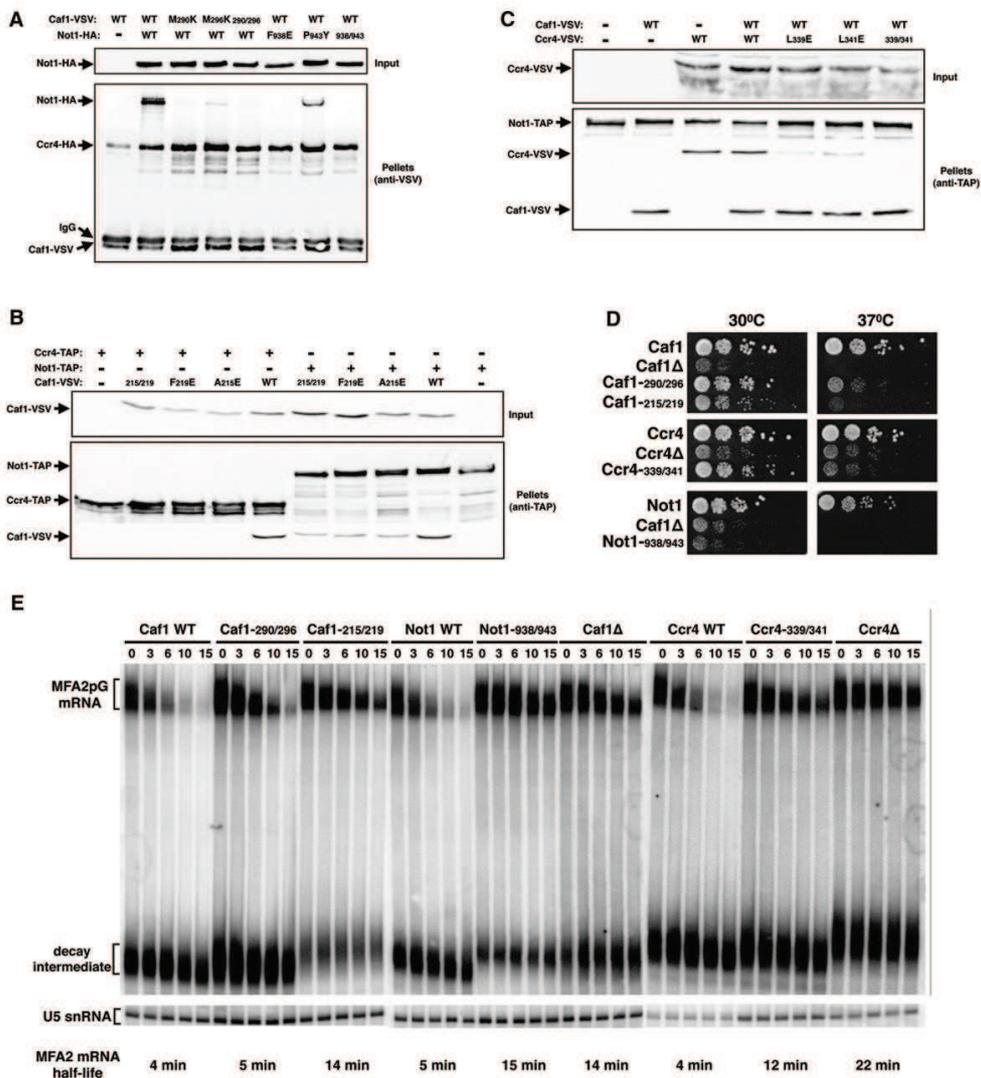


Figure 4. The Interactions in Not1⁷⁵⁴⁻¹⁰⁰⁰-Caf1-Ccr4 Are Essential for Normal Growth and mRNA Decay

(A–C) Coimmunoprecipitation was used to assay Not1–Caf1–Ccr4 interactions.

(A) VSV-tagged wild-type or mutant Caf1 present in strains expressing wild-type HA-tagged Not1 and HA-tagged Ccr4 were immunoprecipitated with anti-VSV antibodies. Parallel anti-VSV immunoprecipitations were performed from strains expressing VSV-tagged wild-type Caf1, HA-tagged mutant Not1, and wild-type HA-tagged Ccr4. Proteins present in the pellet fraction were detected by western blotting with anti-HA (Ccr4 and Not1) and anti-VSV (Caf1) antibodies.

(B) TAP-tagged wild-type Ccr4 (left lanes) or TAP-tagged Not1 (right lanes) present in strains expressing either wild-type or mutant VSV-tagged Caf1 were immunoprecipitated on IgG beads. Proteins present in the pellet fraction were detected by western blotting with anti-VSV antibodies (Caf1) and peroxidase-anti-peroxidase complex (Ccr4 or Not1).

(C) TAP-tagged wild-type Not1 present in strains expressing either VSV-tagged Ccr4 and/or VSV-tagged Caf1 was immunoprecipitated. Either wild-type or the indicated mutant forms of Ccr4 were used for this experiment. Proteins present in the pellet fraction were detected by western blotting with anti-VSV antibodies (Ccr4 and Caf1) and peroxidase-anti-peroxidase complex (Not1).

(D) Growth phenotypes resulting from mutations of interface residues. Plasmids encoding the Caf1, Ccr4, and Not1 mutants were introduced in strains deleted of the cognate wild-type gene, either directly (Caf1 and Ccr4) or by plasmid shuffling (Not1). Growth of the resulting strain was assayed on YPDA media at 30°C and 37°C.

(E) The plasmid encoding the MFA2pG reporter was introduced in the wild-type yeast strain, isogenic deletion strains lacking Ccr4 or Caf1 and point mutant of interface residues. The decay of the MFA2 mRNA was monitored in chase experiments. The MFA2 mRNA, decay intermediates and U5 small nuclear RNA were detected by northern blotting after fractionation in denaturing polyacrylamide gels.

See also Figure S4.

The Recruitment of Caf1 and Ccr4 to Not1 Is Essential for Their Biological Function

We used the structure-based mutants to test the physiological consequences of disrupting the Not1-Caf1 and the Caf1-Ccr4 interactions *in vivo*. To dissociate Caf1 from Not1, we introduced the Caf1-290/296 mutant in a strain carrying a chromosomal deletion of Caf1. Assaying the growth of the resulting strain demonstrated that preventing the association of Caf1 with Not1 results in a poor growth phenotype at high temperature (Figure 4D). In the latter conditions, the Caf1-290/296 mutant barely differed from the strain lacking Caf1.

In a complementary analysis, we introduced the Not1-938/943 mutant in the *not1Δ* haploid strain rescued by a *URA3*-marked *NOT1* plasmid. The Not1-938/943 mutant was able to rescue the lethal phenotype conferred by Not1 inactivation, but the resulting strain grew at reduced rate at normal temperatures and very poorly at high temperatures (Figure 4D). The growth properties of the Not1-938/943 mutant strain were similar to those of an otherwise isogenic Caf1Δ strain (Figure 4D). The more severe phenotype observed with the Not1 mutant may reflect differences in residual interaction and/or the possibility that the mutated region of Not1 could be involved in additional functions. Nevertheless, because these mutations did not alter the level of Caf1, Ccr4, and Not1 or the formation of an active Ccr4-Caf1 dimer, these observations demonstrate that the association of Caf1-Ccr4 with Not1 is required for their biological activity. Interestingly, fusing Caf1 at the N terminus of the Not1-938/943 mutant partly rescued the slow growth phenotype of the Not1-938/943 protein, demonstrating that its inability to recruit Caf1 (and indirectly Ccr4) explains the poor activity of this factor (Figure S4B).

To dissociate Ccr4 from Caf1, we introduced the Caf1-215/219 and the Ccr4-339/341 mutants in strains devoid of Caf1 and Ccr4, respectively (Figure 4D). In both cases, the mutant strains grew at slightly reduced rate, a phenotype that was exacerbated at high temperature. We conclude that the recruitment of Ccr4 to the Ccr4-Not complex is essential for its biological activity.

Detachment of Caf1 and Ccr4 from Not1 Impairs Deadenylation and mRNA Decay

Next, we monitored the effect of the mutations on mRNA half-life. We introduced in the cognate mutants a reporter plasmid encoding the MFA2 mRNA under the control of a galactose-inducible promoter (Decker and Parker, 1993). Yeast cells were grown in galactose containing medium and transferred to glucose media to repress the reporter (time 0). Samples were withdrawn from the culture at various time points, to extract RNA and monitor the decay of the MFA2 reporter by northern blotting. The reporter mRNA was specifically detected with a probe hybridizing to an oligo(G) tract introduced in the reporter mRNA 3' untranslated region (UTR) that also allowed the visualization of mRNA decay fragments by blocking the main 5'-3' exonuclease, Xrn1.

Decay of the MFA2 reporter occurred with normal kinetic in control strains expressing wild-type Caf1, Ccr4 and Not1 (half-life [$t_{1/2}$] of 4–5 min) but was strongly impaired in strains lacking Caf1 or Ccr4 ($t_{1/2}$ of 14 and 22 min, respectively) (Figure 4E). In the absence of Caf1 and Ccr4, longer length intermediates that

carried extended poly(A) tracts compared to their wild-type counterparts were detected (Figure 4E), as reported (Daugeron et al., 2001; Tucker et al., 2002). These intermediates extend from a stem loop in the reporter 3' UTR to the mRNA 3' end and originate from 5'-3' exonucleolytic degradation of the mRNA body after decapping. Their increased length indicates that decapping proceeds despite the fact that poly(A) tails have not been shortened to the same length as in wild-type strains.

Upon assaying the disruption of the Not1-Caf1 interface (Figure 4E), we found that the reporter was only mildly affected in the Caf1-290/296 mutant (normal half life, moderately extended intermediates), while its decay was more severely impaired in the Not1-938/943 strain ($t_{1/2}$ 15min, highly extended intermediates). When assaying the disruption of the Caf1-Ccr4 interface, both the Caf1-215/219 and the Ccr4-339/341 mutant strains resulted in the stabilization of the reporter ($t_{1/2}$ of 14 and 12 min, respectively) and in extended decay intermediates (Figure 4E). These phenotypes were highly reproducible when the independent biological replicates were performed. Altogether, these observations demonstrate that the incorporation of Caf1 and Ccr4 in the Ccr4-Not complex (via the interaction with Not1) is essential for efficient deadenylation *in vivo*, as qualitatively evidenced by altered reporter mRNA half-life and the increased size of the decay intermediate. Quantitative differences observed between the mutants (i.e., the more severe phenotype of the Not1 mutation) may originate from different residual levels of interaction. It is also possible that the incorporation of Caf1-Ccr4 within the complex might differentially affect the deadenylation of different mRNAs. Finally, the Not1 mutant might also affect other properties or activities of the Ccr4-Not complex in addition to the recruitment of the deadenylases.

The Not1₁₅₄₋₇₅₄ Domain Is an Extended Platform of HEAT Repeats

To obtain a more complete view of the N-terminal arm of Not1, we also determined the structure of its largest domain. We solved the crystal structure of Not1₁₅₄₋₇₅₃ using Au-based SAD phasing and refined it at 2.8 Å resolution to an *R*_{free} of 27.3% (see Table 1 for data collection and refinement statistics). The asymmetric unit of the crystals contains two independent copies of Not1₁₅₄₋₇₅₃, which have a r.m.s.d. of 0.62 Å over all α carbon atoms. The final model includes residues 193–745, with the exception of disordered loops at residues 211–232 and 268–270.

The Not1₁₅₄₋₇₅₃ structure is entirely built by antiparallel helices assembled side by side to form an elongated molecule with approximate dimensions of 120 Å × 25 Å × 25 Å (Figure 5A). Analysis of the structure identifies 13 helical hairpins related to HEAT repeats. HEAT repeats 1–9 have an overall parallel arrangement, with the exception of an unusually large rotation between HEAT 4 and 5. HEAT repeats 10–13 are also organized in a parallel fashion, but this region as a whole is rotated by about 90° with respect to HEAT 1–9 (Figure 5A). A search of the structural database using the Dali server (Holm and Rosenström, 2010) identified structural similarity between HEAT 10–13 (residues 571–746) and four HEAT repeats of MIF4G1 (r.m.s.d. of 3.9 Å over 115 Ca atoms after optimal superposition) (Figure S5).

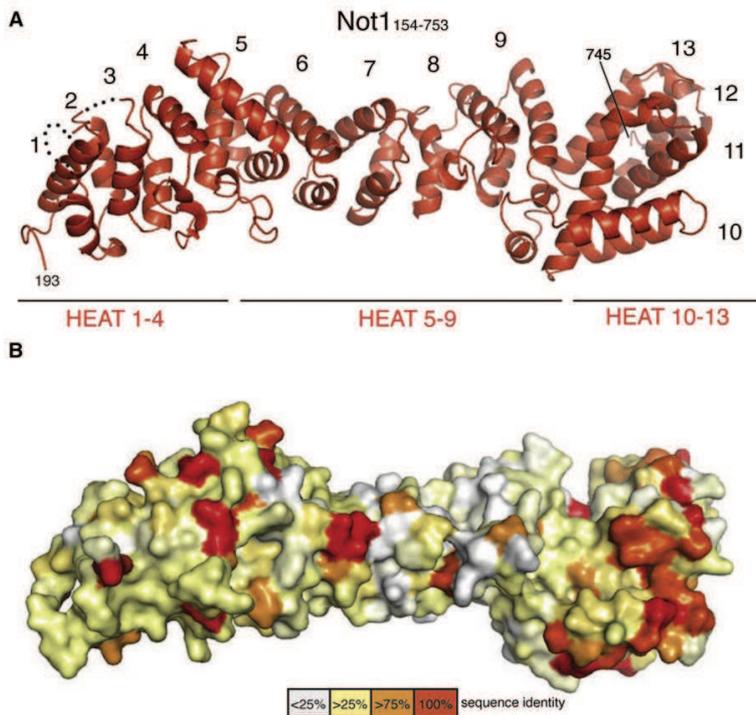


Figure 5. Not1₁₅₄₋₇₅₃ Is a HEAT-Rrepeat Platform

(A) Crystal structure of Not1₁₅₄₋₇₅₃ shown in a ribbon representation in red. The HEAT repeats and the ordered N and C termini are labeled. Disordered loops are indicated with a dotted line. (B) Not1₁₅₄₋₇₅₃ is shown in the same view as a surface representation colored according to sequence conservation as in Figure 2B. See also Figure S5.

HEAT 10–13 is the most conserved portion of the Not1₁₅₄₋₇₅₃ domain (Figures 5B and S2). In particular, residues from the A helices of HEAT 10–11 form an evolutionary conserved patch on the outer surface of this domain. This conserved surface of Not1₁₅₄₋₇₅₃ (Figure 5B) and the conserved surface of Not1₇₅₄₋₁₀₀₀ (Figure 2B, right panel) are prime targets for protein-protein interactions either within the core of the Ccr4-Not complex or for the recruitment of other factors.

Conclusions

The emerging picture from the current structural and functional information is that Not1 scaffolds the Ccr4-Not complex into two structural arms organized around the two halves of its 240 kDa polypeptide chain (Figure 6). In the N-terminal arm, the Not1₁₅₄₋₇₅₃ and Not1₇₅₄₋₁₀₀₀ domains have a HEAT-repeat organization with regions related to MIF4G. This architecture is reminiscent of other scaffolding proteins involved in mRNA metabolism such as eIF4G (Schütz et al., 2008), CBC (Mazza et al., 2001), and UPF2 (Kadlec et al., 2004). The N-terminal arm of Not1 ends with a MIF4G domain that forms a rather rigid platform for the interaction with the Caf1 nuclease. Caf1 is sandwiched between Not1 and the LRR domain of Ccr4, which in turn tethers the nuclease domain with a covalent linkage and intramolecular interactions. The interactions between Not1, Caf1, and Ccr4 are evolutionarily conserved and functionally essential *in vivo* in yeast.

In vitro, Caf1-Ccr4 are active deadenylases in the absence of Not1 (Thore et al., 2003; Clark et al., 2004; Jonstrup et al.,

2007). In the crystal structure, their active sites do not contact Not1₇₅₄₋₁₀₀₀ and are accessible. *In vivo*, however, longer mRNA decay intermediates accumulate when disrupting the Not1-Caf1-Ccr4 interactions. These observations argue that the recruitment of Caf1-Ccr4 in the Ccr4-Not complex (via Not1) is required for full activity in yeast. It is possible that the complex might target the deadenylases to specific substrates and/or that the additional proteins in the Ccr4-Not complex might contribute. Indeed, studies in human cells have shown that depletion of Not2, a protein bound at the C-terminal arm of Not1 (Bai et al., 1999), affects the stability and enzymatic activity of the complex (Ito et al., 2011).

Finally, the incorporation of Caf1-Ccr4 into the Not complex might also impact other processes in the cell, such as ubiquitination by Not4 or transcription.

The interpretation of the available data suggests that Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 connects the two structural arms (Figures 6 and S6). The resulting model has interesting implications. First, it shows that the two nucleases are positioned centrally within the Ccr4-Not complex. Second, it rationalizes how regulators like BTG/Tob factors can be recruited to the surface of the complex to thread RNA substrates to the nuclease active sites (Figure S6). While BTG/Tob proteins bind directly to Caf1 (Horiuchi et al., 2009), it is possible to envisage that other regulators might be recruited to the extended HEAT-repeat platform in the N-terminal arm or to the C-terminal Not module. Finally, the model suggests that the nuclease domain of Ccr4 is adjacent to the Not module (Figures 6 and S6), rationalizing how a crosstalk between the nucleases and the Not proteins might be achieved. Understanding the structure and function of the Not module and the recruitment of regulators are open questions for future studies.

EXPERIMENTAL PROCEDURES

Protein Purification and Structure Determination

Not1 and Caf1 proteins were produced recombinantly in *E. coli*, while Ccr4 was obtained from insect-cell expression (Supplemental Experimental Procedures). The Not1₇₅₄₋₁₀₀₀-Caf1 and Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 complexes were reconstituted by mixing the purified components and by subsequent purification by size exclusion chromatography (Sephadex 200, GE

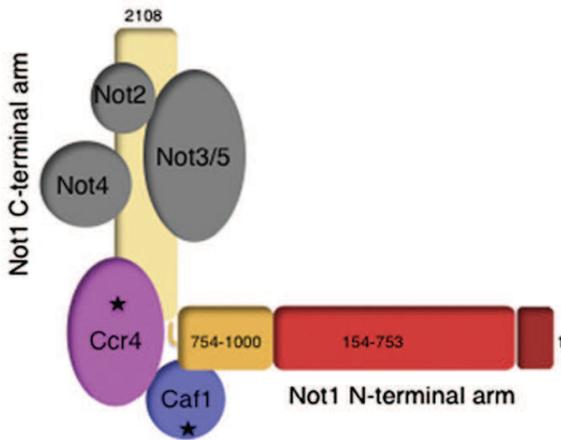


Figure 6. Structural Organization of the Yeast Ccr4-Not Complex

The scheme recapitulates the information from the crystal structures of Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 and Not1₁₅₃₋₇₅₃ reported in this manuscript, combined with structural information from an EM study of the Ccr4-Not complex (Nasertorabi et al., 2011) (Figure S6) and previous interaction studies (reviewed in Collart and Timmers, 2004). Not1 is shown with the two structural arms colored from red (N terminus, residue 1) to yellow (C terminus, residue 2,108). Caf1 and Ccr4 are shown in blue and pink, with stars indicating their active sites. The Not2, Not3, Not4, and Not5 proteins that are part of the C-terminal “Not module” are shown in gray. Not shown here are Caf40 and Caf130, for which the interacting regions within the complex have yet to be precisely identified. See also Figure S6.

Healthcare). The samples were crystallized with the sitting-drop vapor diffusion technique at 18°C. Crystals were cryoprotected in mother liquor supplemented with 20% glycerol, mounted in nylon loops, and flash-frozen in liquid nitrogen for data collection at 100 K. Detailed procedures for crystallization and structure determination are in the [Supplemental Experimental Procedures](#). The data statistics and quality of the models are summarized in Table 1.

Yeast Strains, Plasmids, and Growth Conditions

Yeasts were grown in standard media at 30°C, except when otherwise indicated. Yeast strains are listed in Table S1. Standard cloning, mutagenesis, and transformation procedures were used. Plasmids are listed in Table S2.

Reporter RNA Analysis, Western Blots and Coimmunoprecipitation Assays

RNA chase experiments were essentially performed as described previously (Daugeron et al., 2001). Details are in the [Supplemental Experimental Procedures](#).

For coimmunoprecipitation assays, yeast total proteins were extracted from pellets prepared from 150 ml of yeast culture grown to OD₆₀₀ 1.5. Coimmunoprecipitations were performed with 20 μl IgG-sepharose (GE Healthcare) for TAP-tagged proteins, or using protein G-sepharose preincubated with antiVSV IgG (Roche) for VSV-tagged proteins. Details are in the [Supplemental Experimental Procedures](#). The input and precipitated fractions were analyzed by western blotting with antibodies described in the [Supplemental Experimental Procedures](#).

ACCESSION NUMBERS

The coordinates and the structure factors have been deposited in the Protein Data Bank with accession numbers 4b8c (Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4), 4b8a (Not1₇₅₄₋₁₀₀₀-Caf1), 4b89 (Not1₇₅₄₋₁₀₀₀), and 4b8b (Not1₁₅₄₋₇₅₃).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2012.08.014>.

ACKNOWLEDGMENTS

We are grateful to E. Simon for defining stably expressing functional Ccr4-Caf1 truncations, to C. Carolis and D. Suck for the original expression vector of Ccr4-Caf1, to C. Gaudon, F. Wyers, F. Lacroute, V. Henriot, and C. Faux for plasmids and/or antibodies and the IGBMC for assistance. We gratefully acknowledge the Crystallization Facility and the Core Facilities at the Max Planck Institute of Biochemistry and A. Pauluhn, V. Olieric, and the staff of the PX beamlines at the SLS (Villigen, Zurich, Switzerland) for assistance during data collection. We also thank E. Lorentzen and F. Mauxion for useful discussions as well as members of our labs for critical reading of the manuscript. This study was supported by the Max Planck Gesellschaft, the European Research Council (ERC Advanced Investigator Grant 294371) and the Deutsche Forschungsgemeinschaft (DFG Sonderforschungsbereich SFB646, GRK1721), and the Gottfried Wilhelm Leibniz Program) to E.C. B.S. acknowledges support from the CERBM-IGBMC, the CNRS, the Ligue Contre le Cancer (Equipe Labellisée 2011) and the EU “3D-Repertoire” program (LSHG-CT-2005-512028).

Received: June 25, 2012

Revised: August 6, 2012

Accepted: August 16, 2012

Published online: September 6, 2012

REFERENCES

- Albert, T.K., Lemaire, M., van Berkum, N.L., Gentz, R., Collart, M.A., and Timmers, H.T. (2000). Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. *Nucleic Acids Res.* 28, 809–817.
- Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* 309, 1–18.
- Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* 38 (Web Server issue), W529–W533.
- Bai, Y., Salvatore, C., Chiang, Y.C., Collart, M.A., Liu, H.Y., and Denis, C.L. (1999). The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Mol. Cell Biol.* 19, 6642–6651.
- Bella, J., Hindle, K.L., McEwan, P.A., and Lovell, S.C. (2008). The leucine-rich repeat structure. *Cell. Mol. Life Sci.* 65, 2307–2333.
- Boeck, R., Tarun, S., Jr., Rieger, M., Deardorff, J.A., Müller-Auer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. *J. Biol. Chem.* 271, 432–438.
- Braun, J.E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* 44, 120–133.
- Brown, C.E., and Sachs, A.B. (1998). Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation. *Mol. Cell Biol.* 18, 6548–6559.
- Brown, C.E., Tarun, S.Z., Jr., Boeck, R., and Sachs, A.B. (1996). PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 16, 5744–5753.
- Cao, D., and Parker, R. (2001). Computational modeling of eukaryotic mRNA turnover. *RNA* 7, 1192–1212.
- Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* 18, 1218–1226.

- Chen, J., Rappsilber, J., Chiang, Y.C., Russell, P., Mann, M., and Denis, C.L. (2001). Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. *J. Mol. Biol.* **314**, 683–694.
- Chen, J., Chiang, Y.-C., and Denis, C.L. (2002). CCR4, a 3'-5' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. *EMBO J.* **21**, 1414–1426.
- Clark, L.B., Viswanathan, P., Quigley, G., Chiang, Y.-C., McMahon, J.S., Yao, G., Chen, J., Nelsbach, A., and Denis, C.L. (2004). Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. *J. Biol. Chem.* **279**, 13616–13623.
- Collart, M.A., and Timmers, H.T. (2004). The eukaryotic Ccr4-not complex: a regulatory platform integrating mRNA metabolism with cellular signaling pathways? *Prog. Nucleic Acid Res. Mol. Biol.* **77**, 289–322.
- Daugeron, M.C., Mauxion, F., and Séraphin, B. (2001). The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res.* **29**, 2448–2455.
- de Moor, C.H., Meijer, H., and Lissenden, S. (2005). Mechanisms of translational control by the 3' UTR in development and differentiation. *Semin. Cell Dev. Biol.* **16**, 49–58.
- Decker, C.J., and Parker, R. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes Dev.* **7**, 1632–1643.
- Deluen, C., James, N., Maillet, L., Molinete, M., Theiler, G., Lemaire, M., Paquet, N., and Collart, M.A. (2002). The Ccr4-not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions. *Mol. Cell. Biol.* **22**, 6735–6749.
- Denis, C.L., Chiang, Y.C., Cui, Y., and Chen, J. (2001). Genetic evidence supports a role for the yeast CCR4-NOT complex in transcriptional elongation. *Genetics* **158**, 627–634.
- Draper, M.P., Salvatore, C., and Denis, C.L. (1995). Identification of a mouse protein whose homolog in *Saccharomyces cerevisiae* is a component of the CCR4 transcriptional regulatory complex. *Mol. Cell. Biol.* **15**, 3487–3495.
- Dupressoir, A., Morel, A.P., Barbot, W., Loireau, M.P., Corbo, L., and Heidmann, T. (2001). Identification of four families of yCCR4- and Mg2+-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. *BMC Genomics* **2**, 9.
- Eckmann, C.R., Rammelt, C., and Wahle, E. (2011). Control of poly(A) tail length. *Wiley Interdiscip Rev RNA* **2**, 348–361.
- Edmonds, M., Vaughan, M.H., Jr., and Nakazato, H. (1971). Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship. *Proc. Natl. Acad. Sci. USA* **68**, 1336–1340.
- Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., and Sonenberg, N. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217.
- Gameau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* **8**, 113–126.
- Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. *Nat. Rev. Mol. Cell Biol.* **9**, 337–344.
- Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38** (Web Server issue), W545–W549.
- Horiuchi, M., Takeuchi, K., Noda, N., Muroya, N., Suzuki, T., Nakamura, T., Kawamura-Tsuzuku, J., Takahashi, K., Yamamoto, T., and Inagaki, F. (2009). Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. *J. Biol. Chem.* **284**, 13244–13255.
- Ito, K., Inoue, T., Yokoyama, K., Morita, M., Suzuki, T., and Yamamoto, T. (2011). CNOT2 depletion disrupts and inhibits the CCR4-NOT deadenylase complex and induces apoptotic cell death. *Genes Cells* **16**, 368–379.
- Jonstrup, A.T., Andersen, K.R., Van, L.B., and Brodersen, D.E. (2007). The 1.4-Å crystal structure of the *S. pombe* Pop2p deadenylase subunit unveils the configuration of an active enzyme. *Nucleic Acids Res.* **35**, 3153–3164.
- Kadlec, J., Izaurralde, E., and Cusack, S. (2004). The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat. Struct. Mol. Biol.* **11**, 330–337.
- Kruk, J.A., Dutta, A., Fu, J., Gilmour, D.S., and Reese, J.C. (2011). The multifunctional Ccr4-Not complex directly promotes transcription elongation. *Genes Dev.* **25**, 581–593.
- Lau, N.C., Kolkman, A., van Schaik, F.M., Mulder, K.W., Pijnappel, W.W., Heck, A.J., and Timmers, H.T. (2009). Human Ccr4-Not complexes contain variable deadenylase subunits. *Biochem. J.* **422**, 443–453.
- Liu, H.Y., Badarinarayana, V., Audino, D.C., Rappsilber, J., Mann, M., and Denis, C.L. (1998). The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *EMBO J.* **17**, 1096–1106.
- Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O., and Collart, M.A. (2000). The essential function of Not1 lies within the Ccr4-Not complex. *J. Mol. Biol.* **303**, 131–143.
- Malvar, T., Biron, R.W., Kaback, D.B., and Denis, C.L. (1992). The CCR4 protein from *Saccharomyces cerevisiae* contains a leucine-rich repeat region which is required for its control of ADH2 gene expression. *Genetics* **132**, 951–962.
- Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4**, 223.
- Mauxion, F., Faux, C., and Séraphin, B. (2008). The BTG2 protein is a general activator of mRNA deadenylation. *EMBO J.* **27**, 1039–1048.
- Mazza, C., Ohno, M., Segref, A., Mattaj, I.W., and Cusack, S. (2001). Crystal structure of the human nuclear cap binding complex. *Mol. Cell* **8**, 383–396.
- Nasertorabi, F., Batisse, C., Diepholz, M., Suck, D., and Böttcher, B. (2011). Insights into the structure of the CCR4-NOT complex by electron microscopy. *FEBS Lett.* **585**, 2182–2186.
- Ohn, T., Chiang, Y.-C., Lee, D.J., Yao, G., Zhang, C., and Denis, C.L. (2007). CAF1 plays an important role in mRNA deadenylation separate from its contact to CCR4. *Nucleic Acids Res.* **35**, 3002–3015.
- Sandler, H., Kreth, J., Timmers, H.T., and Stoecklin, G. (2011). Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res.* **39**, 4373–4386.
- Schütz, P., Bumann, M., Oberholzer, A.E., Bieniossek, C., Trachsel, H., Altmann, M., and Baumann, U. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **105**, 9564–9569.
- Schwede, A., Ellis, L., Luther, J., Carrington, M., Stoecklin, G., and Clayton, C. (2008). A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res.* **36**, 3374–3388.
- Simón, E., and Séraphin, B. (2007). A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay. *Nucleic Acids Res.* **35**, 6017–6028.
- Sonenberg, N., and Dever, T.E. (2003). Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.* **13**, 56–63.
- Sun, M., Schwab, B., Schulz, D., Pirkel, N., Eitzold, S., Larivière, L., Maier, K.C., Seiz, M., Tresch, A., and Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res.* **22**, 1350–1359.
- Tarun, S.Z., Jr., Wells, S.E., Deardorff, J.A., and Sachs, A.B. (1997). Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proc. Natl. Acad. Sci. USA* **94**, 9046–9051.
- Temme, C., Zaessinger, S., Meyer, S., Simonelig, M., and Wahle, E. (2004). A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in *Drosophila*. *EMBO J.* **23**, 2862–2871.

- Thore, S., Mauxion, F., Séraphin, B., and Suck, D. (2003). X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylation complex. *EMBO Rep.* *4*, 1150–1155.
- Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* *104*, 377–386.
- Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrud, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. *EMBO J.* *21*, 1427–1436.
- Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M., et al. (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *EMBO J.* *29*, 2566–2576.
- Yamashita, A., Chang, T.-C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.-Y.A., and Shyu, A.-B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* *12*, 1054–1063.
- Zhang, X., Virtanen, A., and Kleiman, F.E. (2010). To polyadenylate or to deadenylate: that is the question. *Cell Cycle* *9*, 4437–4449.

Supplemental Information

Architecture of the Nuclease Module

of the Yeast Ccr4-Not Complex:

the Not1-Caf1-Ccr4 Interaction

Jérôme Basquin, Vladimir V. Roudko, Michaela Rode, Claire Basquin, Bertrand Séraphin, and Elena Conti

Supplemental Experimental Procedures

Protein purification

Not1₁₋₁₀₀₀, Not1₁₋₁₅₃, Not1₁₅₄₋₇₅₃, Not1₇₅₄₋₁₀₀₀ and Caf1 were subcloned with a cleavable SUMO-His tag at the N-terminus. Each protein was individually expressed in *E. coli* BL21(DE3) cells. The pellets were resuspended in 20 mM Tris pH 7.5, 250 mM NaCl, 20 mM imidazole and lysed by sonication. The proteins were purified with a Ni-nitriloacetate (NTA) affinity step followed by tag cleavage with the SUMO-protease Senp2 (in a 1:40 protease:protein ratio) and by heparin Sepharose chromatography (GE Healthcare). The final purification step by size-exclusion chromatography was carried out using a Sephadex 75 (for Not1₁₋₁₅₃, Not1₇₅₄₋₁₀₀₀ and Caf1) or a Sephadex 200 (for Not1₁₋₁₀₀₀ and Not1₁₅₄₋₇₅₃) column (GE Healthcare).

For Ccr4 insect-cell expression, a derivative of a pFastBac DUAL vector (Invitrogen) inserted with His-tagged Ccr4 (residues 110-837) and Caf1 (residues 146-433) at its two different promoters (kind gift of Carlo Carolis) was used to generate bacmids and viruses. The proteins were co-expressed in SF9 cells according to standard procedures. Cells were lysed after resuspending the pellet in a low-salt buffer (20 mM Tris pH 7.5, 10 mM NaCl, 2mM MgCl₂). The clear lysate was supplemented with 250 mM NaCl and 20 mM imidazole, and purified by consecutive Ni-NTA Sepharose, heparin Sepharose and size-exclusion chromatography (Sephadex 200, GE Healthcare). The Not1₇₅₄₋₁₀₀₀-Caf1 and Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 complexes were reconstituted by mixing the purified components described above and by subsequent purification by size exclusion chromatography (Sephadex 200, GE Healthcare). In all cases, the final purification steps by size-exclusion chromatography were carried out in 20 mM Tris pH 7.5, 250 mM NaCl and 1 mM DTT.

Crystal structures determination

Not1₁₅₄₋₇₅₃ was concentrated to 15 mg/ml and crystallized in 2.8 M NaCl and 50 mM Tris pH 7.2. Not1₁₅₄₋₇₅₃ crystals grew to about 0.3 x 0.3 x 0.3 mm in size, but with severe macroscopic defects. Phases were obtained from a crystal derivatized by soaking with 6 mM KAu(CN)₂ for 2 hours and then manually dissected into small fragments. Using one of the fragments, we collected a SAD data set to 2.8 Å resolution with synchrotron radiation tuned at the peak wavelength of the gold absorption edge (Table 1). After substructure solution and phasing with Phenix AutoSol (Adams et al., 2010), an initial model was built automatically with BUCCANEER (Cowtan, 2006) and extended manually in COOT (Emsley et al.,

2010). Structure refinement was carried out using phenix.refine with TLS refinement (Adams et al., 2010).

Not1₇₅₄₋₁₀₀₀ was concentrated to 20 mg/ml and crystallized in 20% (w/v) PEG3350, 0.5 M NaI and 50 mM Tris pH 7.5. We obtained phases from a data set collected to 2.6 Å resolution using in-house CuK α radiation, exploiting the anomalous signal of iodine. Substructure solution and phasing were carried out with Phenix AutoSol. The model was built with COOT and refined using phenix.refine against a native data set to 1.4 Å resolution collected at the Swiss Light Source (SLS) synchrotron (Table 1).

The Not1₇₅₄₋₁₀₀₀-Caf1 binary complex was concentrated to 20 mg/ml and crystallized in 8% (w/v) PEG 8000, 0.2 M Na malonate, 50 mM Tris pH 8.0 and 0.1% low-melting agarose. The crystals diffracted to 2.3 Å resolution at SLS. The structure was solved with the program PHASER (McCoy et al., 2007) implemented in Phenix AutoMR using as search models the Not1₇₅₄₋₁₀₀₀ coordinates (see above) and the Caf1 coordinates (PDB 1UOC, (Thore et al., 2003)).

The Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 ternary complex was concentrated to 15 mg/ml and crystallized in 8% PEG 3350, 75 mM ammonium acetate and 10% glycerol (w/v). Iterative rounds of seeding were used to improve the size and morphology of the crystals, which were then screened for diffraction at SLS. The best crystal diffracted to 3.3 Å resolution. Initial phases were obtained by molecular replacement with Phenix AutoMR using the coordinates of Not1₇₅₄₋₁₀₀₀-Caf1 as search model (see above). We subsequently placed a model for the LRR domain of Ccr4 obtained with the program PHYRE (ref) and part of the model of the nuclease domain (residues 512-555, 557-568, 576-579, 600-607, 766-779, 797-820) aided by the crystal structure of the human orthologue (ref). The model was completed the model by cyclic rounds of model building (with COOT) and refinement (with Phenix).

Reporter RNA analysis

1mL aliquots of cells harboring the MFA2pG reporter (Decker and Parker, 1993) were collected at the indicated time points. Cell pellets were frozen in liquid nitrogen and stored at -80°C. Yeast total RNA was obtained by hot acid phenol extraction. Northern blot analyses were carried out after polyacrylamide gel electrophoresis. The reporter RNA was detected using a radiolabeled poly(C) probe oRP121 (Boeck et al., 1998). Signals were measured and quantified using PhosphorImager (GE Healthcare). The mRNA half-lives were calculated using the best fit to an exponential decay. U5 snRNA, detected with a specific probe (EM19, 5'GGCCCACAGTTCTTGATGTTG 3'), was used as loading control.

Co-immunoprecipitation assays

Yeast total proteins were extracted from pellets prepared from 150 ml of yeast culture grown to OD₆₀₀ 1.5, that were stored at -80°C. 1mL of glass beads, 1.5 mL of buffer A (150mM NaCl, 40 mM Tris-Cl pH 7.5, 0,1% NP-40, 12% glycerol, 1mM DTT, protease inhibitors (1 tablet of protease inhibitor cocktail (Roche), added per 100mL of buffer) and 100 μ L of 1M PMSF were added to each pellet in a Corex tube. After 3 rounds of vortexing at maximum speed for 30 seconds at 4°C, large cell debris were removed by low speed centrifugation. 1.4 mL of lysate fraction was transferred to a cold Eppendorf tube and spun at 14krpm for 20 min at 4°C. 1.2 mL of the supernatant was transferred to new tubes and spun again at 14krpm for 10 min. 1ml of clarified lysate was flash frozen in liquid nitrogen and stored at -80°C.

Co-immunoprecipitations were performed with 20 μ l of IgG-sepharose 6 Fast Flow (GE Healthcare) for TAP-tagged proteins, or using protein G Sepharose, preincubated

with antiVSV IgG (Roche) for VSV-tagged proteins, in IPP150 buffer (50mM Tris-Cl pH8.0, 150mM NaCl, 0.1% NP40). Beads were rotated with 4mg of total extract for 2 hours at 4°C. After spinning, the flow-through fractions were kept and beads were washed 3 times with IPP150 at 4°C. Bound proteins were eluted with IPP150 + 1% SDS. Protein present in input and precipitated fractions were analyzed by western blotting following standard procedure. Monoclonal antibodies against VSV and HA tags were used (anti VSV-G antibody clone P5D4 from Roche; anti HA antibody clone HA.11 from Covance). Polyclonal antibodies against Stm1 (gift of F. Wyers) were used to ascertain equal loading. Peroxidase-anti-peroxidase complexes (Sigma) were used for detection TAP-tag containing proteins.

Yeast-two-hybrid assays

MAV203 competent yeast strain (MAT α , Invitrogen) and HF7c yeast strain (MAT α , (Feilotter et al., 1994)) were transformed with constructs in pDEST32 (DNA-binding domain, Invitrogen) and pDEST22 (Activation domain, Invitrogen) vector, respectively. After crossing, diploids were grown in 2ml of selective media and β -galactosidase activity was monitored using the NovaBright β -galactosidase Enzyme Reporter Gene Chemiluminescent Detection Kit (Invitrogen) following the manufacturer recommendations. Resulting values were normalized to cell densities monitored by OD₆₀₀ readings.

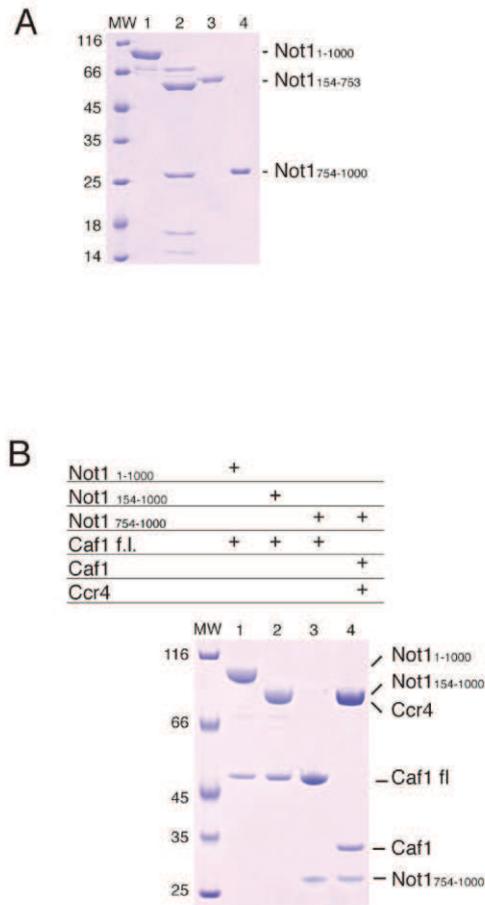


Figure S1. *In Vitro* Characterization of Not1 Domains, Related to Figure 1

- A)** Identification and purification of individual domains present in Not1₁₋₁₀₀₀. The 12.5 % (w/v) SDS gel (stained with Coomassie) shows in lanes 1, 3 and 4 the peak fractions from size exclusion chromatography. Lane 2 shows the proteolysed Not1₁₋₁₀₀₀ after treatment with trypsin. We note that since we could not obtain full-length Not1 in a recombinant form, we identified computationally the rough boundaries of about half of the Not1 protein (Not1₁₋₁₀₀₀), which could be expressed and purified. We then used experimental approaches (limited proteolysis) to map the precise boundaries within this region.
- B)** Identification of the minimal interacting core of Not1, Caf1 and Ccr4. All lanes correspond to peak fractions from size exclusion chromatography. The 12.5 % (w/v) SDS gel was stained with Coomassie. Upon mixing individually purified proteins, full-length Caf1 co-eluted with Not1₁₋₁₀₀₀, Not1₁₅₄₋₁₀₀₀ or Not1₇₅₄₋₁₀₀₀ in size exclusion chromatography (lanes 1 to 3). However, Caf1 failed to interact with Not1₁₋₁₅₃ and with Not1₁₁₋₇₅₃ (data not shown). We used insect cells to co-express and purify a construct of Ccr4 encompassing the functionally important residues (110-837, referred to as Ccr4) together with

the nuclease domain of Caf1 (residues 146-433, referred to as Caf1). Upon mixing, the Caf1-Ccr4 complex and Not₁₇₅₄₋₁₀₀₀ co-eluted in size exclusion chromatography, indicating that they form a ternary complex (lane 4).

```

1      10      20      30      40      50      60
Not1_Sc/1-1000  MLSARYRDLNATASLETSKERQAAQIVIAQISLLETLNNDNFESVEREIRHILDRS...SVDIYIKVM
Not1_Hs/340-1309  VLDVLRKELNPLMFK.....EVTYELDHPGQIRDSKGLNVTYGLQRGLGHEVFP.VDLIYRPMK
Not1_Dm/415-1373  VFDVALKEVPLMFK.....EVTYELDHPGQIRDSKGLNVTYGLQRGLGHEVFP.VDLIYRPMK
Not1_Dr/340-1306  VLDVLRKELNPLMFK.....EVTYELDHPGQIRDSKGLNVTYGLQRGLGHEVFP.VDLIYRPMK
Not1_Xt/340-1300  VLDVLRKELNPLMFK.....EVTYELDHPGQIRDSKGLNVTYGLQRGLGHEVFP.VDLIYRPMK

70      80      90      100     110     120     130
Not1_Sc/1-1000  ..BRLVTLSSRDITLQACKLLEQENLLRLLLEPAQLDCKKSTDLIELLKERTFNQGFQKQFGTISLFI
Not1_Hs/340-1309  HACCQSFQIHSLINPEICFADYPCHTVATDILAPDED.....DNRRIATWKS.....
Not1_Dm/415-1373  HACCQSFQIHSLINPEICFADYPCHTVATDILAPDED.....DNRRIATWKS.....
Not1_Dr/340-1306  HACCQSFQIHSLINPEICFADYPCHTVATDILAPDED.....DNRRIATWKS.....
Not1_Xt/340-1300  HACCQSFQIHSLINPEICFADYPCHTVATDILAPDED.....DNRRIATWKS.....

140     150     160     170     180     190     200
Not1_Sc/1-1000  DLPDKSANKDITISDRSSQINDFKTKMHTNYRNFPLQTTDREESNRDQL..LHSREESLNDLQ
Not1_Hs/340-1309  .....DLISGLRLAEVQVEQVKQ...DFRFPKHCDDMDVLALQINTSWHTRHHELISDM
Not1_Dm/415-1373  .....DLVSGVLSIADKCYTQVRE...DFKFPANQCDDMDVLALQINTSWHTRHHELISDM
Not1_Dr/340-1306  .....DLVSGVLSIADKCYTQVRE...DFKFPANQCDDMDVLALQINTSWHTRHHELISDM
Not1_Xt/340-1300  .....DLISGLRLAEVQVEQVKQ...DFRFPKHCDDMDVLALQINTSWHTRHHELISDM

210     220     230     240     250     260
Not1_Sc/1-1000  ALLISE.....LSPRSQNLQNDPFRWLTTPMVLDAITRGNVIARSISLQANQINNNRFFNLHSDRY
Not1_Hs/340-1309  FIFGMPHNSAIEHYAHGQGGSPSIROLIMHMAEWYMRG.....QYDARLSRLDVAQDLK
Not1_Dm/415-1373  FIFGMPHNSAIEHYAHGQGGSPSIROLIMHMAEWYMRG.....QYDARLSRLDVAQDLK
Not1_Dr/340-1306  FIFGMPHNSAIEHYAHGQGGSPSIROLIMHMAEWYMRG.....QYDARLSRLDVAQDLK
Not1_Xt/340-1300  FIFGMPHNSAIEHYAHGQGGSPSIROLIMHMAEWYMRG.....QYDARLSRLDVAQDLK

270     280     290     300     310     320     330
Not1_Sc/1-1000  FCGAPMPHTASLSCFLAALHDGPIVDFEFCMVKVFXLDLAIQLHNSVSGNCCFLNLNCFRRVSEIT
Not1_Hs/340-1309  ALGMLNNGTFFAF.....VDLAAALSRREVLKLD.....KWLTD...DKIREGEPFIQAC
Not1_Dm/415-1373  ALGMLNNGTFFAF.....VDLAAALSRREVLKLD.....KWLTD...DKIREGEPFIQAC
Not1_Dr/340-1306  ALGMLNNGTFFAF.....VDLAAALSRREVLKLD.....KWLTD...DKIREGEPFIQAC
Not1_Xt/340-1300  ALGMLNNGTFFAF.....VDLAAALSRREVLKLD.....KWLTD...DKIREGEPFIQAC

340     350     360     370     380     390
Not1_Sc/1-1000  IPMKRSLLVLEIASLMLSEF.....QNRSLSDGPIVDFEFCMVKVFXLDLAIQLHNSVSGNCCFLNLNCFRRVSEIT
Not1_Hs/340-1309  HTFLKRRCPFSILGGLAPEKQD..PKSAQPPETLAT..MLACLQAGSVSQ..ELSETLTHVANCSN..
Not1_Dm/415-1373  IYVLRKRRCPQVINKAKVPE..DQLPPKQAGLPEVTVT..MINCLQAGSVSQ..ELSETLTHVANCSN..
Not1_Dr/340-1306  VYFLKRRCPFSILGGLAPEKQD..PKSAQPPETLAT..MLACLQAGSVSQ..ELSETLTHVANCSN..
Not1_Xt/340-1300  HTFLKRRCPFSILGGLAPEKQD..PKSAQPPETLAT..MLACLQAGSVSQ..ELSETLTHVANCSN..

400     410     420     430     440     450     460
Not1_Sc/1-1000  LLINRTVIDEILITLLIQVNRNSPSSFKDVIS.TITDQSKIVDAKIIINSDDAPIANFLKSLDTCGRU
Not1_Hs/340-1309  .....VNNAR.....QPPGVMPKGRFPSSLDALISPVQIDPLAHTSE
Not1_Dm/415-1373  .....MANNARAQQQQPGLVPPPPPILRGRHMDLPGGIVPPFPQOFPFS..GNL
Not1_Dr/340-1306  .....VNNAR.....QPPGVMPKGRFPSTSLDALISPVQMDPLSAGS
Not1_Xt/340-1300  .....VNNAR.....QPPGVMPKGRFPSTSLDALISPVQIDPLAANAS

470     480     490     500     510     520     530
Not1_Sc/1-1000  DTVINKLFFNEAFKILPCARQIGWEGDFLRTKVSNSVDVVDESLEVQTRMTDNTFFRSK....E
Not1_Hs/340-1309  SLDGSAAPHTSQMGFF.....PNLCSAFSTPQSPAKAFPPSTPMTAFSCDGLSSQDFVGGLCQ
Not1_Dm/415-1373  NADMGFGDPTPNSM.....NLACLRLCGFMGRFPGMSTPPELMT.PCANPFLKSH.QMPOA
Not1_Dr/340-1306  SLDGSAAPHTSQMGFF.....PNLCSAFSTPQSPAKAFPPSTPMTAFSCDGLSSQDFVGGLCQ
Not1_Xt/340-1300  SLDGSAAPHTSQMGFF.....PNLCSAFSTPQSPAKAFPPSTPMTAFSCDGLSSQDFVGGLCQ

540     550     560     570     580     590
Not1_Sc/1-1000  FDLFAPHSLIEVINKCFLDVLQQRPELEFSLLIAPRLINFGFGHDEATLNCDTAGI.....NN
Not1_Hs/340-1309  GSLTGIQMGALGLPAVNDPFFVQRKLGTS.....GLMQPPEQDTLSQVWPEANQHFSA
Not1_Dm/415-1373  FPPNVQMLGRMLPGGQQTFPPFTAP.....PNNPWNADLQIP.....VSK
Not1_Dr/340-1306  .SLSGICSG.LCMTVSSDVSARKNSFP.....GLMPPEQDTLSQVWPEANQHFSA
Not1_Xt/340-1300  GSLTGMCGGLSLPTVNSDPPFGQRKLGTS.....GLMQPPEQDTLSQVWPEANQHFSA

```

600 610 620 630 640 650 660

Notl_Sc/1-1000 D I R K M Q R N V P O R K M S G E . . . L A I K D V I E L L R R L R D S D L R Q Q E Y F C I T T A V I A S F F Q D Y P L D D L A F

Notl_Bs/340-1309 E I D D A N S F P O R I N H P P H P T M S V D V L E M L O R F K D S T I K R R E R V E N C M L R N L F E Y F F P O Y P D K E S H I

Notl_Dm/415-1373 E V E D A N S F P O R I N H P P N P T S I D E V L D I D O R F K E S S R R R E Q E V L L O M L R N L F E Y F F P O Y P D K E S H I

Notl_Dr/340-1306 E I D D A N S F P O R I N H P P H P T M S V D V L E M L O R F K D S T I K R R E R V E N C M L R N L F E Y F F P O Y P D K E S H I

Notl_Xt/340-1300 E I D D A N S F P O R I N H P P H P T M S V D V L E M L O R F K D S T I K R R E R V E N C M L R N L F E Y F F P O Y P D K E S H I

670 680 690 700 710 720

Notl_Sc/1-1000 I S V L P G S M I F Q D L R C F V L D V A F R I N R F A K E P P S K M F K A V O A I Y A P R I R L A Y P O Y C R D L L

Notl_Bs/340-1309 I A C L P G I I E R C K G V T Y M A L C L A L R V L E A L R K P F C S K M Y P G I A L D R E K N L K D P O Y C H L A S I S H F

Notl_Dm/415-1373 I A C L P G I I D R N V P P F V A L C L S L R C V L D A L R K P D C S K L Y P G V T A L D R E K N L K D P O Y C H L A S I S H F

Notl_Dr/340-1306 I A C L P G I I E R C K G V T Y M A L C L A L R V L E A L R K P F C S K M Y P G I A L D R E K N L K D P O Y C H L A S I S H F

Notl_Xt/340-1300 I A C L P G I I E R C K G V T Y M A L C L A L R V L E A L R K P F C S K M Y P G I A L D R E K N L K D P O Y C H L A S I S H F

730 740

Notl_Sc/1-1000 R D V P A L N S Q A Q V Y S I V E A A D L A

Notl_Bs/340-1309 H O P P H L Q E Y I E Y G Q S R D P V . K M G . S I T T P G S I A L A Q A Q A Q V P A K A P L A G Q V S T M W T L S T R P T V

Notl_Dm/415-1373 S D F P P H I Q V E Y C M G Q E P P C L I G S W T I P S A I S S C F G T E T H K N S S L C M D A R S P G

Notl_Dr/340-1306 L O P P H L Q E Y I E Y G Q S R D P V . K M G . S I T T P G S L A L A Q A Q A Q S O . P F K A P Q G A S T L V T A T T R P T A

Notl_Xt/340-1300 H O P P H L Q E Y I E Y G Q S R D P V . K M G . S I T T P G S I A L S Q A Q T P A K . P A S S S A W T P T R P T P

750 760 770 780 790 800

Notl_Sc/1-1000 N A P K E R S R P V Q E M I P L R F F A V D E V S C I N O R C A D R D V E K V L V L N V T L A N L N K V D E L

Notl_Bs/340-1309 A K T V N T R P P G V S F K K D V P P S I N T T I D L L V A T D Q E R I V E P E N N I Q E K I A I F N N S Q S M P Q K V E S L

Notl_Dm/415-1373 S Q P K S N A V S H A T R M K S I A W A T N I D L L V A . N O E R K V T V P E P F Q D R T A I F N S Q L I P Q C D E I

Notl_Dr/340-1306 A K T T I T R P P A V G F K K D V P P S I N T T I D L L V A T D Q E R I V E P E N N I Q E K I A I F N N S Q S M P Q K V E S L

Notl_Xt/340-1300 A K T I T I T R P P G V S F K K D V P P S I N T T I D L L V A T D L A C I V E P E N N I Q E K I A I F N N S Q S M P Q K V E S L

810 820 830 840 850 860 870

Notl_Sc/1-1000 K K L L P N T P S P S T L V T Q A K P D P P I D Y K R V I V A M S C E L L H Q F N V N V L G O L V L S F R D R O A I . . .

Notl_Bs/340-1309 K E T V K E F P P V S Q Y L V M K R V S I E P P S L Y S N F L D L K N P E F V K M V L N E Y R N I K V L L T S D K A A A N F S

Notl_Dm/415-1373 K E I M T E F P P V L A Q Y L V L K R A S H E P P P L Y N F L D A L K N G E I N R V T K E D L R I K V L L R S D K G V I N F S

Notl_Dr/340-1306 K E T V K E F P P V S Q Y L V M K R V S I E P P S L Y S N F L D L K N P E F V K M V L N E Y R N I K V L L T S D K A A A N F S

Notl_Xt/340-1300 K E T V K D D P P V S Q Y L V M K R V S I E P P S L Y S N F L D A L K H L E F N S M V L A E D I R I K V L L T S D K A A A N F S

880 890 900 910 920 930 940

Notl_Sc/1-1000 D K K I L K N L A S M E G C I T T A L N K P I K I N T A F R E M D E A Y K E N L I D I V V P P V K I I D R A S E S K I I R P P N P

Notl_Bs/340-1309 D R S L L K N L G H M L G I I L A K N R P L H T D L D L K S L I E A Y V K C Q O E L L Y V V P P V A K V E S S I R S V V R P P P

Notl_Dm/415-1373 D R S L L K N L G H M L G I I L A K N R P L Q I D L D L K S L I E A Y V K C Q O E L L Y V V P P V A K V E S S I R S V V R P P P

Notl_Dr/340-1306 D R S L L K N L G H M L G I I L A K N R P L Y T D L E L K S L I E A Y V K C Q O E L L Y V V P P V A K V E S S I R S V I R P P N P

Notl_Xt/340-1300 D R S L L K N L G H M L G I I L A K N R P L H T D L D L K S L I E A Y V K C Q O E L L Y V V P P V A K V E S S V S V V R P P N P

950 960 970 980 990 1000

Notl_Sc/1-1000 H V V G D K L I S I N E R A N K L S G P P E S V L N S F N L D K S K R P S N F I N T P E W I E P L S G

Notl_Bs/340-1309 H V M A M N V A B E H Q E D I L K N R F E I V L C N L A L D I N E K P G N L K D K D R L K N L D E

Notl_Dm/415-1373 H V M G M N V A B E H Q E D I L K N R F E I V L C N L N L E L A R Q V I Y L K D P N R I H R I E E

Notl_Dr/340-1306 H V M G M N V A B E H Q E S I D I L K N R F E I V L C N L S W D I T T K P C S L R K D K R I K T L E E

Notl_Xt/340-1300 H V M A M N V A B E H Q E D I L K N R F E I V L C N L V I D I N D K P G T L L K D K D R L K S L D E

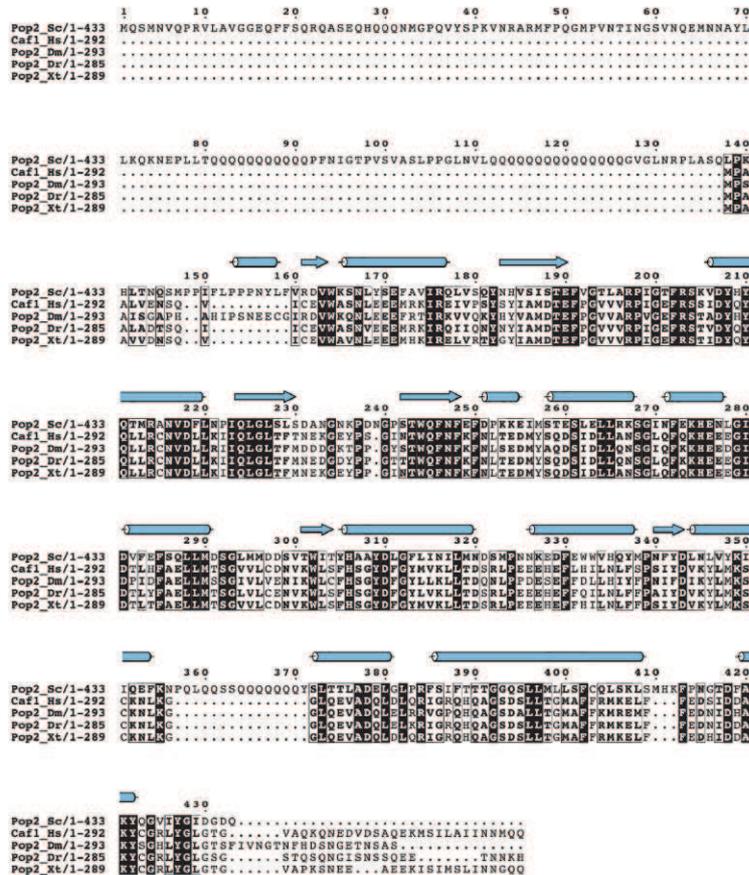


Figure S2. Structure-Based Sequence Alignment, Related to Figure 2

Sequence alignments including of Not1₁₋₁₀₀₀, full-length Caf1 and Crr4 including orthologues from *S. cerevisiae* (Sc), *H. sapiens* (Hs), *D. melanogaster* (Dm), *D. rerio* and *X. tropicalis*. Conserved residues are highlighted in gray. The secondary structure elements are shown above the sequences (as colored-filled cylinders for α -helices and arrows for β -strands). The colors of the secondary structure elements is according to the schematics in Figure 1A, with red for Not1₁₅₄₋₇₅₃, orange for Not1₇₅₄₋₁₀₀₀, blue for Caf1 and pink for Ccr4. In the Ccr4 alignment, the secondary structure elements from the structure of the human orthologue (Wang et al., 2010) are shown in gray.

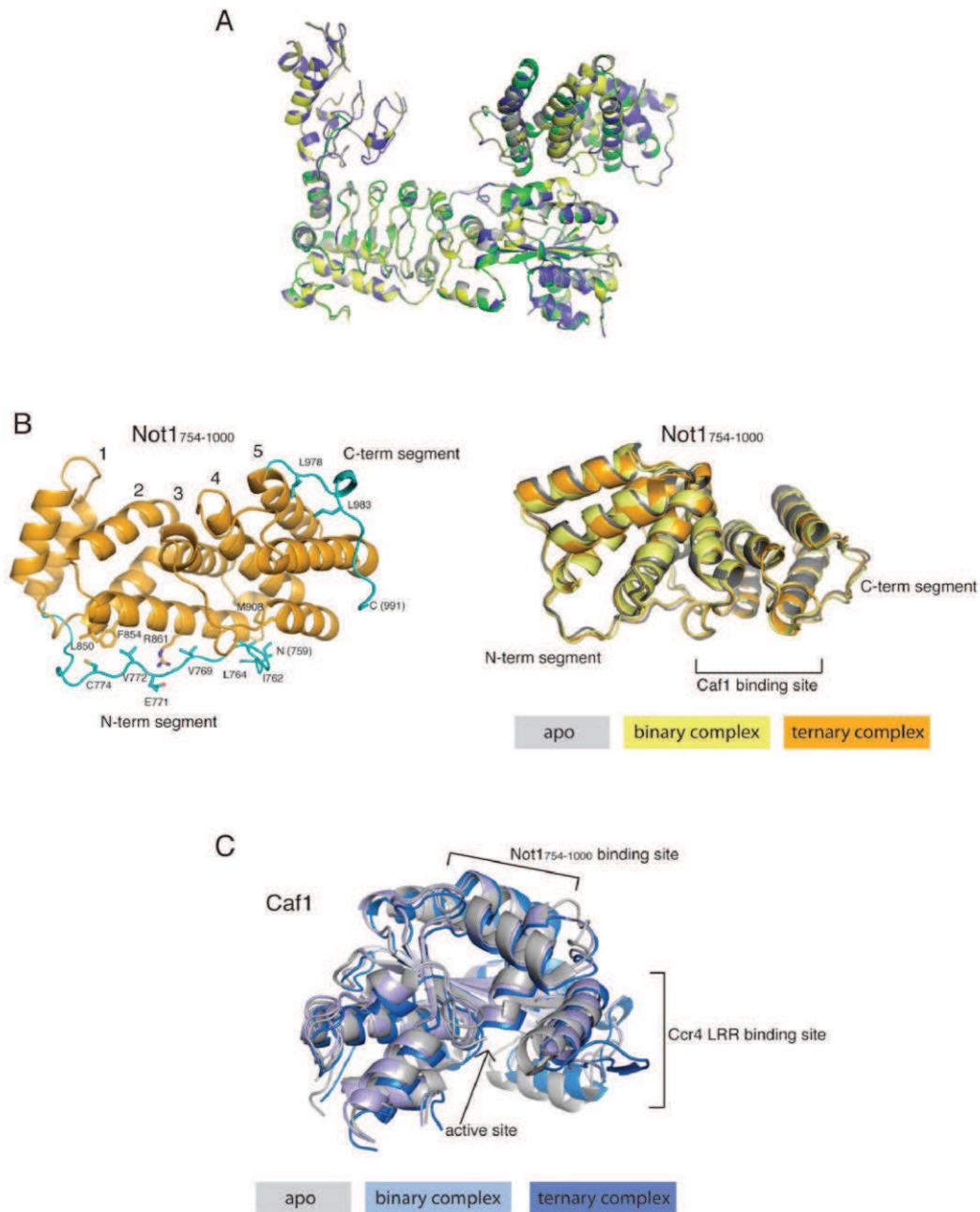


Figure S3. Similarity of the Not1₇₅₄₋₁₀₀₀ and Caf1 Proteins in the Bound and Unbound Forms, Related to Figure 3

A) Superposition of the 4 independent copies of Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 in the crystal structure. Each copy is shown in a different colour (gray, yellow, green and blue).

- B) On the left panel, the interactions of the N- and C-terminal segments of Not1₇₅₄₋₁₀₀₀ (in cyan) are shown with the underlying helices of the HEAT repeats. The N-terminal segment consists of 20 residues and wraps around the MIF4G domain, stretching across the A helices of HEAT 2 to 4. The C-terminal segment consists of 15 residues and packs against the side of HEAT 5. Highlighted are conserved residues involved in hydrophobic and polar interactions. To view the N-terminal segment, the structure has been rotated about 45° clockwise with respect to the view in Figure 2A, right panel. The right panel shows the superposition of the structure of Not1₇₅₄₋₁₀₀₀ in isolation (in gray) with the structure of Not1₇₅₄₋₁₀₀₀ in the binary complex with Caf1 (in yellow) and with the structure of Not1₇₅₄₋₁₀₀₀ in a ternary complex with Caf1 and Ccr4 (in orange). In both cases, Not1 superposes with an rmsd of less than 0.53 Å overall all C α atoms. The structure is viewed in a similar orientation as in Figure 2A, right panel.
- C) Superposition of Caf1 in isolation (in gray) (Thore et al., 2003) with the structure of Caf1 in the binary complex with Not1 (in light blue) and with the structure of Caf1 in the ternary complex with Not1 and Ccr4 (right side, in blue).

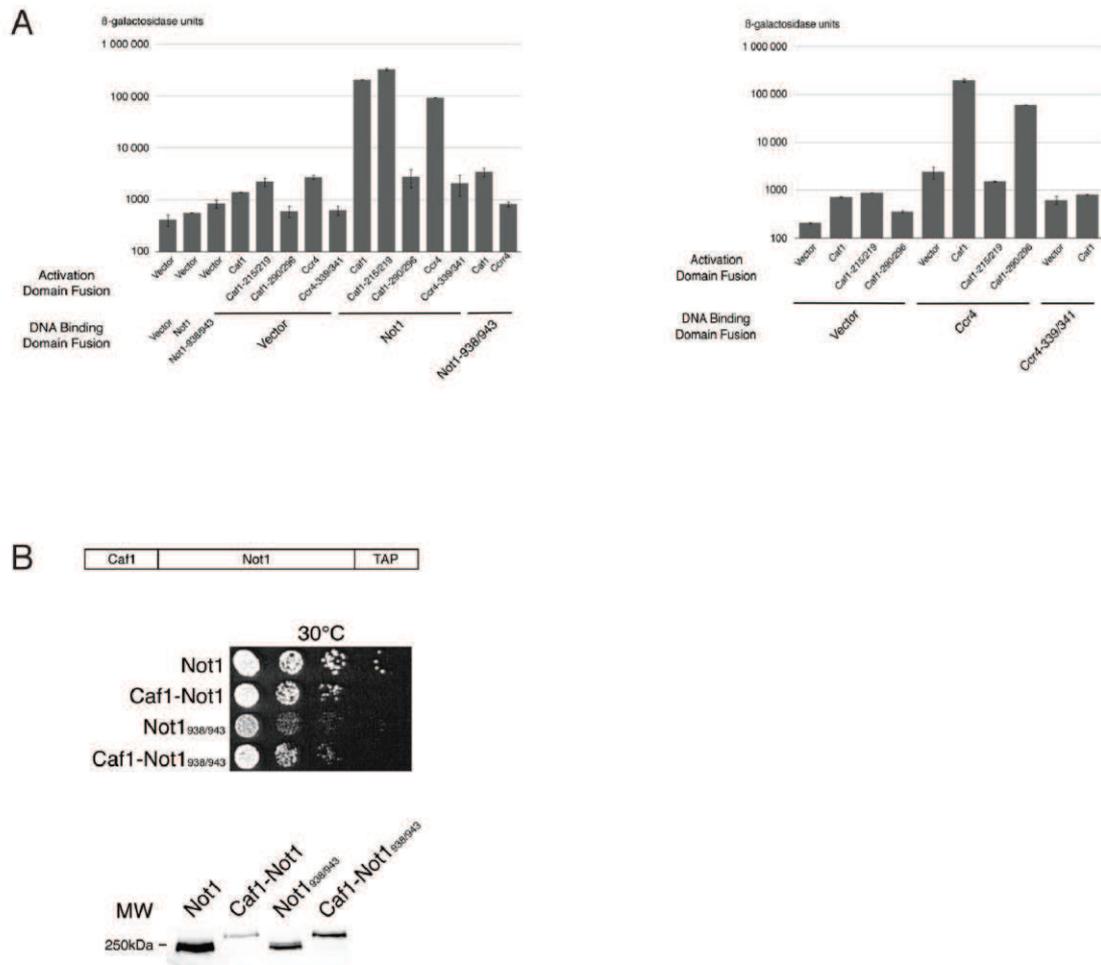


Figure S4. Effect of Structure-Based Mutations Monitored Using the Two-Hybrid Assay, Related to Figure 4

- A)** Two-hybrid interactions between wild type and mutant Not1 (as DNA binding domain fusions) and either Caf1 (wild type or mutant derivatives) or Ccr4 (wild type or mutant derivatives) (as activation domain fusions) monitored by β -galactosidase production. The average of two measurements (each in duplicate) is plotted, together with the error bars. A logarithmic scale is used for β -galactosidase activities to allow the simultaneous visualization of low background activities and strong positive interactions. Details of the experiments are in Supplemental material. Full length proteins were used for these assays.
- B)** Same as panel A, except that two-hybrid interactions between wild type and mutant Ccr4 (as DNA binding domain fusions) and wild type and mutant Caf1 (as activation domain fusions) were monitored.

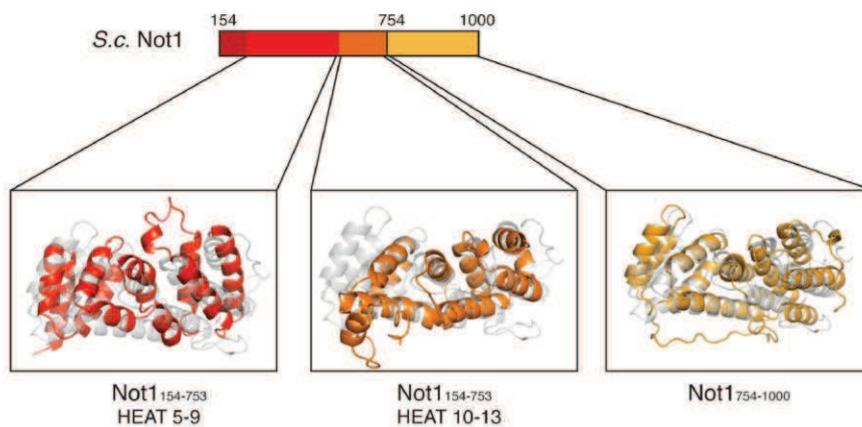


Figure S5. The N-Terminal Half of Not1 Contains Domains Related to MIF4G, Related to Figure 5

The scheme of N-terminal half of Not1 shows the Not1₇₅₄₋₁₀₀₀ domain in orange and the Not1₁₅₄₋₇₅₃ domain in three different shades of red, according to the division of HEAT repeats described in the text. The MIF4G domain of Not1₇₅₄₋₁₀₀₀ is canonical (right panel), as shown with the superposition of MIF4G1 (in gray). The four HEAT repeats at the C-terminus of Not1₁₅₄₋₇₅₃ (HEAT repeats 10-13) also superpose well with four of the HEAT repeats of MIF4G1 (central panel). HEAT repeats 5 to 9 of Not1₁₅₄₋₇₅₃ are instead more divergent (left panel).

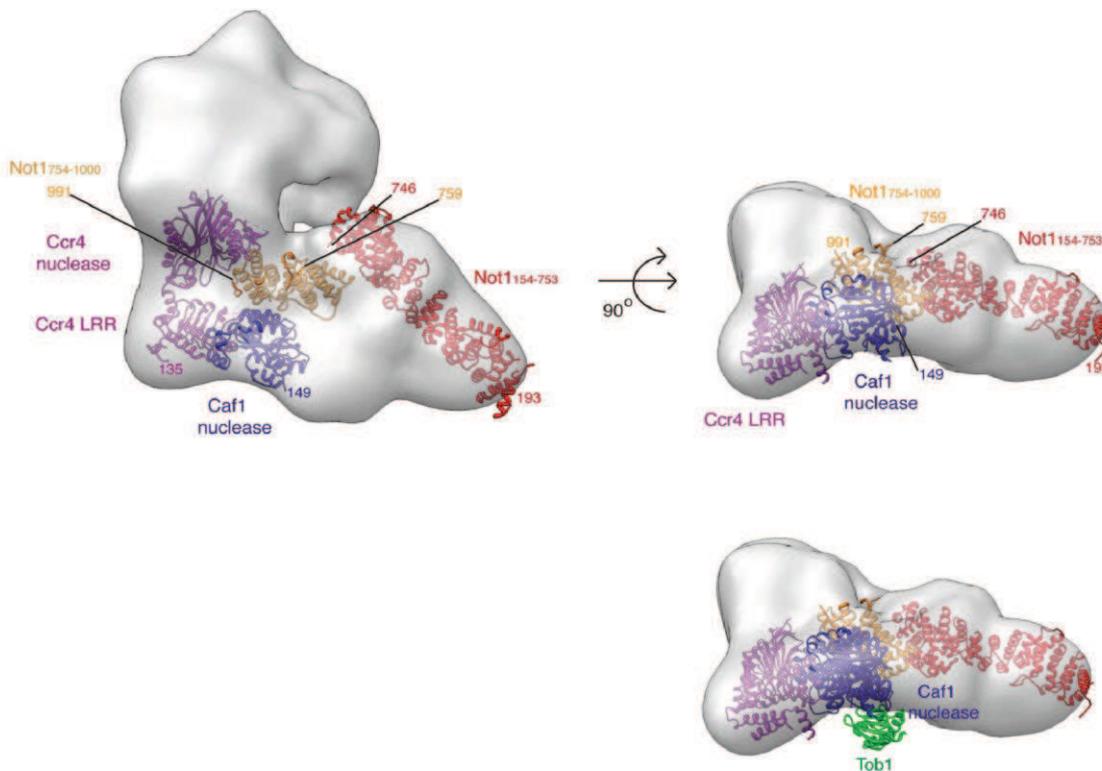


Figure S6. Pseudo-Atomic Model of the Nuclease Module of the Ccr4-Not Complex, Related to Figure 6

In the upper panels, the negative-stain EM map of the Ccr4-Not complex (EMDB code 1901) (Nasertorabi et al., 2011) is shown as a gray surface in two views related by a 90° rotation around a horizontal axis. The atomic coordinates of the Not1₁₅₄₋₇₅₃ and the Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 structures were tentatively docked into the EM map and colored as in Figures 2A and 5A.

As the Not1₇₅₄₋₁₀₀₀ domain is located centrally within the Not1 sequence, we docked the Not1₇₅₄₋₁₀₀₀-Caf1-Ccr4 structure centrally in the EM map. We positioned Ccr4 at the hinge region of the EM density based on the similarity of the L-shape surfaces and on the observation that flexibility in this region was observed both in the crystal structure and in the EM study (Nasertorabi et al., 2011). In this model, the C-terminus of Not1₇₅₄₋₁₀₀₀ would point towards the thicker arm of the Ccr4-Not complex, where the protein-dense Not module is expected to reside based on the components of complex and their known binding sites on Not1 (Collart and Timmers, 2004; Nasertorabi et al., 2011). In the model, the Not1₇₅₄₋₁₀₀₀ N-terminus would be positioned within the thinner arm of the EM density, pointing towards an extended platform-like ridge. Here, the Not1₁₅₄₋₇₅₃ crystal structure could be docked with good agreement with the features of the map and with a compatible distance to Not1₇₅₄₋₁₀₀₀. Although the docking is speculative and is unlikely to be precise in details, it suggests that the deadenylase module (comprising the N-terminal half of Not1, Caf1 and Ccr4) would explain most of the density of the thinner arm observed in the Ccr4-Not EM analysis. This region of the molecule might be even more extended in solution in the

absence of crosslinking and could contribute a large surface for protein-protein interactions.

The lower panel shows the additional superposition of the Caf1-Tob1 structure (Horiuchi et al., 2009). Tob1, a regulator that accelerates mRNA decay, is not present in yeast. Upon Caf1 superposition, Tob1 (in green) resides outside the EM density. The figure panels were generated with the program Chimera (Goddard and Ferrin, 2007).

Table S1. Yeast Strains Used in This Work, Related to the Experimental Procedures

Strain	Genotype
BMA64	<i>MATa, ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1</i> (Baudin-Baillieu et al. 1997)
T26N28	<i>MATa ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, ΔNOT1::HIS3 pFL 38 (NOT1)</i>
T23N3	<i>MATa, ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, POP2::TRP1</i>
T23N10	<i>MATa ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, CCR4::TRP1</i>
BSY2681	<i>MATa ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, ΔNOT1::HIS3 pFL39-NOT1-3HA</i>
BSY2682	<i>MATa ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, ΔNOT1::HIS3 pFL39-not1-F938E-3HA</i>
BSY2683	<i>MATa ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, ΔNOT1::HIS3 pFL39-not1-P943Y-3HA</i>
BSY2684	<i>MATa ade2-1, leu2-3,112, his3-11,15, ura3-1, Δtrp1, ΔNOT1::HIS3 pFL39-not1-938/943-3HA</i>

Table S2. Plasmids Used in This Work, Related to the Experimental Procedures

Plasmid	Description	Comments
pBS4721	pRS413-Caf1-6His-3VSV	Insertion of tag in pBS4003 (pRS413-Caf1) using pU6H3VSV (De Antoni and Gallwitz, 2000) as a source of tag. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4562-OBS4567
pBS4725	pRS413-Caf1-M290K-6His-3VSV	Site-directed mutagenesis of pBS4721 performed using QuickChange method (Stratagene) and oligonucleotides OBS4657 and OBS4658
pBS4726	pRS413-Caf1-M296K-6His-3VSV	Site-directed mutagenesis of pBS4721 performed using QuickChange method (Stratagene) and oligonucleotides OBS4659 and OBS4660
pBS4727	pRS413-Caf1-M290K, M296K-6His-3VSV	Site-directed mutagenesis of pBS4721 performed using QuickChange method (Stratagene) and oligonucleotides OBS4661 and OBS4662
pBS4728	pRS413-Caf1-A215E-6His-3VSV	Site-directed mutagenesis of pBS4721 performed using QuickChange method (Stratagene) and oligonucleotides OBS5462 and OBS5463
pBS4729	pRS413-Caf1-F219E-6His-3VSV	Site-directed mutagenesis of pBS4721 performed using QuickChange method (Stratagene) and oligonucleotides OBS5464 and OBS5465
pBS4730	pRS413-Caf1-A215E, F219E-6His-3VSV	Site-directed mutagenesis of pBS4721 performed using QuickChange method (Stratagene) and oligonucleotides OBS5466 and OBS5467
pBS2618	pDONR-Caf1	Gateway cloning
pBS4736	pDONR-Caf1-M290K, M296K	Site-directed mutagenesis of pBS2618 performed using QuickChange method (Stratagene) and oligonucleotides OBS4661 and OBS4662
pBS4739	pDONR-Caf1-A215E, F219E	Site-directed mutagenesis of pBS2618 performed using QuickChange method (Stratagene) and oligonucleotides OBS5466 and OBS5467
pBS2672	pDEST22-Caf1	LR reaction between pDEST22 and pBS2618
pBS4756	pDEST22-Caf1-M290K, M296K	LR reaction between pDEST22 and pBS4736
pBS4762	pDEST22-Caf1-A215E, F219E	LR reaction between pDEST22 and pBS4739
pBS2661	pDEST32-Caf1	LR reaction between pDEST32 and pBS2618
pBS4757	pDEST32-Caf1-M290K, M296K	LR reaction between pDEST32 and pBS4736
pBS4763	pDEST32-Caf1-A215E, F219E	LR reaction between pDEST32 and pBS4739
pBS4779	pRS415-Ccr4	Cloning of PCR fragment amplified from yeast genomic DNA in pRS415
pBS4781	pRS415-Ccr4-6His-3VSV	Insertion of tag in pBS4779 (pRS415-Ccr4) using pU6H3VSV (De Antoni and Gallwitz, 2000) as a source of tag. Cloning was performed using the ColdFusion kit according to the manufacturer

		recommendation and oligonucleotides OBS5429-OBS5434
pBS4783	pRS415-Ccr4-L339E-6His-3VSV	Site-directed mutagenesis of pBS4781 performed using QuickChange method (Stratagene) and oligonucleotides OBS5456 and OBS5457
pBS4784	pRS415-Ccr4-L341E-6His-3VSV	Site-directed mutagenesis of pBS4781 performed using QuickChange method (Stratagene) and oligonucleotides OBS5458 and OBS5459
pBS4785	pRS415-Ccr4-L339E, L341E-6His-3VSV	Site-directed mutagenesis of pBS4781 performed using QuickChange method (Stratagene) and oligonucleotides OBS5460 and OBS5461
pBS2683	pDONR-Ccr4	Gateway cloning
pBS4792	pDONR-Ccr4-L339E, L341E	Site-directed mutagenesis of pBS2683 performed using QuickChange method (Stratagene) and oligonucleotides OBS5460 and OBS5461
pBS2702	pDEST22-Ccr4	LR reaction between pDEST22 and pBS2683
pBS4798	pDEST22-Ccr4-L339E, L341E	LR reaction between pDEST22 and pBS4792
pBS2703	pDEST32-Ccr4	LR reaction between pDEST32 and pBS2683
pBS4799	pDEST32-Ccr4-L339E, L341E	LR reaction between pDEST32 and pBS4792
pBS4040	pFL39-Not1	Not1 inserted in pFL39 backbone
pBS4802	pFL39-Not1-6His-3HA	Insertion of tag in pBS4040 (pFL39-Not1) using pU6H3HA (De Antoni and Gallwitz, 2000) as a source of tag. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4570-OBS4575
pBS4803	pFL39-Not1-F938E-6His-3HA	Mutant derived from pBS4802. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4589-OBS4594. An expression plasmid containing the mutated Caf1-binding domain of Not1 served as source of mutation
pBS4804	pFL39-Not1-P943Y-6His-3HA	Mutant derived from pBS4802. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4589-OBS4594. An expression plasmid containing the mutated Caf1-binding domain of Not1 served as source of mutation
pBS4805	pFL39-Not1-F938E, P943Y-6His-3HA	Mutant derived from pBS4802. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4589-OBS4594. An expression plasmid containing the mutated Caf1-binding domain of Not1 served as source of mutation
pBS2873	pDONR-Not1	Gateway cloning

pBS4812	pDONR-Not1-F938E, P943Y	Mutant derived from pBS2873. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4589-OBS4594. An expression plasmid containing the mutated Caf1-binding domain of Not1 served as source of mutation
pBS2872	pDEST22-Not1	LR reaction between pDEST22 and pBS2873
pBS4817	pDEST22-Not1-F938E, P943Y	LR reaction between pDEST22 and pBS4812
pBS2876	pDEST32-Not1	LR reaction between pDEST32 and pBS2873
pBS4818	pDEST32-Not1-F938E, P943Y	LR reaction between pDEST32 and pBS4812
pBS4806	pFL39-Not1-TAP	Insertion of tag in pBS4040 (pFL39-Not1) using pBS1479 (Rigaut et al., 1999) as a source of tag. Cloning was performed using the ColdFusion kit according to the manufacturer recommendation and oligonucleotides OBS4570, OBS4922, OBS4923, OBS4573, OBS4924 and OBS4925
pBS4823	pFL39-Not1 (1-1696)- TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS482 and OBS5053
pBS4824	pFL39-Not1 (981-2108)- TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS5072, OBS5076, OBS5075 and OBS5077
pBS4825	pFL39-Not1 (1081- 2108)-TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS5072, OBS5078, OBS5075 and OBS5079
pBS4828	pFL39-Not1 (154-2108)- TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS5072, OBS5468, OBS5469 and OBS5075
pBS4829	pFL39-Not1 (754-2108)- TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS5072, OBS5471, OBS5472 and OBS5075
pBS4830	pFL39-Not1 Δ (754- 1000)-TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS5474, OBS5475, OBS5476 and OBS5473
pBS4831	pFL39-Not1 Δ (154-753)- TAP	Deletion of pBS4806 constructed by PCR-fusion using oligonucleotides OBS5072, OBS5477, OBS5478 and OBS5479

Supplemental References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* *66*, 213–221.

Baudin-Baillieu, A., Guillemet, E., Cullin, C. and Lacroute, F. (1997) Construction of a yeast strain deleted for the TRP1 promoter and coding region that enhances the efficiency of the polymerase chain reaction-disruption method. *Yeast* *13*, 353-356.

Boeck, R., Lapeyre, B., Brown, C.E., and Sachs, A.B. (1998). Capped mRNA degradation intermediates accumulate in the yeast *spb8-2* mutant. *Mol Cell Biol.* *18*, 5062-5072.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* *62*, 1002–1011.

De Antoni, A., and Gallwitz, D. (2000). A novel multi-purpose cassette for repeated integrative epitope tagging of genes in *Saccharomyces cerevisiae*. *Gene* *246*, 179–185.

Decker, C.J., and Parker, R. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes Dev.* *7*, 1632-1643.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* *66*, 486–501.

Feilotter, H.E., Hannon, G.J., Ruddell, C.J., and Beach, D. (1994). Construction of an improved host strain for two hybrid screening. *Nucleic Acids Res* *22*, 1502–1503.

Goddard, T.D., and Ferrin, T.E. (2007). Visualization software for molecular assemblies. *Curr Opin Struct Biol* *17*, 587–595.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* *40*, 658–674.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* *17*, 1030–1032.

Détermination de la structure du module
NOT1-NOT2-NOT5 du complexe CCR4-NOT de
levure.

E.2. Détermination de la structure du module NOT1-NOT2-NOT5 du complexe Ccr4-Not de levure

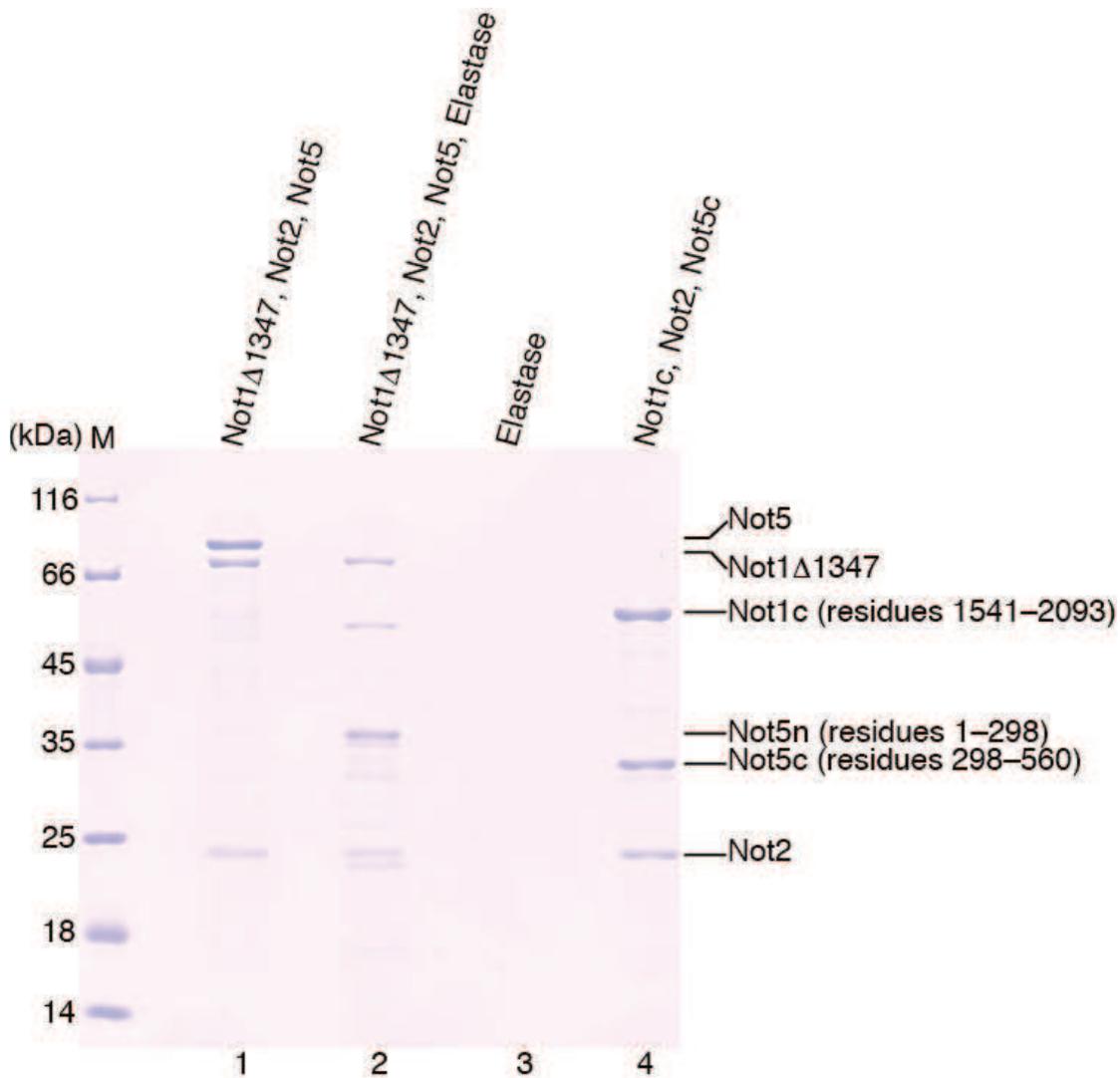
E.2.1. Contexte et objectifs du projet

Le complexe CCR4-NOT est impliqué dans différents aspects de l'expression des gènes incluant la dégradation des ARN, la répression de la traduction et la transcription. Notre laboratoire a conduit une étude structurale qui a mené à la résolution de la structure d'un sous complexe de 120 KD, nommé module NOT, du complexe CCR4-NOT de *S. cerevisiae*. Ce module comprend la partie C terminale de NOT1 complexée avec les protéines NOT2 et NOT5.

E.2.2. Approches expérimentales

Production du complexe NOT1-NOT2-NOT5

J'ai commencé par sous-cloner dans des vecteurs de surexpression chez *E.coli* un complexe comprenant les 750 derniers résidus de NOT1 et les protéines NOT2 et NOT5 individuellement. Le complexe a été reconstitué en mélangeant les protéines séparément purifiées, puis le mélange a été purifié sur une colonne de d'exclusion de taille. Ce complexe a été soumis à une étude de protéolyse ménagée où la protéase élastase a permis d'identifier un complexe plus compact composé des fragments NOT1 (1541–2093), NOT2 FL and NOT5 (298–560). Cette approche s'est révélée particulièrement fructueuse. Elle est désormais utilisée en routine dans notre laboratoire afin de générer des complexes plus compacts, meilleurs candidats pour des essais de cristallisation. Le gel ci dessous illustre cette approche.



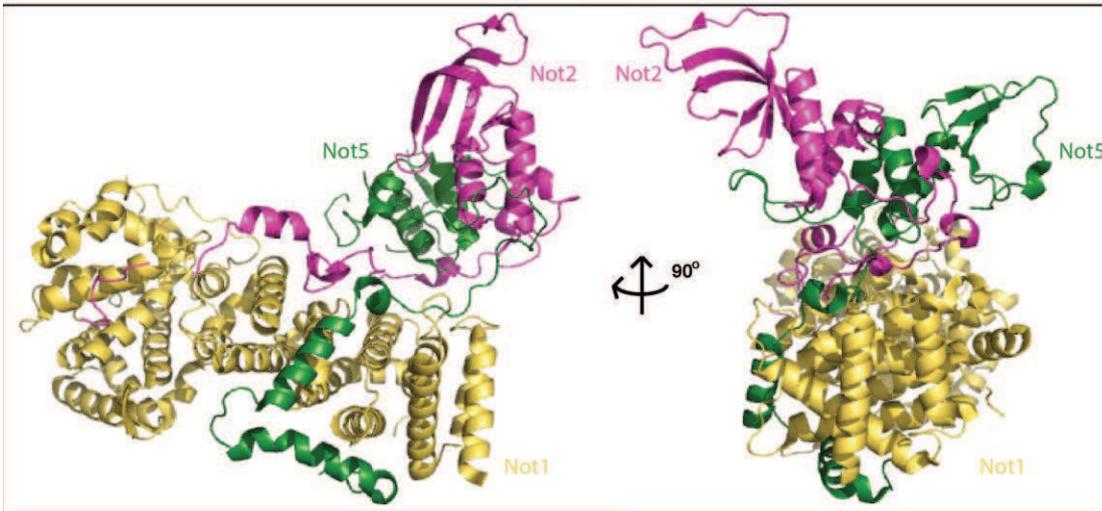
Approche rationnelle de sélection des atomes lourds

Il a été ensuite possible de produire et d'optimiser des cristaux compatibles avec des analyses de cristallographie aux rayons X. Comme il n'existait pas de modèle permettant une approche de résolution de structures par remplacement moléculaire, il a été nécessaire de déterminer un jeu de phases expérimentales. Nous avons tenté de produire des protéines sélénométhioninées mais le faible rendement de production était insuffisant pour la production de cristaux. Afin de contourner ce problème, nous avons décidé de mettre en place une approche rationnelle de dérivatisation des cristaux avec des métaux lourds. En effet, cette étape nécessite très souvent un criblage long et aléatoire par la technique de trempage des cristaux dans des solutions de métaux lourds. Nous avons procédé à un criblage en mélangeant la solution de notre complexe avec différents métaux lourds. Après un échange de tampon, chacun de ces mélanges a été analysé par spectrométrie de masse. Cette stratégie nous a permis d'identifier un dérivé mercurique et de déterminer le nombre de sites de liaison du mercure à notre complexe. Nous avons ensuite réalisé

une co-cristallisation de notre complexe avec une stœchiométrie calculée en considérant le nombre de sites présents sur notre complexe. Cette stratégie a permis de générer les cristaux nécessaires à l'enregistrement de plusieurs jeux de données anomaux. Un modèle initial a été tracé dans une carte expérimentale à 3,2 Å et le modèle final a été affiné à partir d'un jeu de données natif à une résolution de 2,2 Å.

E.2.3. Résumé de l'étude

NOT1 est une protéine d'échafaudage composée de répétitions HEAT. Les protéines NOT2 et NOT5 adoptent un repliement caractéristique des protéines Sm formant un domaine globulaire appelé Boîte Not. Ces deux domaines interagissent entre eux pour former une interface de dimérisation atypique. L'introduction de mutation sur cette interface conduit à des défauts sévères de croissance chez la levure. De plus NOT2 et NOT5 arborent des extensions N-terminales étendues qui viennent s'ancrer sur la protéine NOT1. Le complexe ternaire forme une surface composite capable de se lier à l'ARN *in vitro*. L'ensemble des résultats de cette étude nous a conduit à proposer que le complexe du module NOT forme une plateforme favorisant les interactions macromoléculaires.



Structure cristalline du complexe ternaire Not1-Not2-Not5

Résolution: 2,8 Å

Phasage expérimental avec un composé contenant du mercure(ethyl mercury phosphate EMP)

pdb: 4B8A

Structure and RNA-binding properties of the Not1–Not2–Not5 module of the yeast Ccr4–Not complex

Varun Bhaskar¹, Vladimir Roudko^{2,3}, Jérôme Basquin¹, Kundan Sharma⁴, Henning Urlaub⁴, Bertrand Séraphin^{2,3} & Elena Conti¹

The Ccr4–Not complex is involved in several aspects of gene expression, including mRNA decay, translational repression and transcription. We determined the 2.8-Å-resolution crystal structure of a 120-kDa core complex of the *Saccharomyces cerevisiae* Not module comprising the C-terminal arm of Not1, Not2 and Not5. Not1 is a HEAT-repeat scaffold. Not2 and Not5 have extended regions that wrap around Not1 and around their globular domains, the Not boxes. The Not boxes resemble Sm folds and interact with each other with a noncanonical dimerization surface. Disruption of the interactions within the ternary complex has severe effects on growth *in vivo*. The ternary complex forms a composite surface that binds poly(U) RNA *in vitro*, with a site at the Not5 Not box. The results suggest that the Not module forms a versatile platform for macromolecular interactions.

The Ccr4–Not complex is a large assembly that regulates eukaryotic gene expression at multiple levels. The best-studied function of Ccr4–Not relates to its action as the major deadenylase involved in shortening the poly(A) tail of cellular mRNAs in the cytoplasm (reviewed in ref. 1). Deadenylation by Ccr4–Not is a key step in the constitutive and regulated turnover of mRNAs^{2,3}. Ccr4–Not can also be targeted to *cis*-acting elements in the 3' untranslated region (UTR) of specific transcripts to accelerate their decay (for example, in the case of ARE-containing mRNAs)^{4,5} or to mediate microRNA-dependent repression^{6–8} or translational repression (examples in refs. 9,10). In addition, Ccr4–Not has been implicated in transcription initiation and elongation in the nucleus as well as in ubiquitylation (reviewed in refs. 11,12). The nuclear and cytoplasmic functions of Ccr4–Not have long been thought of as disconnected. However, recent evidence is converging on the functional coupling between mRNA synthesis and degradation¹³.

Ccr4–Not contains several evolutionarily conserved proteins (Not1, Caf1 (also called Pop2), Not2, Not3 or Not5 and Caf40) that are constitutive components of the complex in all species examined to date (yeast^{14,15}, humans^{16,17}, flies^{18,19} and trypanosoma²⁰). Other bona fide subunits of Ccr4–Not are peripheral (for example, Ccr4 and Not4)^{16,18,21} and/or species specific^{15,19,22,23}. Variants of the core complex are likely to exist, as homologs are present both in yeast (Not3 and Not5)²⁴ and humans (for Caf1 and for Ccr4)^{16,17}. The core complex is built around Not1, a large scaffold protein of ~240 kDa (refs. 21,25). The N-terminal half of Not1 associates with the Caf1 and Ccr4 RNases and is involved in the formation of the deadenylase module of the complex^{21,26}. The C-terminal half of Not1 binds Not2, Not3, Not4 and Not5 to form the so-called Not module^{16,21,27}.

Synthetic lethality between the yeast deadenylase subunits and Not subunits suggests that they have separate or only partially overlapping functions²¹. The deadenylase module is connected to the cytoplasmic activities of Ccr4–Not (reviewed in refs. 1,11) and has been studied at the structural and functional level^{28,29}. How the Not module is structured and how it functions are far less clear (reviewed in ref. 30).

Genetic and biochemical studies have shown that Not2, Not3 and Not5 are closely associated^{19,21}. *S. cerevisiae* Not3 and Not5 are currently thought of as paralogous proteins³⁰. Yeast Not5 is reported to be crucial for vegetative growth, whereas Not3 deletion has milder phenotypes²⁴. The only Not3 and Not5 homolog in metazoans (known as Not3) is essential in mice for embryonic development and for control of heart function³¹ and metabolism³² in adults. In metazoans, Not2 is believed to recruit Not3 into the complex^{17,19}, to be important for the integrity of Ccr4–Not^{33,34} and to act as a repressor of promoter activity in the nucleus³⁵. In yeast, Not2 and Not5 have been reported to interact with components of the transcription machinery, specifically with subunits of TFIID^{36–38} and SAGA³³. In addition to data pointing to connections with transcription (reviewed in refs. 11,12), the Not module has also been implicated in mRNA-decay pathways in the cytoplasm^{18,39}. To shed light on how the Not module can mediate these different functions, we have determined the structure and biochemical properties of a core complex from *S. cerevisiae*.

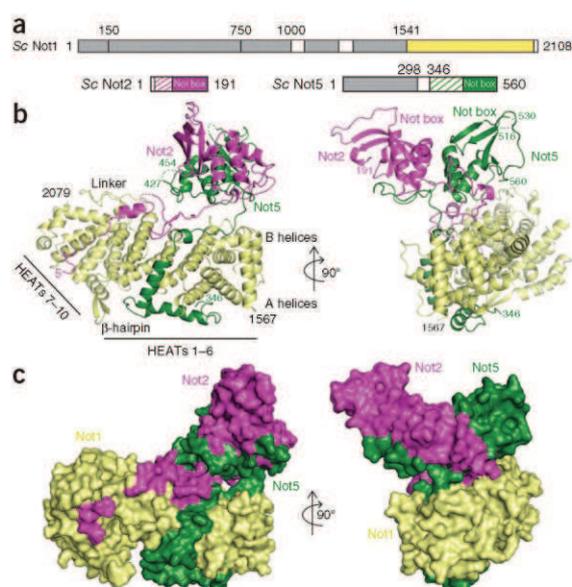
RESULTS

Structure determination of a yeast Not1–Not2–Not5 complex

Yeast two-hybrid and coimmunoprecipitation assays have shown that Not1 (2,108 residues in *S. cerevisiae*) binds Not2, Not3 and Not5 in a region that spans approximately the last 700 residues^{19,21,25} (Fig. 1a).

¹Department of Structural Cell Biology, Max Planck Institute of Biochemistry (MPI Biochemistry), Martinsried, Germany. ²Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Centre National de Recherche Scientifique (CNRS) UMR 7104, Institut National de Santé et de Recherche Médicale (INSERM), U964, Illkirch, France. ³Université de Strasbourg, Illkirch, France. ⁴Max Planck Institute of Biophysical Chemistry, Göttingen, Germany. Correspondence should be addressed to E.C. (conti@biochem.mpg.de).

Received 21 May; accepted 5 September; published online 13 October 2013; doi:10.1038/nsmb.2686



Not2 (191 residues in *S. cerevisiae*) is predicted to contain a poorly structured N-terminal region followed by a conserved domain known as the Not-box domain³⁵ (Fig. 1a). Not5 (560 residues) contains an N-terminal coiled-coil region followed by a low-complexity linker and a C-terminal Not-box domain³⁵ (Fig. 1a). *S. cerevisiae* Not3 has a similar domain architecture as does Not5, but it could not be expressed as full length in a soluble form (V.B. and E.C., unpublished observations). We purified and reconstituted a complex containing the last ~750 residues of Not1, full-length (FL) Not2 and FL Not5, subjected it to limited proteolysis and identified the core complex composed of Not1 residues 1541–2093, Not2 FL and Not5 residues 298–560 (hereafter defined as Not1c, Not2 and Not5c, respectively) (Supplementary Fig. 1).

The Not1c–Not2–Not5c complex yielded crystals diffracting to 2.8-Å resolution. We determined the structure by SAD, using crystals derivatized with mercury, and refined it to an R_{free} of 22.6% and R_{work} of 18.1% with good stereochemistry (Table 1). The two independent copies of the Not1c–Not2–Not5c complex present in the crystal asymmetric unit are virtually identical (superposing with an r.m.s. deviation of 0.85 Å over all C α atoms). The final model includes Not1 residues 1567–2079, Not2 residues 5–191 (with the exception of two disordered segments at residues 14–29 and 44–48) and Not5 residues 346–560 (with the exception of two loops at residues 428–453 and 517–529) (Fig. 1b and Supplementary Fig. 2).

The C-terminal region of Not1 is a scaffold of HEAT repeats

The Not1c–Not2–Not5c structure is organized around Not1 (Fig. 1b). Not1c is built almost entirely of antiparallel α -helices, forming an elongated molecule of the approximate dimensions 85 Å \times 35 Å \times 40 Å. The topology of the secondary-structure elements in Not1c is typical of that observed in HEAT-repeat proteins. Canonical HEAT repeats are characterized by a helix A–turn–helix B motif and are arranged in tandem in an almost-parallel fashion, with a 15° rotation between consecutive repeats⁴⁰. Multiple repeats typically give rise to superhelical structures with a convex layer formed by the A helices and a concave layer formed by the B helices. Not1c contains ten HEAT repeats, which can be grouped into two units.

Figure 1 Structure of a yeast Not1–Not2–Not5 core complex.

(a) Schematic representation of the domain organization of *S. cerevisiae* (Sc) Not1, Not2 and Not5. Color-filled rectangles indicate globular domains present in the crystal structure (yellow, Not1; magenta, Not2; green, Not5). Dashed rectangles indicate low-complexity regions of the molecules with ordered electron density. Gray rectangles indicate globular domains either from previous structures²⁸ or predicted from sequence analysis. (b) Structure of the complex shown in cartoon representation in two orientations (right, front view of the Not boxes; left, side view). Not1 features are labeled in black. Disordered regions are shown as dotted lines. The N- and C-terminal residues are indicated. The labeled linker and β -hairpin refer to the HEAT 6–7 and the HEAT 7–8 inter-repeat loops. This and all other cartoon drawings were generated with PyMOL (<http://www.pymol.org/>). (c) Surface representations of the complex in the same orientations and colors as in b.

The first unit, comprising HEATs 1–6 (residues 1567–1849), has a regular architecture, albeit one less curved than for canonical HEAT-repeat proteins (Fig. 1b). The second unit, comprising HEATs 7–10 (residues 1888–2058), also adopts a regular architecture, with the exception of a long β -hairpin connecting HEATs 7 and 8 and of an additional C-terminal helix (residues 2059–2079). This unit contains four of the five HEAT repeats characteristic of the middle domain of eukaryotic initiation factor 4G (MIF4G)⁴¹ and can therefore be described as an MIF4G-like domain. The 40-residue linker connecting HEATs 6 and 7 wraps around both units and contributes to formation of the extensive hydrophobic core of Not1c. The two HEAT-repeat units pack against each other in a perpendicular fashion resulting in a T-shaped architecture (Fig. 1b and Supplementary Fig. 3a). Interestingly, the domain formed by residues 193–745 in the N-terminal

Table 1 Data collection and refinement statistics

	Native	Hg derivative
Data collection^a		
Space group	$P2_1$	$P2_1$
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	110.45, 109.17, 133.62	109.67, 106.19, 136.02
α , β , γ (°)	90, 94.7, 90	90, 94.0, 90
		Peak
Wavelength	1.00004	1.00606
Resolution (Å)	49.15–2.80 (2.95–2.80)	47.77–3.20 (3.37–3.20)
R_{merge}	6.50 (42.90)	16.70 (80.60)
$I / \sigma I$	17.30 (2.90)	10.60 (2.40)
Completeness (%)	99.50 (97)	100 (100)
Redundancy	4.80 (4.50)	6.90 (7.20)
Refinement		
Resolution (Å)	49.15–2.80 (2.83–2.80)	
No. reflections	77,882	51,653
$R_{\text{work}} / R_{\text{free}}$	0.1812 / 0.2258	
No. atoms	14,019	
Protein	13,978	
Ligand/ion	36	
Water	5	
<i>B</i> factors	67.00	
Protein	67.10	
Ligand/ion	62.20	
Water	39.90	
r.m.s. deviations		
Bond lengths (Å)	0.009	
Bond angles (°)	1.12	

^aOne native and one Hg-derivative crystal were used for data collection. Values in parentheses are for highest-resolution shell.

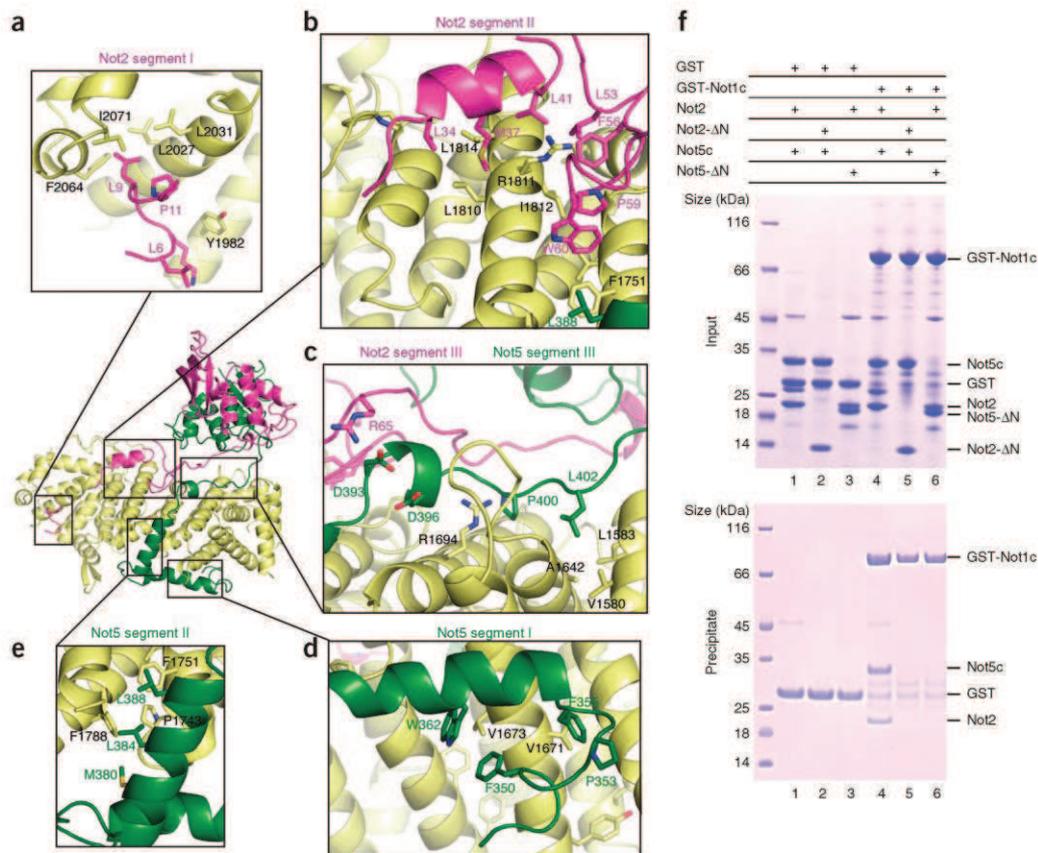


Figure 2 Not1 interacts with extended regions of Not2 and Not5. (a–e) Close-up views of the interactions of Not1 with Not2 and Not5 showing the three segments (I, II and III) of the N-terminal extensions of Not2 and Not5 that form the Not1-binding domains. The positions of the individual close-up views within the complex are indicated at center left. Interacting residues involved in evolutionarily conserved interactions are indicated and labeled (conservation shown in **Supplementary Fig. 2**). (f) Pull-down experiments of GST-tagged Not1c with untagged Not2, Not5c, Not2-ΔN and Not5-ΔN (ΔN refers to the removal of the N-terminal extension involved in Not1 binding). GST is shown as a control. Input samples (top) and samples precipitated on glutathione-agarose beads (bottom), analyzed on 4–12% Bis-Tris NuPAGE gel with MES running buffer, are shown. The proteins corresponding to the bands are indicated on the right.

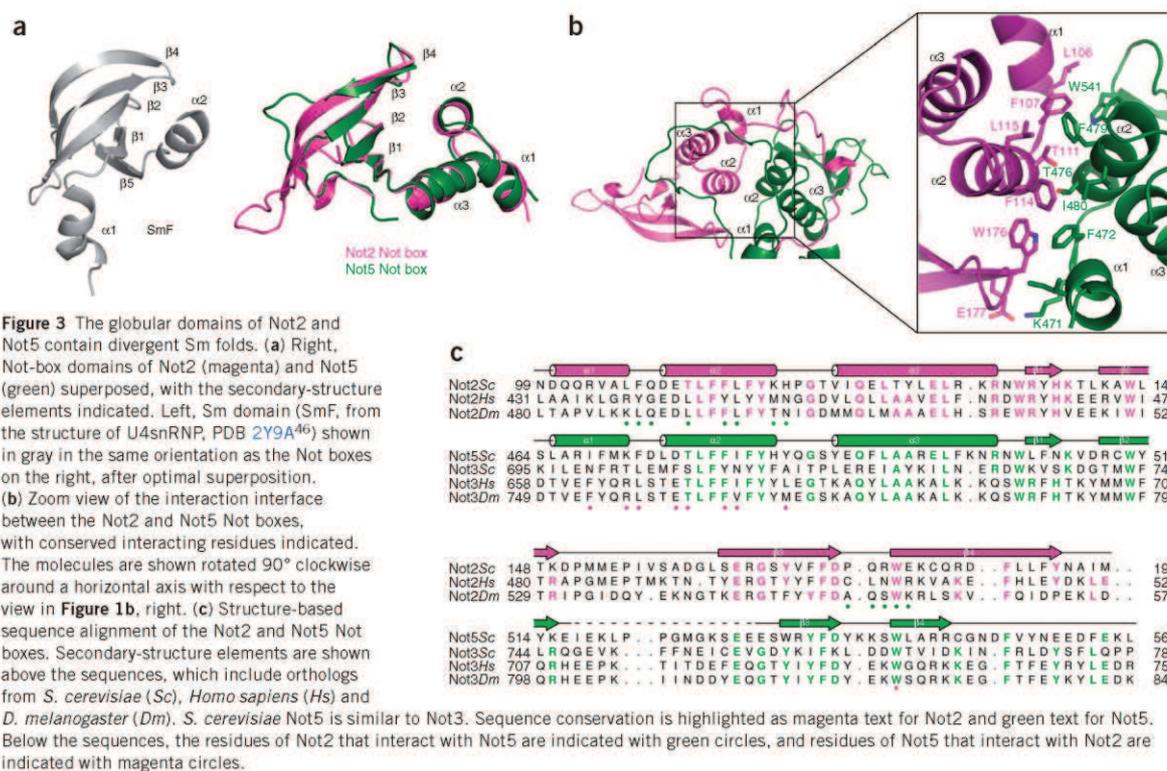
arm of yeast Not1 is also formed by a MIF4G-like unit and a longer HEAT-repeat unit arranged perpendicularly to each other²⁸. Although the relative orientations of the individual units differ in detail, the N-terminal and C-terminal arms of Not1 are built with remarkably similar principles (**Supplementary Fig. 3a**).

Extended regions of Not2 and Not5 wrap around Not1

Not2 and Not5 both contain a globular domain preceded by N-terminal extensions (**Fig. 1b,c**). In the N-terminal extensions, Not2 residues 5–75 and Not5 residues 346–404 mediate binding to Not1c, covering a distance of >100 Å each and burying a total surface of ~3,700 Å². The Not1-binding domain of Not2 can be described as composed of three segments (**Fig. 2**). The first segment (Not2 residues 5–13) binds the MIF4G-like unit of Not1c, mainly at the A helices of HEATs 9 and 10. Here, a conserved hydrophobic pocket of Not1 recognizes Not2 Leu9 (**Fig. 2a**), a conserved residue that has been shown to be functionally important in *in vivo* studies³³. The second segment (Not2 residues 31–64) binds Not1c at the adjacent HEAT-repeat unit, zigzagging over the B helices of HEATs 4–6 (**Fig. 2b**). This segment of Not2

forms a short helix and a hairpin. The helix docks with hydrophobic residues on the conserved surface of HEAT 5 centered at Arg181 and Leu1814. The hairpin wedges into another set of hydrophobic residues in a conserved groove at HEATs 4–5 (from Phe1751 to Ile1812). The third segment (Not2 residues 65–75) extends over the B helices of Not1c toward HEAT 3 (**Fig. 2c**).

The Not1-binding domain of Not5 wraps around HEATs 1–5 of Not1 (**Fig. 1b**) and can also be subdivided into three segments. The first segment (Not5 residues 346–373) contains an α-helix and binds the A helices of Not1c with apolar interactions (**Fig. 2d**). The second segment (Not5 residues 374–391) contains another α-helix and binds the edge of Not1c formed at the intrarepeat connections of HEATs 3–5 through hydrophobic interactions (**Fig. 2e**). The third segment (Not5 residues 392–404) stretches over the B helices of Not1c between HEATs 1–3, making both polar and apolar contacts (**Fig. 2c**). The third segment of Not5 flanks the third segment of Not2 and directly interacts with it through a salt bridge (between Asp393 and Arg65) (**Fig. 2c**). The structure suggests that the Not1-binding domains of Not2 and Not5 bind Not1c synergistically. We tested the effect of



deleting either domain on Not1 binding in pull-down assays with purified proteins. As a control, Not2 and Not5c coprecipitated with glutathione S-transferase (GST)-tagged Not1c (**Fig. 2f**, lane 4). In this assay, Not5c was not coprecipitated with Not1c when Not2 was truncated (to Not2-ΔN, residues 76–191) (**Fig. 2f**, lane 5). Analogously, Not2 did not coprecipitate with Not1c when Not5c was truncated (to Not5-ΔN, residues 405–560) (**Fig. 2f**, lane 6). We concluded that formation of the core of the Not module requires the cooperative binding of Not2 and Not5.

The Not boxes of Not2 and Not5 have divergent Sm-like folds

The globular domains of Not2 and Not5 are positioned on top of the B helices of Not1 HEATs 1–4, sandwiching in between parts of the Not2 and Not5 N-terminal extensions (**Fig. 1b,c**). The globular domains contain the so-called Not boxes. The Not box of Not2 (residues 99–191) consists of three N-terminal helices ($\alpha 1$, $\alpha 2$ and $\alpha 3$) and a β -sheet formed by four antiparallel β -strands adjacent to each other (**Fig. 3a**, **Supplementary Fig. 3b** and **Supplementary Fig. 3c**). The β -sheet is highly bent: strands $\beta 3$ and $\beta 4$ are long and curved, with a conserved glycine residue (Gly166) at the bending point of $\beta 3$. A short C-terminal extension packs against $\beta 1$, creating a small β -barrel. The Not box of Not5 (residues 464–560) is similar in structure to that of Not2, superposing with an r.m.s. deviation of $<1.3 \text{ \AA}$ over all C α atoms (**Fig. 3a** and **Supplementary Fig. 3b**). The main difference is that in Not5 all the β -strands are short, thus resulting in a rather flat β -sheet.

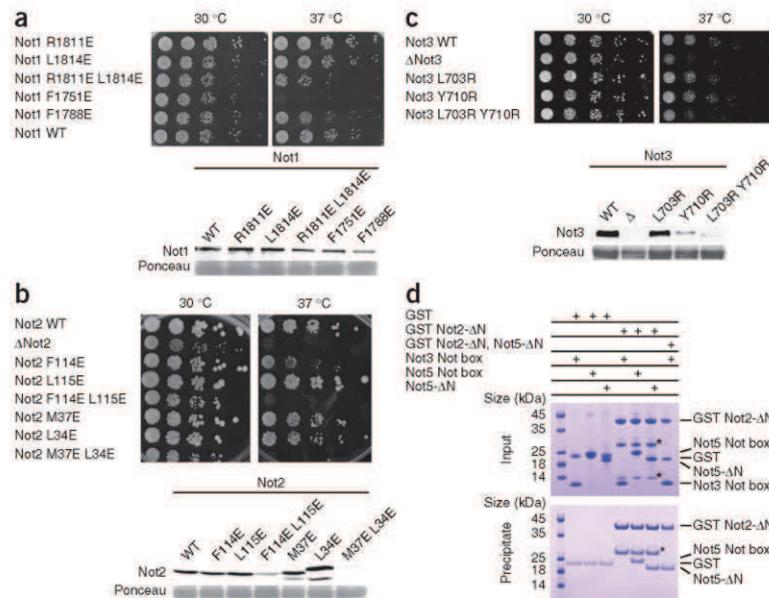
Database searches on the DALI server⁴² for structural similarities to the Not-box domains identified Sm domains as the most similar in fold (r.m.s. deviation of 2.2 \AA and 2.7 \AA with SmD3 and SmF,

respectively⁴³) (**Fig. 3a**). The Not boxes, however, differ from canonical Sm folds in several aspects. First, they lack the characteristic Sm1 and Sm2 signature motifs in the amino acid sequence. At the structural level, the Not boxes lack the fifth β -strand that in Sm proteins mediates the interaction forming dimeric Sm–Sm subcomplexes⁴³ and ring-like structures^{44–46} (**Supplementary Fig. 3b**). The Not boxes of Not2 and Not5 also interact with each other, but in the absence of a fifth β -strand they do so with a different dimerization mechanism that involves the N-terminal α -helices (**Fig. 3b** and **Supplementary Fig. 3b**). Helix $\alpha 1$ of the Not2 Not box packs against the base of the β -sheet of the Not5 Not box and vice versa. Between them, the $\alpha 2$ helices of Not2 and Not5 pack against each other. The dimerization interface is mediated by extensive interactions centered at the conserved Phe114 and Leu115 of Not2 and the corresponding Phe479 and Ile480 of Not5 (**Fig. 3b,c**). Finally, the globular domains are also formed by parts of the N-terminal extensions. Residues 67–93 of Not2 wrap around the Not box of Not5, and residues 408–427 of Not5 wrap around the Not box of Not2 (**Fig. 3b**). The interactions of the N-terminal extensions effectively clamp the Not boxes on the Not1 scaffold (**Fig. 1b,c**).

Not1–Not2–Not5 mutations lead to growth defects *in vivo*

It has previously been shown that deletion of ~400 residues from the C terminus of Not1 is lethal in yeast^{25,28}. In hindsight, these deletions generated Not1 proteins that lacked the last eight HEAT repeats (HEATs 3–10 in the Not1c structure). To test the functional importance of the Not module, we used the structural information to design point mutations that would disrupt specific interactions in the context of tagged full-length proteins.

Figure 4 Analysis of mutants targeting interaction surfaces of the Not module. **(a)** Top, growth assays of Not1 mutants. Serial dilutions of cultures of strains expressing the indicated mutants incubated on YPDA medium at the indicated temperature are shown. Bottom, western blot analysis of Not1-TAP protein from cells expressing the wild-type (WT) or mutant proteins at 30 °C detected by peroxidase-antiperoxidase complex (PAP) prepared in rabbit. Ponceau staining of the membrane, used to assess equal loading in the different lanes, is shown. **(b)** Top, growth assays of Not2 mutants. Mutant strains were analyzed as for Not1 in **a**. Bottom, western blot analysis of Not2 mutant protein levels. Mutant strains were analyzed as for Not1 in **a**, with anti-VSV antibodies. **(c)** Top, growth assays of Not3 mutants. Serial dilutions of cultures of strains expressing the indicated mutant on YPDA medium containing 1 M NaCl at the indicated temperature is shown. Bottom, western blot analysis of Not3 mutant protein levels. Mutant strains were analyzed as for Not2 in **b**. **(d)** Pulldown experiment of GST Not2- Δ N with Not3 Not box, Not5 Not box and Not5- Δ N. The experiment was carried out as in **Figure 2f**. The Not3 Not box, Not5- Δ N and Not5 Not box include residues 685–800, 405–560 and 460–560 (with an N-terminal His-Z tag). The asterisks indicate a degradation product of GST Not2- Δ N.



We constructed four substitutions of Not1 residues contributing to the interaction with Not2 and Not5 (R1811E, L1814E, F1751E or F1788E) and a double mutant (R1811E L1814E) in a tandem affinity purification (TAP)-tagged plasmidic copy of the gene. R1811E and L1814E target the conserved binding site for the second segment of Not2 (**Fig. 2b**). Phe1751 is sandwiched between Not2 Trp60 and Not5 Leu388 and thus is expected to affect the binding of both proteins (**Fig. 2b,e**). F1788E targets the binding to the second segment of Not5 (**Fig. 2e**). Mutants were introduced in a *not1 Δ* strain rescued by a *NOT1* gene (official symbol CDC39) on a *URA3*-marked plasmid. We recovered strains expressing only the mutant protein after counter selection for the *URA3* plasmid and scored the growth phenotypes at different temperatures. This revealed that Not1 R1811E or L1814E had little effect on cell growth at 30 °C and 37 °C, whereas strains expressing Not1 R1811E L1814E, F1751E or F1788E had a slow-growth phenotype at 30 °C that was exacerbated at 37 °C, particularly for the F1751E mutant (**Fig. 4a**). Western blot analyses demonstrated that the Not1 mutant proteins were expressed at comparable levels to those of the wild type (**Fig. 4a**). Coimmunoprecipitation experiments showed that the Not1 R1811E L1814E and F1751E mutants indeed blocked the interaction of Not1 with Not2 but maintained a normal interaction with Pop2 (**Supplementary Fig. 4a**).

Next, we engineered substitutions in Not2. The Not2 mutants L34E, M37E and the double mutant combining these substitutions target the Not1-binding site (**Fig. 2b**). The Not2 mutants F114E, L115E and combination of these substitutions target the binding to the Not5 Not box (**Fig. 3b**). We did not test the corresponding mutations in Not5 because of the lack of easily scorable phenotypes of Not5 mutants in our strain background (V.R. and B.S., unpublished observations). We introduced the Not2 mutants in yeast cells with the corresponding wild-type gene deleted and assayed the growth phenotypes of the resulting strains at various temperatures on appropriate medium. At 30 °C, the different mutations had no detectable effect, whereas growth of the double-mutant strains was severely impaired at 37 °C (**Fig. 4b**). Western blot analyses revealed that the double-mutant

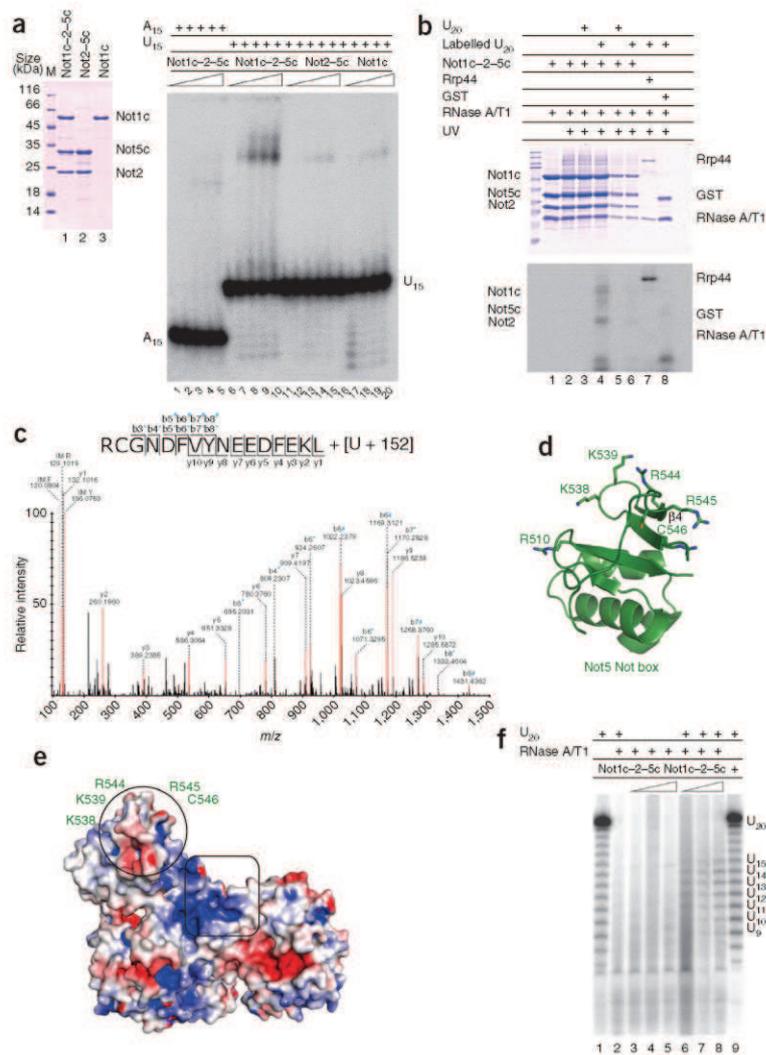
proteins were barely detectable (**Fig. 4b**). This observation that interfering with the dimerization of the Not2 Not box destabilizes Not2 is consistent with previous analyses of Not2 mutants showing protein instability with a concomitant reduction of Not5 protein level³³.

Differences between yeast and human Not3

In coimmunoprecipitation experiments, mutation on the surface of Not1 that interacts with Not2–Not5 also prevented the association with Not3 (**Supplementary Fig. 4a**), a result reinforcing the parallel between Not3 and Not5. The Not box of Not3 is predicted to have a similar fold and dimerization interface as those of Not5 or Not2. The central residues at the putative dimerization interface of the Not3 Not box are conserved, including Leu703 and Tyr710 (which are equivalent to *S. cerevisiae* Not5 Phe472 and Phe479; **Fig. 3b,c**). Using strategies described above, we constructed and evaluated yeast strains with the Not3 L703R and Y710R substitutions and a combination of both. As in the case of Not2, the single mutants had no growth phenotypes, whereas the growth of the double-mutant strain was severely impaired at 37 °C (**Fig. 4c**). Similarly as for the Not2 double mutants, low levels of the Not3 protein were present for the Not3 L703R Y710R double mutant (**Fig. 4c**). These results are consistent with the notion that Not3 is also destabilized if the Not-box domain is mutated at the putative dimerization interface.

Yeast Not3 and Not2 have been shown to associate *in vivo*^{21,47,48} (**Supplementary Fig. 4b**), although with the caveat that the interaction might be indirect. The interaction between human Not2 and Not3 Not boxes has been shown to be direct by *in vitro* assays¹⁹. To test for direct interactions of the yeast proteins *in vitro*, we engineered a fragment of yeast Not3 encompassing a minimal Not-box region (residues 685–800). In contrast to that of Not5, the Not3 Not box failed to interact with GST Not2- Δ N in pulldown assays with purified proteins (**Fig. 4d**, comparison of lane 4 with lanes 5 and 6). The Not3 Not box also did not bind on top of the GST Not2- Δ N–Not5- Δ N complex (**Fig. 4d**, lane 7). A close inspection of the amino acid sequences revealed that a subset of residues at the putative dimerization interface

Figure 5 Not1c–Not2–Not5c binds poly(U) RNA. **(a)** Left, Coomassie-stained 13.5% SDS-PAGE gel with the protein samples used for the biochemical assays. Right, EMSA with A_{15} or U_{15} RNA (50 nM) labeled at the 5' end with $[\gamma\text{-}^{32}\text{P}]$ phosphate and incubated with increasing amounts of proteins. M, molecular weight marker. **(b)** Protein-RNA cross-linking. Proteins and $\gamma\text{-}^{32}\text{P}$ body-labeled poly(U) 20-mer RNAs cross-linked under UV light and separated on 13.5% SDS PAGE are shown. The gel was stained with Coomassie blue (top) and analyzed by phosphorimaging (bottom). A/T1, mixture of RNases A and T1. **(c)** Tandem mass spectrum of Not5 residues 545–560, identifying an additional mass of 476.0338 Da corresponding to a U nucleoside with an adduct of 152 Da (associated with a cysteine). Peptide sequence and fragment ions are indicated at top. b ions with a mass shift corresponding to $\text{U-H}_3\text{PO}_4 + 152$ and to $\text{U} + 152$ are shown with an asterisk and hash mark, respectively. IM, immonium ions. **(d)** Structure of the Not5 Not box showing the position of the U-cross-linked Cys546 surrounded by a patch of positively charged residues. **(e)** Surface representation of the ternary complex colored by electrostatic potential (positive in blue and negative in red), calculated with PyMOL APBS tools. The structure is shown after a 180° rotation around a vertical axis with respect to **Figure 1b**, left. Circle, RNA-binding site; square, positively charged surface patch at the intersection of Not1, Not2 and Not5. **(f)** RNase protection assay. The protected RNA fragments obtained after treatment with a mixture of RNase A and T1 (A/T1), labeled at the 5' end with $\gamma\text{-}^{32}\text{P}$, resolved by denaturing PAGE and analyzed by phosphorimaging are shown.



is conserved between Not5 and metazoan Not3 but diverges in *S. cerevisiae* Not3 (for example, *S. cerevisiae* Not3 Thr702, Phe706, Asn711 and Ala715 in **Fig. 3c**), thus rationalizing the different behavior of the yeast Not3 protein.

The Not1–Not2–Not5 complex is a binding platform for proteins

The interaction of the Not2 and Not5 Not boxes creates a V-shaped surface (**Fig. 1b,c**). In one molecule of the asymmetric unit, the β -sheet of Not2 is extended by a loop that mediates a crystal contact with the $\beta 4$ strand of Not5 from a symmetry-related molecule (**Supplementary Fig. 3c**). Another interaction is present between the $\beta 4$ strand of Not2 and the β -hairpin of Not1 from a symmetry copy (**Supplementary Fig. 3c**). These crystal-packing contacts are somewhat reminiscent of canonical Sm–Sm interactions and point to the Not boxes as possible interaction surfaces. Genetic evidence suggests that the Not box of Not5 interacts with Not4 (ref. 24), a conserved subunit of the complex with ubiquitin-ligase activity¹⁹. The Not box of Not2 interacts with ADA2, a component of the transcription-regulatory histone-acetylation complex SAGA³³. Mutation of Not2 Arg165 has been shown to abrogate the interaction with ADA2 without affecting the integrity of the Ccr4–Not complex in yeast³³. In the structure, Arg165 protrudes on the surface of the β -barrel and is indeed accessible to solvent (**Supplementary Fig. 5a**).

Not1–Not2–Not5 is a binding platform for poly(U) RNA

Mapping of the electrostatic potential on the molecular surface of the Not1c–Not2–Not5c complex showed patches of positively charged residues. We therefore tested whether the Not module can mediate RNA binding. In electrophoretic mobility shift assays (EMSA), a single-stranded poly(U) 15-mer RNA (U_{15}) bound the Not1c–Not2–Not5c complex, whereas we detected no binding with a poly(A) 15-mer RNA (A_{15}) (**Fig. 5a**). The Not module recognized poly(U) RNA specifically, albeit with low affinity (in the low-micromolar range; **Supplementary Fig. 5b**). The EMSAs showed no binding of U_{15} RNA to either Not1c or the Not2–Not5c subcomplex in isolation (**Fig. 5a**), suggesting that the different portions of the Not module contribute together to RNA recognition. Indeed, after incubation of the Not1c–Not2–Not5c complex with a body-labeled U_{20} RNA and exposure to UV irradiation at 254 nm, all bands showed RNA cross-linking, which was strong in the case of Not1c and Not2 and less pronounced in the case of Not5c (**Fig. 5b**). In this experiment, Rrp44 and GST were positive and negative controls, respectively (**Fig. 5b**).

Next, we used MS to identify residues of the complex cross-linked to the U₂₀ RNA. This approach is based on the detection and sequencing by LC-MS/MS of peptides conjugated to the mass of an RNA nucleotide (reviewed in ref. 50). The advantage of this approach is that RNA-contact sites are determined in an unbiased manner. The caveat is that the identification is limited to sites where the ribonucleotide is in proximity to amino acids with reactive groups (for example, thiol groups in cysteine residues) and is limited by the amounts of the cross-linked species and the complexity of the spectra. The MS analysis identified a Not5 peptide corresponding to residues 545–560, with a single U nucleoside cross-linked to Cys546 (Fig. 5c). In the structure, Cys546 is positioned at the top of the Not-box β -sheet and is part of a surface patch with positively charged residues (Lys515, Arg533, Arg544 and Arg545; Fig. 5d). This RNA-binding site (circle in Fig. 5e) differs from the U nucleoside-binding site of canonical Sm folds (Supplementary Fig. 5a) and is contiguous to a positively charged surface patch at the intersection of Not1, Not2 and Not5 (square in Fig. 5e). To estimate the length of the RNA-binding path on the complex, we carried out RNase protection assays. We found that fragments of 11–15 nucleotides accumulated in the presence of Not1c–Not2–Not5c (Fig. 5f, lanes 6–8). Fragments of this size could easily span a distance of 40–60 Å.

DISCUSSION

The core of the Not module that we investigated in this work is built around the C-terminal arm of Not1 by the cooperative binding of Not2 and Not5. The C-terminal arm of Not1 has a HEAT-repeat architecture similar to that found in the N-terminal arm²⁸. It is thus possible to imagine that the two arms of Not1 might have originated from a duplication event. Not2 and Not5 interact through their C-terminal Not-box domains. At the structural level, the Not boxes resemble Sm folds. The similarity extends to their biochemical properties in terms of the ability of the Not boxes to dimerize and to bind poly(U) RNA stretches, although the interaction mechanisms have diverged from those of canonical Sm folds. The heterodimerization of the two Not boxes in the Not1–Not2–Not5 complex serves multiple purposes.

First, heterodimerization of the Not boxes tethers the N-terminal regions of Not2 and Not5, promoting their synergistic binding to Not1. Previous studies have shown that the N-terminal region of Not2 is essential for the structural integrity of the Not module because it recruits Not5 into the complex³³. We found that the N-terminal region of Not5 is equally important in recruiting Not2. The two Not boxes thus contribute indirectly to Not1 binding by bringing the two separate N-terminal regions into spatial proximity, thus probably increasing their effective local concentration. Not-box heterodimerization is also important for the stability of the individual proteins *in vivo*, as shown by mutational analysis of Not2 as well as Not3. Yeast Not3 and Not5 are currently considered homologs with partially redundant functions^{24,30}. Unexpectedly, we found that the yeast Not3 Not box has diverged from that of Not5 and does not interact with the Not2 Not box *in vitro*. The Not2–Not5 dimerization interface is instead conserved in Not2–Not3 from higher eukaryotes, thus suggesting that the protein that we currently refer to as metazoan Not3 is an ortholog of *S. cerevisiae* Not5. The identity of the direct interactions that mediate the recruitment of *S. cerevisiae* Not3 Not box in the complex is currently unclear and is an important question for future studies.

Second, the Not boxes together with Not1 form a composite platform for macromolecular interactions. Extensive data indicate that the Not module is closely connected to the transcriptional machinery

and physically recruits transcription factors, such as ADA2 (ref. 33; reviewed in refs. 11,12). Evidence is also accumulating on the ability of the Not modules to mediate protein-protein interactions important for cytoplasmic mRNA metabolism. For example, in *Drosophila melanogaster* Not3 binds the translational repressor BicC⁵¹, and in mice the C-terminal arm of Not1 binds the mRNA developmental regulator NANOS2 (ref. 52). We found that the Not module creates a composite RNA-binding surface for U nucleosides, with a specific site in the Not box of Not5. Although this RNA-binding activity of the Not module was unexpected, it rationalizes previous observations. In yeast, the decay of the Edc1 mRNA has been shown to proceed through a deadenylation-independent decapping pathway that depends on the Not proteins and on a poly-U tract in its 3' UTR³⁹. A model is conceivable in which binding of the Not module to this 3' UTR might bring the mRNA into proximity of Dhh1, a decapping activator (known as DDX6 or RCK in metazoans) that is recruited to Ccr4–Not^{27,53,54}. Interestingly, mouse Not3 has been shown to regulate the deadenylation of specific mRNAs by recruiting their 3' UTRs³², which also contain U-rich stretches. The emerging picture is that the Not module of the Ccr4–Not complex creates a platform for protein and nucleic acid interactions that is able to contribute to the many functions of the Ccr4–Not complex, including the degradation of specific mRNAs.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Coordinates and structural factors have been deposited in the Protein Data Bank under accession code [4BY6](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the MPI Biochemistry Crystallization Facility and Core Facility. We thank F. Bonneau and C. Basquin (MPI Biochemistry) for help with biochemical assays and for [Supplementary Figure 5b](#); F. Lacroute, F. Gabriel and M.C. Daugeron (Centre de Génétique Moléculaire) for yeast strains; the staff of the Swiss Light Source synchrotron for assistance during data collection; and members of our laboratories for discussions and for critical reading of the manuscript. E.C. acknowledges support from the Max Planck Gesellschaft, the European Research Council (ERC Advanced Investigator Grant 294371, Marie Curie Initial Training Network RNpnet 289007) and the Deutsche Forschungsgemeinschaft (DFG SFB646, SFB1035, GRK1721, FOR1680, CIPSM). B.S. acknowledges support from the Centre Européen de Recherche en Biologie et en Médecine (CERBM)-IGBMC, the CNRS and the Ligue Contre le Cancer (Equipe Labellisée 2011).

AUTHOR CONTRIBUTIONS

V.B. and J.B. carried out the structure determination and the *in vitro* experiments; V.R. carried out the *in vivo* experiments; K.S. and H.U. carried out the MS analysis; E.C. and B.S. supervised the project; and E.C., V.B. and B.S. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Wahle, E. & Winkler, G.S. RNA decay machines: deadenylation by the Ccr4-Not and Pan2-Pan3 complexes. *Biochim. Biophys. Acta* **1829**, 561–570 (2013).
2. Tucker, M. *et al.* The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell* **104**, 377–386 (2001).
3. Daugeron, M.C., Mauxion, F. & Séraphin, B. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res.* **29**, 2448–2455 (2001).

4. Sandler, H., Kreth, J., Timmers, H.T.M. & Stoeklin, G. Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res.* **39**, 4373–4386 (2011).
5. Fabian, M.R. *et al.* Structural basis for the recruitment of the human CCR4–NOT deadenylase complex by tristetraprolin. *Nat. Struct. Mol. Biol.* **20**, 735–739 (2013).
6. Braun, J.E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44**, 120–133 (2011).
7. Chekulaeva, M. *et al.* miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226 (2011).
8. Fabian, M.R. *et al.* miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217 (2011).
9. Goldstrohm, A.C., Hook, B.A., Seay, D.J. & Wickens, M. PUF proteins bind Pop2p to regulate messenger RNAs. *Nat. Struct. Mol. Biol.* **13**, 533–539 (2006).
10. Suzuki, A., Igarashi, K., Aisaki, K.-I., Kanno, J. & Saga, Y. NANOS2 interacts with the CCR4–NOT deadenylation complex and leads to suppression of specific RNAs. *Proc. Natl. Acad. Sci. USA* **107**, 3594–3599 (2010).
11. Collart, M.A. & Panasenko, O.O. The Ccr4–Not complex. *Gene* **492**, 42–53 (2012).
12. Miller, J.E. & Reese, J.C. Ccr4–Not complex: the control freak of eukaryotic cells. *Crit. Rev. Biochem. Mol. Biol.* **47**, 315–333 (2012).
13. Sun, M. *et al.* Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res.* **22**, 1350–1359 (2012).
14. Liu, H.Y. *et al.* The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *EMBO J.* **17**, 1096–1106 (1998).
15. Chen, J. *et al.* Purification and characterization of the 1.0 MDa CCR4–NOT complex identifies two novel components of the complex. *J. Mol. Biol.* **314**, 683–694 (2001).
16. Albert, T.K. *et al.* Isolation and characterization of human orthologs of yeast CCR4–NOT complex subunits. *Nucleic Acids Res.* **28**, 809–817 (2000).
17. Lau, N.C. *et al.* Human Ccr4–Not complexes contain variable deadenylase subunits. *Biochem. J.* **422**, 443–453 (2009).
18. Temme, C. *et al.* Subunits of the *Drosophila* CCR4–NOT complex and their roles in mRNA deadenylation. *RNA* **16**, 1356–1370 (2010).
19. Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S. & Izaurralde, E. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4–NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol.* **10**, 228–244 (2013).
20. Schwede, A. *et al.* A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res.* **36**, 3374–3388 (2008).
21. Bai, Y. *et al.* The CCR4 and CAF1 proteins of the CCR4–NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Mol. Cell Biol.* **19**, 6642–6651 (1999).
22. Mauxion, F., Prève, B. & Seraphin, B. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4–NOT complex. *RNA Biol.* **10**, 267–276 (2013).
23. Färber, V., Erben, E., Sharma, S., Stoeklin, G. & Clayton, C. Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs. *Nucleic Acids Res.* **41**, 1211–1222 (2013).
24. Oberholzer, U. & Collart, M.A. Characterization of NOT5 that encodes a new component of the Not protein complex. *Gene* **207**, 61–69 (1998).
25. Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O. & Collart, M.A. The essential function of Not1 lies within the Ccr4–Not complex. *J. Mol. Biol.* **303**, 131–143 (2000).
26. Draper, M.P., Liu, H.Y., Nelsbach, A.H., Mosley, S.P. & Denis, C.L. CCR4 is a glucose-regulated transcription factor whose leucine-rich repeat binds several proteins important for placing CCR4 in its proper promoter context. *Mol. Cell Biol.* **14**, 4522–4531 (1994).
27. Maillet, L. & Collart, M.A. Interaction between Not1p, a component of the Ccr4–Not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. *J. Biol. Chem.* **277**, 2835–2842 (2002).
28. Basquin, J. *et al.* Architecture of the nuclease module of the yeast Ccr4–Not complex: the Not1–Caf1–Ccr4 interaction. *Mol. Cell* **48**, 207–218 (2012).
29. Petit, A.-P. *et al.* The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4–NOT deadenylase complex. *Nucleic Acids Res.* **40**, 11058–11072 (2012).
30. Collart, M.A., Panasenko, O.O. & Nikolaev, S.I. The Not3/5 subunit of the Ccr4–Not complex: a central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cell Signal.* **25**, 743–751 (2013).
31. Neely, G.G. *et al.* A global *in vivo* *Drosophila* RNAi screen identifies NOT3 as a conserved regulator of heart function. *Cell* **141**, 142–153 (2010).
32. Morita, M. *et al.* Obesity resistance and increased hepatic expression of catabolism-related mRNAs in *Cnot3*^{+/−} mice. *EMBO J.* **30**, 4678–4691 (2011).
33. Russell, P., Benson, J.D. & Denis, C.L. Characterization of mutations in NOT2 indicates that it plays an important role in maintaining the integrity of the CCR4–NOT complex. *J. Mol. Biol.* **322**, 27–39 (2002).
34. Ito, K. *et al.* CNOT2 depletion disrupts and inhibits the CCR4–NOT deadenylase complex and induces apoptotic cell death. *Genes Cells* **16**, 368–379 (2011).
35. Zwartjes, C.G.M., Jayne, S., van den Berg, D.L.C. & Timmers, H.T.M. Repression of promoter activity by CNOT2, a subunit of the transcription regulatory Ccr4–Not complex. *J. Biol. Chem.* **279**, 10848–10854 (2004).
36. Badarinarayana, V., Chiang, Y.C. & Denis, C.L. Functional interaction of CCR4–NOT proteins with TATAA-binding protein (TBP) and its associated factors in yeast. *Genetics* **155**, 1045–1054 (2000).
37. Lemaire, M. & Collart, M.A. The TATA-binding protein-associated factor yTafII19p functionally interacts with components of the global transcriptional regulator Ccr4–Not complex and physically interacts with the Not5 subunit. *J. Biol. Chem.* **275**, 26925–26934 (2000).
38. Deluen, C. *et al.* The Ccr4–not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions. *Mol. Cell Biol.* **22**, 6735–6749 (2002).
39. Muhrad, D. & Parker, R. The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. *EMBO J.* **24**, 1033–1045 (2005).
40. Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W. & Bork, P. Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* **309**, 1–18 (2001).
41. Marcotrigiano, J. *et al.* A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol. Cell* **7**, 193–203 (2001).
42. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
43. Kambach, C. *et al.* Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* **96**, 375–387 (1999).
44. Toró, I. *et al.* RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J.* **20**, 2293–2303 (2001).
45. Khushf, P., Plaag, R. & Zieve, G.W. LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem. Sci.* **30**, 522–528 (2005).
46. Leung, A.K.W., Nagai, K. & Li, J. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473**, 536–539 (2011).
47. Azzouz, N. *et al.* Specific roles for the Ccr4–Not complex subunits in expression of the genome. *RNA* **15**, 377–383 (2009).
48. Tarassov, K. *et al.* An *in vivo* map of the yeast protein interactome. *Science* **320**, 1465–1470 (2008).
49. Albert, T.K. *et al.* Identification of a ubiquitin–protein ligase subunit within the CCR4–NOT transcription repressor complex. *EMBO J.* **21**, 355–364 (2002).
50. Schmidt, C., Kramer, K. & Urlaub, H. Investigation of protein–RNA interactions by mass spectrometry: techniques and applications. *J. Proteomics* **75**, 3478–3494 (2012).
51. Chicoine, J. *et al.* Bicaudal-C recruits CCR4–NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. *Dev. Cell* **13**, 691–704 (2007).
52. Suzuki, A., Saba, R., Miyoshi, K., Morita, Y. & Saga, Y. Interaction between NANOS2 and the CCR4–NOT deadenylation complex is essential for male germ cell development in mouse. *PLoS ONE* **7**, e33558 (2012).
53. Hata, H. *et al.* Dhh1p, a putative RNA helicase, associates with the general transcription factors Pop2p and Ccr4p from *Saccharomyces cerevisiae*. *Genetics* **148**, 571–579 (1998).
54. Collier, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A. & Parker, R. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* **7**, 1717–1727 (2001).

ONLINE METHODS

Protein purification. All the proteins were cloned and expressed individually in *E. coli* BL21(DE3) pLysS cells (Stratagene) in TB medium with IPTG induction overnight at 18 °C. Not1_{1541–2093} (Not1c), Not5 FL and Not5_{298–560} (Not5c) were expressed with an N-terminal His-SUMO tag (cleavable with the Snp2 protease). The Not1 C-terminal arm (starting at 1348), Not2 FL and Not5_{460–560} (Not5 Not box) were expressed with an N-terminal His-Z tag (cleavable with TEV protease). Not3_{685–800} (Not3 Not box) was expressed with an N-terminal His tag (cleavable with TEV protease). The cells were lysed in buffer A (50 mM sodium phosphate buffer, pH 7.5, 250 mM NaCl and 20 mM imidazole) by sonication. The lysates were cleared by centrifugation and were loaded on a 5-ml His-trap column (GE Healthcare). The column was washed with buffer B (50 mM phosphate buffer, pH 6.5, 1 M NaCl, 20 mM imidazole, 50 mM KCl, 10 mM MgSO₄ and 2 mM ATP) and with buffer A. The proteins were eluted by a gradient of buffer A and buffer C (buffer A supplemented with 500 mM imidazole). Except for the Not5 Not box and Not2, all other proteins were dialyzed overnight in gel-filtration buffer (without DTT) in the presence of TEV or Snp2 proteases and were then applied to the His-trap column to remove the cleaved tag (second affinity step). The proteins were further purified by size-exclusion chromatography in the gel-filtration buffer (20 mM Tris-Cl buffer, pH 7.5, 250 mM NaCl and 2 mM DTT). The complex of Not1c–Not2–Not5c was formed by mixing the purified protein in a 1:1.25:1 molar ratio and was incubated with TEV protease overnight at 4 °C to cleave the N-terminal His-Z tag of Not2. The complex was applied onto the 5-ml His-trap column (GE Healthcare) to remove the cleaved tag and was purified by gel filtration (Superdex 200 16/60, GE Healthcare) in the gel-filtration buffer.

For the pulldown assays, Not5c, Not5-ΔN and Not2-ΔN were expressed as N-terminal His-GST fusion proteins, whose tags were cleavable with 3C protease. Not2, Not5c and Not5-ΔN were affinity purified with a 5-ml His-trap column (GE Healthcare) as described above. Not2-ΔN was affinity purified at a pH of 8.5 (with buffer A and C at pH 8.5) instead of pH 7.5. Not2–Not5c, Not2-ΔN–Not5c and Not2–Not5-ΔN complexes were formed by mixing a 1:1.5 molar ratio of the larger to smaller protein and dialyzed in gel-filtration buffer (without 2 mM DTT) in the presence of 3C protease and TEV protease. The dialyzed proteins were subjected to a second His-affinity purification with a 5-ml His-trap column (GE Healthcare) and subsequent incubation with glutathione-agarose beads (Protino) for 2 h at 4 °C to remove the GST-tag contamination. The proteins were then purified by gel filtration (Superdex 75 10/30, GE Healthcare) in gel-filtration buffer. Not1c-GST was affinity purified as described above. The protein was dialyzed against heparin buffer A (20 mM Tris, pH 7.5, and 100 mM NaCl), applied onto the 5-ml heparin column (GE Healthcare) and purified with a gradient elution with heparin buffer A and heparin buffer A supplemented with 1 M NaCl. Not1c-GST was further purified by gel filtration (Superdex 200 10/30, GE Healthcare). The Not3 Not box was expressed as TEV protease-cleavable His₆ fusion protein and purified in a similar way as mentioned above.

For the RNA-binding experiments, Not2 and Not5c were expressed and purified as individual proteins as described above. The Not2–Not5c complex was formed by mixing the proteins in a 1.25:1 molar ratio and subsequent overnight incubation with TEV protease. The cleaved protein was subjected to His-affinity purification to remove the cleaved tag and a subsequent heparin-column purification. The complex was further purified by gel filtration (Superdex 200 16/60, GE Healthcare). Not1c was expressed and purified as above, with an additional step of heparin purification included after the second His-affinity step. Not1c–Not2–Not5c was purified by mixing Not1c and Not2–Not5c in a 1:1.25 molar ratio and subsequent gel filtration (Superdex 200 16/60, GE Healthcare).

Limited proteolysis experiment. 0.6 mg/ml of the Not1Δ1347–Not2–Not5 complex was incubated with elastase (Roche) at a 1:10 (w/w) enzyme/protein ratio for 30 min on ice. The products of the proteolysis were identified by N-terminal sequencing and MS analysis. The interacting core complex was identified by size-exclusion chromatography of the proteolyzed sample.

Crystallization and structure solution. The Not1c–Not2–Not5c complex was concentrated to 16 mg/ml and crystallized at room temperature in 8.5% (w/v) PEG 8000, 100 mM MES, pH 6.5, and 200 mM calcium acetate. The mercury derivative was prepared by cocrystallization of a solution of Not1c–Not2–Not5c with ethyl mercury phosphate (EMP) at 0.55 mM final concentration. The crystals were frozen in the presence of 20% glycerol as cryoprotectant. X-ray data were collected at 100 K

at the SLS synchrotron (PXII and PXIII beamlines), with tuning of the wavelength at the Hg edge in the case of the EMP-containing crystals for SAD data collection. The data were processed with XDS⁵⁵. The crystals belong to a monoclinic space group (P₂) with two molecules per asymmetric unit. We used PHENIX.autosol⁵⁶ for phasing and Buccaneer⁵⁷ for the initial automatic model building. We completed the model with iterative rounds of manual model building with Coot⁵⁸ and restrained refinement with PHENIX⁵⁶. The final model has 97.3% residues in the most-favored regions of the Ramachandran plot, as calculated with MolProbity⁵⁹.

Pulldown assays. For the experiments in Figure 2f, 50 pmol of bait (GST or GST-Not1c) were incubated with 100 pmol of prey (Not2–Not5c, Not2-ΔN–Not5c and Not2–Not5-ΔN) for 1 h at 4 °C in 40 mM Tris, pH 7.5, 150 mM NaCl, 2 mM DTT, 12.5% (v/v) glycerol and 0.1% (w/v) NP-40 (binding buffer). The protein mix was incubated with 20 μl of GSH-agarose beads (Protino) for 1 h with gentle rocking at 4 °C. The resin was washed three times with the binding buffer, and the proteins were eluted in 15 μl of binding buffer containing 100 mM glutathione. Input and precipitates were mixed with SDS loading dye, resolved on 4–12% Bis-Tris NuPAGE gel (Invitrogen) with MES as running buffer, and visualized by Coomassie-blue staining. A similar protocol was used for the GST pulldown assays in Figure 4d with 40 mM Tris, pH 8.5, 100 mM NaCl, 2 mM DTT and 12.5% (v/v) glycerol as binding buffer.

Sequence alignments and superpositions. All the sequence alignments were done with ClustalW⁶⁰ and ALINE⁶¹, and the structural superpositions were done with SSM in Coot⁵⁸. The r.m.s. deviations reported are from the output of Coot. Structure-based sequence alignment was done in STRAP⁶² with the Aligner3D method and manually edited in ALINE.

Yeast strains. Yeast strains used in this study are all derivatives of W303 (*ade2-1, can1-100, leu2-3,112, his3-11,15, ura3-1, trp1-1*). Genes differing from W303 are as follows: T26N28 (*MATa, Δtrp1, ΔNOT1::HIS3 pFL 38 (NOT1)*), BSY1110 (*MATa, Δtrp1, not2::HISMx6*), BSY1111 (*MATa, Δtrp1, not3::HISMx6*), BSY1230 (*MATa, Δtrp1, NOT3-VSV, NOT2-3HA::hisMX6*), BSY1231 (*MATa, Δtrp1, POP2-VSV, NOT2-3HA::hisMX6*), BSY1240 (*MATa, Δtrp1, NOT3-TAP::TRP1_{K6}, POP2-VSV, NOT4-HA::hisMX6*) and BSY1242 (*MATa, Δtrp1, NOT3-TAP::TRP1_{K6}, POP2-VSV, NOT2-HA::hisMX6*).

Coprecipitation assays. Protein extract preparation and coimmunoprecipitation were performed as described previously²⁸.

Mutant analyses. Mutations in Not1, Not2 and Not3 were constructed in plasmids pBS4806 (ref. 28) (Not1-TAP), pBS4968 (Not2-VSV) and pBS4975 (Not3-VSV) by one or multiple rounds of site-directed mutagenesis. The presence of the desired mutation was ascertained by sequencing. The resulting plasmids were introduced into yeast strains with the lithium acetate transformation procedure. Plasmid shuffling, growth assays, protein extraction and western blot analyses were performed with standard procedures as previously described²⁸.

Electrophoretic mobility shift assays. Proteins at 3, 10, 13 and 20 μM concentration (30, 100, 130 and 200 pmol, respectively) were incubated with 50 nM (0.5 pmol) of 5'-labeled RNA (A₁₅ or U₁₅) at 4 °C overnight in 20 mM HEPES, pH 7.5, 5 mM EDTA, 0.1% NP-40 and 2 mM DTT (EMSA buffer). The reaction mixtures were complemented with gel-filtration buffer to a final NaCl concentration of 54 mM, resolved on a 6% (w/v) native PAGE and visualized by phosphorimaging.

RNA cross-linking. 200 pmol (20 μM) of Not1–Not2–Not5c complex were incubated with 2.5 pmol (250 nM) of body-labeled U₂₀ RNA overnight in EMSA buffer at 4 °C. The cross-linking was performed by irradiation of the mix at a wavelength of 254 nm for 5 min on ice. The mixture was then treated with 1% SDS and 0.5 μl of RNase A/T1 mixture at 37 °C for 5 min. The samples were heated with SDS loading dye at 70 °C for 2 min, separated on 13.5% (w/v) SDS-PAGE gel and visualized by phosphorimaging and Coomassie-blue staining.

Mass spectrometry. UV-induced protein-RNA cross-linking and enrichment of cross-linked peptides. UV cross-linking and enrichment of cross-linked peptides was performed according to the established protocols described in ref. 63. Briefly, 1 nmol of the single-stranded U₁₅ RNA oligonucleotide and 1 nmol of

Not1c–Not2–Not5c complex were mixed in a 1:1 molar ratio, and the total reaction volume was brought to 100 μ l in 20 mM HEPES, pH 7.5, 50 mM NaCl, 2 mM DTT and 5 mM EDTA. The mixture was incubated on ice overnight for complex formation. The samples were then transferred to black polypropylene microplates (Greiner Bio-One) and irradiated at 254 nm for 10 min. After ethanol precipitation, the samples were denatured in 4 M urea and 50 mM Tris-HCl, pH 7.9, and digested for 2 h at 52 °C with 1 μ g RNase A (Ambion, Applied Biosystems). After RNA digestion, proteolysis with trypsin (Promega) was performed overnight at 37 °C. The sample was desalted on an in-house-prepared C18 (Dr. Maisch GmbH) column, and the cross-linked peptides were enriched on an in-house-prepared TiO₂ (GL Sciences) column with the protocol described in ref. 63. The samples were dried and then resuspended in 10 μ l sample solvent (5% v/v ACN and 1% v/v FA) for MS analysis.

Nano-liquid chromatography and MS analysis. 5 μ l of the above sample was injected onto a nano-liquid chromatography system (Agilent 1100 series, Agilent Technologies) including a C18 trapping column of length ~2 cm and inner diameter 150 μ m, in line with a C18 analytical column of length ~15 cm and inner diameter 75 μ m (both packed in house; C18 AQ 120 Å 5 μ m, Dr. Maisch GmbH). Analytes were loaded on the trapping column at a flow rate of 10 μ l/min in buffer A (0.1% v/v FA) and subsequently eluted and separated on the analytical column with a gradient of 7–38% buffer B (95% v/v acetonitrile and 0.1% v/v FA) with an elution time of 33 min (0.87%/min) and a flow rate of 300 nL/min. Online ESI-MS was performed with an LTQ-Orbitrap Velos instrument (Thermo Scientific), operated in data-dependent mode with a TOP10 method. MS scans were recorded in the m/z range of 350–1,600 and for subsequent MS/MS the top ten most-intense ions were selected. Both precursor ions as well as fragment ions were scanned in the Orbitrap. Fragment ions were generated by higher-energy collision dissociation (HCD) activation (normalized collision energy = 40) and recorded from $m/z = 100$. As precursor ions as well as fragment ions were scanned in the Orbitrap, the resulting spectra were measured with high accuracy (<5 p.p.m.), both in the MS and MS/MS level.

Data analysis. The MS raw files were converted into the .mzML format with msconvert⁶⁴. Protein-RNA cross-links were analyzed with OpenMS^{65,66} and OMSSA⁶⁷ as a search engine. Data-analysis workflows were assembled as described¹¹. The high-scoring cross-linked peptides were manually annotated for confirmation. Protein-RNA interactions between the complex and poly(U) RNA were analyzed with UV-induced protein-RNA cross-linking followed by MS. Peptide RCGNDFVYNEEDFEKL in Not5 (position 545–560) was observed carrying an additional mass of 476.0338 Da corresponding to U nucleoside with an adduct of 152. The γ -ion series could be observed from 1 to 10, unshifted. In contrast, b ions from b3 until b8 were observed with a mass shift corresponding to U-H₃PO₄ and 152 adduct (Fig. 5c). Also, the b ions from b5 until b8 were observed with a mass shift corresponding to U and 152. We have always observed that the 152 adduct is observed as a shift associated with cysteine, which could be the amino acid that is cross-linked. In the corresponding figure (Fig. 5c), the b ions that were observed with a mass shift corresponding to U-H₃PO₄ + 152 and to U + 152 are shown with an asterisk (*) and hash (#), respectively and the immonium ions with IM.

RNase protection assays. 100, 150 and 200 pmol (10, 15 and 20 μ M) of Not1c–Not2–Not5c complex were incubated with 0.5 pmol (50 nM) of U₂₀ RNA in the EMSA buffer overnight at 4 °C. The reaction mixtures were treated with 0.5 μ l of RNase A/T1 mix for 30 min at 4 °C. RNA was purified with phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. The purified RNA was 5' labeled with [γ -³²P]ATP with T4 polynucleotide kinase, repurified by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation, separated on 22% (w/v) denaturing PAGE with 5 M urea and visualized by phosphorimaging.

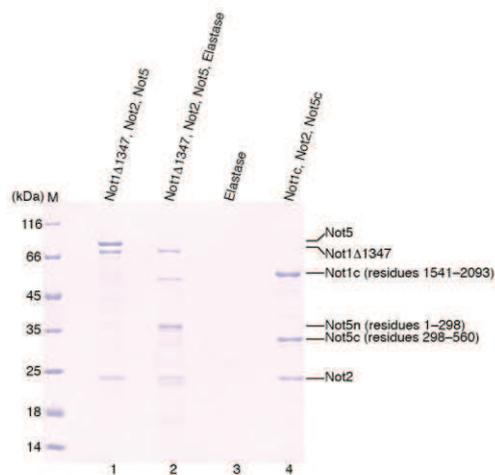
Fluorescence anisotropy. 5'-6-carboxy-fluorescein (6-FAM)-labeled U₁₅ RNA was used in fluorescence anisotropy measurements at 20 °C with Genios Pro (Tecan). RNA was 9.1 nM at final concentration and was incubated with varying concentrations of Not1c–Not2–Not5c complex in the gel-filtration buffer supplemented with 10 mM EDTA. We used 250 mM NaCl in the buffer for the measurement because the protein was not stable at 100 or 150 mM salt in such high concentration at 20 °C. The excitation and emission wavelengths were 485 nm and 535 nm, respectively. Each titration point was measured three times with ten reads with an integration time of 40 μ sec. The data were analyzed by nonlinear regression fitting with Origin (OriginLab; <http://www.originlab.com/>).

55. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 133–144 (2010).
56. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
57. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
58. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
59. Davis, I.W., Murray, L.W., Richardson, J.S. & Richardson, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619 (2004).
60. Larkin, M.A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
61. Bond, C.S. & Schüttelkopf, A.W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 510–512 (2009).
62. Gille, C. & Frömmel, C. STRAP: editor for structural alignments of proteins. *Bioinformatics* **17**, 377–378 (2001).
63. Kramer, K. *et al.* Mass-spectrometric analysis of proteins cross-linked to 4-thio-uracil- and 5-bromo-uracil-substituted RNA. *Int. J. Mass Spectrom.* **304**, 184–194 (2011).
64. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536 (2008).
65. Sturm, M. *et al.* OpenMS: an open-source software framework for mass spectrometry. *BMC Bioinformatics* **9**, 163 (2008).
66. Bertsch, A., Gröpl, C., Reinert, K. & Kohlbacher, O. OpenMS and TOPP: open source software for LC-MS data analysis. *Methods Mol. Biol.* **696**, 353–367 (2011).
67. Geer, L.Y. *et al.* Open mass spectrometry search algorithm. *J. Proteome Res.* **3**, 958–964 (2004).

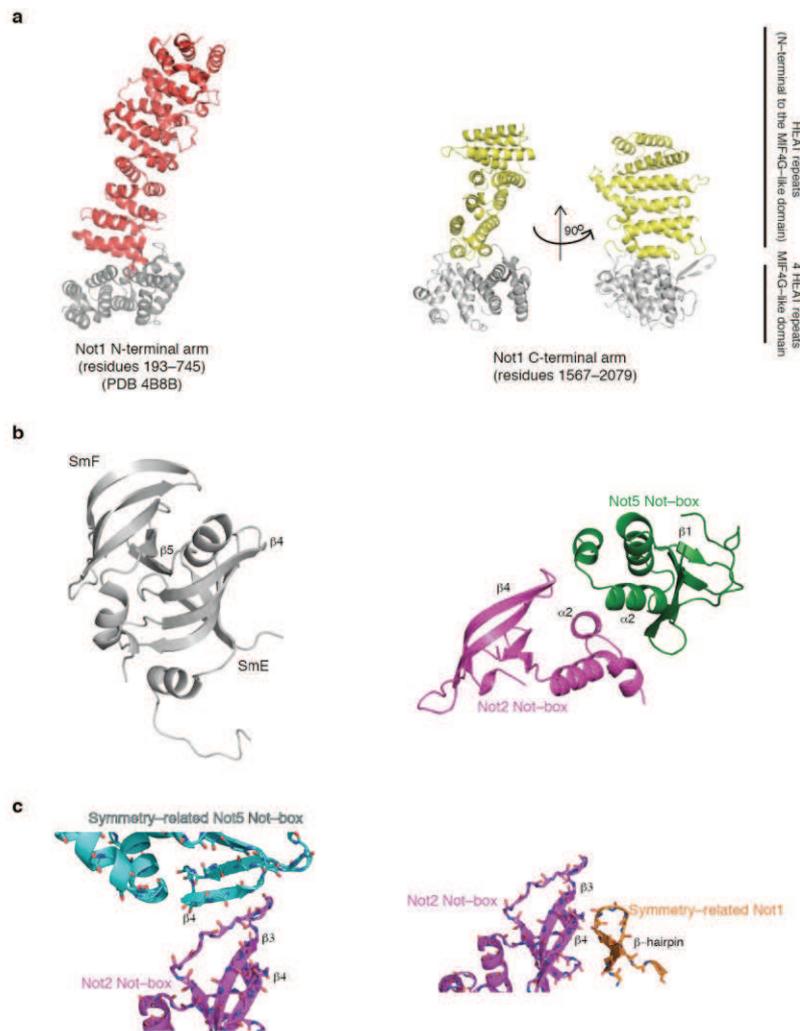
**Structure and RNA-binding properties
of the Not1–Not2–Not5 module of the yeast Ccr4–Not complex**

Varun Bhaskar¹, Vladimir Roudko^{2,3}, Jerome Basquin¹, Kundan Sharma⁴, Henning Urlaub⁴,
Bertrand Seraphin^{2,3} and Elena Conti^{1*}

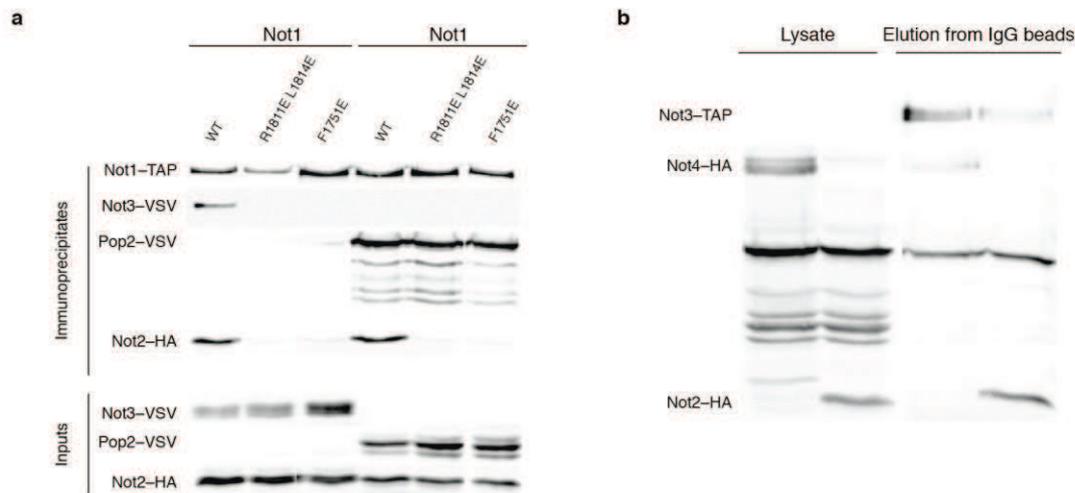
Supplementary Information



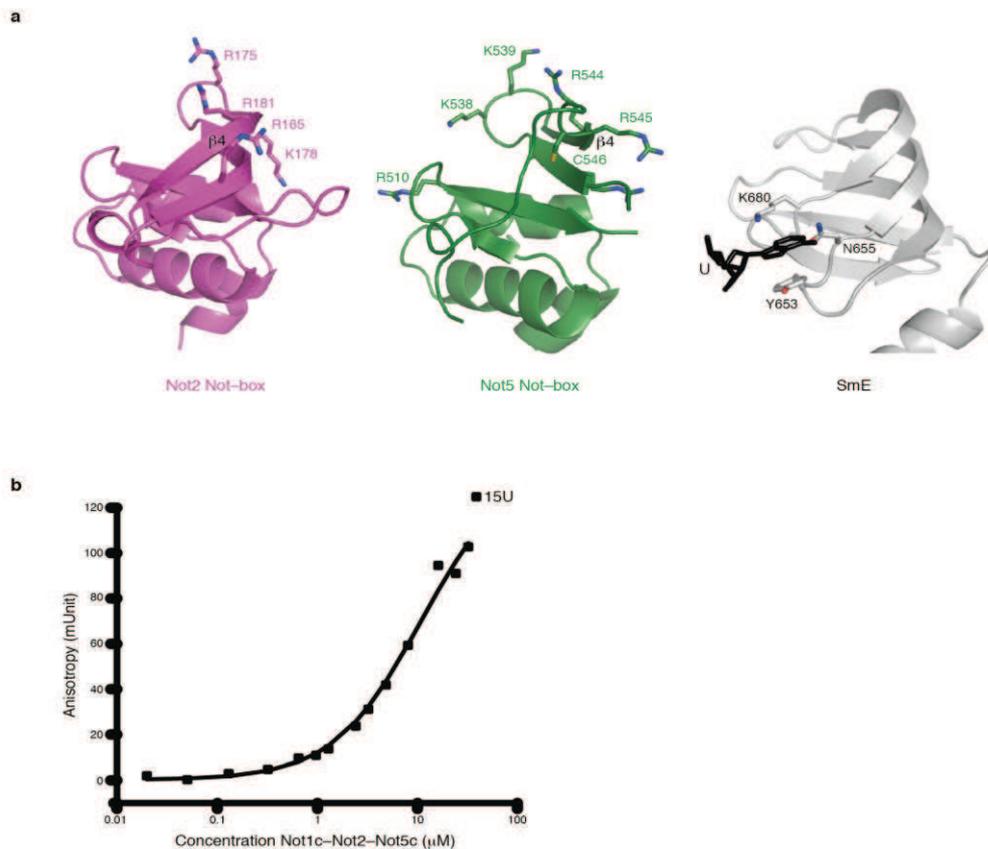
Supplementary Figure 1 Identification of the core of the *S. cerevisiae* Not1–Not2–Not5 interaction. The 12% SDS PAGE gel shows in lane 1 the larger complex that we initially purified (Not1 starting at residue 1348, Not2 f.l. and Not5 f.l.). Lane 2 shows the result of the limited proteolysis of the complex in lane 1 with elastase. Lane 3 shows the protease alone as a control. Lane 4 shows the complex reconstituted with the minimal interacting regions of Not1, Not2 and Not5 that yielded diffracting crystals (Not1c, Not2 and Not5c).



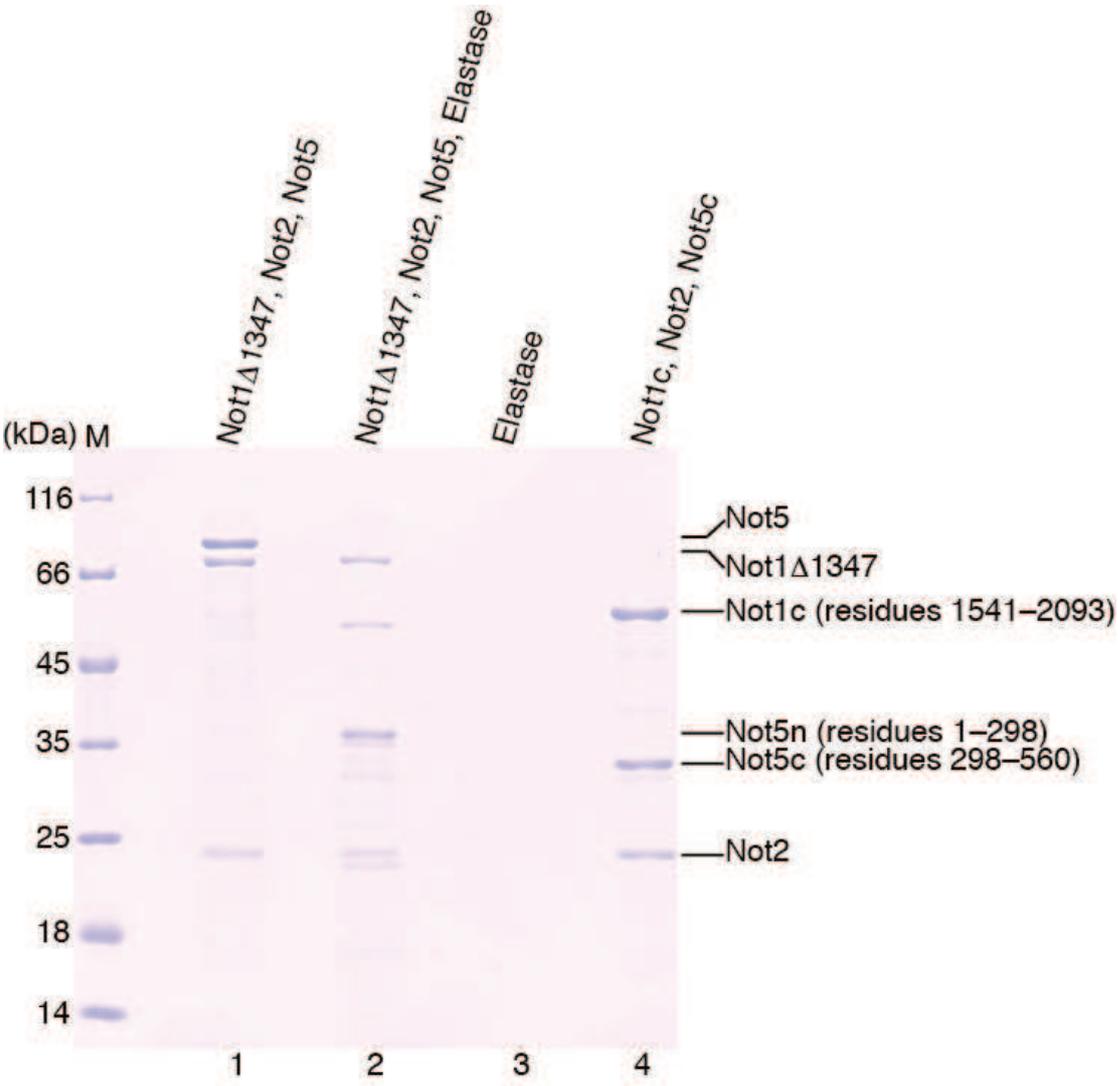
Supplementary Figure 3 The HEAT and Sm folds of Not1–Not2–Not5. **(a)** Comparison of the HEAT-repeat architecture in the C-terminal arm of Not1 (on the right) and the N-terminal arm of Not1 (on the left, PDB code 4B8B¹). The MIF4G-like folds are shown in gray. The longer HEAT-repeat units perpendicular to the MIF4G-like folds are shown in yellow and red for the C-terminal and N-terminal arms, respectively. **(b)** Dimerization properties of Not-box domains. The subcomplex of SmF and SmE (PDB code 2Y9A²) is shown on the left in gray. The β 4 strand of one monomer (SmE) interacts with strand β 5 of the other monomer (SmF). On the right, dimerization of Not2–Not5 leaves strand β 4 exposed to solvent. **(c)** Lattice contacts are reminiscent of Sm–Sm interactions. In the upper panel, the loop between strands β 2 and β 3 of a Not2 molecule (in magenta) has an extended conformation and interacts both with the β 4 strand of a symmetry-related Not5 molecule (in cyan). In the lower panel, the β 4 strand of a Not2 molecule (in magenta) interacts with the β -hairpin of a symmetry-related Not1 molecule (in orange).



Supplementary Figure 4 *In vivo* interactions of Not proteins **(a)** Immunoprecipitation of Not1 from yeast strains harbouring a TAP tagged Not1 wild-type (WT) or indicated mutants and carrying tagged chromosomal variant of Not2 and Not3 (BSY1230), or Not2 and Pop2/Caf1 (BSY1231). Coimmunoprecipitation of Not2-HA, Not3-VSV or Pop2-VSV was assayed by western blotting. As a control, the presence of the tagged protein in the starting extracts was also assayed. **(b)** Immunoprecipitation of Not3 from yeast strains carrying tagged chromosomal variant of Not3, Pop2/Caf1 and Not4 (BSY1240), or Not3, Caf1/Pop2 and Not2 (BSY1242). Coimmunoprecipitation of Not2-HA or Not4-HA was assayed by western blotting. As a control, the level of the tagged protein in the starting extracts was also assayed.



Supplementary Figure 5 Protein and RNA interactions at the Not-boxes **(a)** Surface features of the Not2 and Not5 Not-boxes. On the left is the structure of Not2, showing Arg165 (putative ADA2-binding residue) as well as positively charged residues at a similar position as in Not5. In the central panel is the Not-box of Not5, in the same orientation, showing the uridine-crosslinked residue Cys546 as well as the surrounding positively charged residues (as in **Fig. 5d**). On the right is the structure of RNA-bound SmE (U4 snRNP, PDB code 2Y9A²) oriented in a similar view as the structures in the left and central panels (after optimal superimposition), showing a bound uridine nucleotide (in black). **(b)** Quantification of the RNA-binding properties of Not1c-Not2-Not5c by fluorescence anisotropy. The K_d of the Not1c-Not2-Not5c complex to 6-FAM-labeled U₁₅ RNA under these conditions was found to be 9.47 ± 0.95 μ M.



Reference

- 1 Basquin, J. *et al.* Architecture of the Nuclease Module of the Yeast Ccr4–Not Complex: the Not1–Caf1–Ccr4 Interaction. *Mol Cell* **48**, 207–218 (2012).
- 2 Leung, A. K. W., Nagai, K. & Li, J. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473**, 536–539 (2011).

Methods

Protein purification.

All the proteins were cloned and expressed individually in *E. coli* BL21(DE3) pLysS cells (Stratagene) in TB medium with IPTG induction overnight at 18 °C. Not11541–2093 (Not1c), Not5 FL and Not5298–560 (Not5c) were expressed with an N-terminal His-SUMO tag (cleavable with the Senp2 protease). The Not1 C-terminal arm (starting at 1348), Not2 FL and Not5460–560 (Not5 Not box) were expressed with an N-terminal His-Z tag (cleavable with TEV protease). Not3685–800 (Not3 Not box) was expressed with an N-terminal His tag (cleavable with TEV protease). The cells were lysed in buffer A (50 mM sodium phosphate buffer, pH 7.5, 250 mM NaCl and 20 mM imidazole) by sonication. The lysates were cleared by centrifugation and were loaded on a 5-ml His-trap column (GE Healthcare). The column was washed with buffer B (50 mM phosphate buffer, pH 6.5, 1 M NaCl, 20 mM imidazole, 50 mM KCl, 10 mM MgSO₄ and 2 mM ATP) and with buffer A. The proteins were eluted by a gradient of buffer A and buffer C (buffer A supplemented with 500 mM imidazole). Except for the Not5 Not box and Not2, all other proteins were dialyzed overnight in gel-filtration buffer (without DTT) in the presence of TEV or Senp2 proteases and were then applied to the His-trap column to remove the cleaved tag (second affinity step). The proteins were further purified by size-exclusion chromatography in the gel-filtration buffer (20 mM Tris-Cl buffer, pH 7.5, 250 mM NaCl and 2 mM DTT). The complex of Not1c–Not2–Not5c was formed by mixing the purified protein in a 1:1.25:1 molar ratio and was incubated with TEV protease overnight at 4 °C to cleave the N-terminal His-Z tag of Not2. The complex was applied onto the 5-ml His-trap column (GE Healthcare) to remove the cleaved tag and was purified by gel filtration (Superdex 200 16/60, GE Healthcare) in the gel-filtration buffer.

For the pulldown assays, Not5c, Not5-ΔN and Not2-ΔN were expressed as N-terminal His-GST fusion proteins, whose tags were cleavable with 3C protease. Not2, Not5c and Not5-ΔN were affinity purified with a 5-ml His-trap column (GE Healthcare) as described above. Not2-ΔN was affinity purified at a pH of 8.5 (with buffer A and C at pH 8.5) instead of pH 7.5. Not2–Not5c, Not2-ΔN–Not5c and Not2–Not5-ΔN complexes were formed by mixing a 1:1.5 molar ratio of the larger to smaller protein and dialyzed in gel-filtration buffer (without 2 mM DTT) in the presence of 3C protease and TEV protease. The dialyzed proteins were subjected to a second His-affinity purification with a 5-ml His-trap column (GE Healthcare) and subsequent incubation with glutathione-agarose beads (Protino) for 2 h at 4 °C to remove the GST-tag contamination. The proteins were then purified by gel filtration (Superdex 75 10/30, GE Healthcare) in gel-filtration buffer. Not1c-GST was affinity purified as described above. The protein was dialyzed against heparin buffer A (20 mM Tris, pH 7.5, and 100 mM NaCl), applied onto the 5-ml heparin column (GE Healthcare) and purified with a gradient elution with heparin buffer A and heparin buffer A supplemented with 1 M NaCl. Not1c-GST was further purified by gel filtration (Superdex 200 10/30, GE Healthcare). The Not3 Not box was expressed as TEV protease–cleavable His6 fusion protein and purified in a similar way as mentioned above.

For the RNA-binding experiments, Not2 and Not5c were expressed and purified as individual proteins as described above. The Not2–Not5c complex was formed by mixing the proteins in a 1.25:1 molar ratio and subsequent overnight incubation with TEV protease. The cleaved protein was subjected to His-affinity purification to remove the cleaved tag and a subsequent heparin-column purification. The complex was further purified by gel filtration (Superdex 200 16/60, GE Healthcare). Not1c was expressed and purified as above, with an additional step of heparin purification included after the second His-affinity step. Not1c–Not2–Not5c was purified by mixing Not1c and Not2–Not5c in a 1:1.25 molar ratio and subsequent gel filtration (Superdex 200 16/60, GE Healthcare).

Limited proteolysis experiment.

0.6 mg/ml of the Not1Δ1347–Not2–Not5 complex was incubated with elastase (Roche) at a 1:10 (w/w) enzyme/protein ratio for 30 min on ice. The products of the proteolysis were identified by N-terminal sequencing and MS analysis. The interacting core complex was identified by size-exclusion chromatography of the proteolyzed sample.

Crystallization and structure solution.

The Not1c–Not2–Not5c complex was concentrated to 16 mg/ml and crystallized at room temperature in 8.5% (w/v) PEG 8000, 100 mM MES, pH 6.5, and 200 mM calcium acetate. The mercury derivative was prepared by cocrystallization of a solution of Not1c–Not2–Not5c with ethyl mercury phosphate (EMP) at 0.55 mM final concentration. The crystals were frozen in the presence of 20% glycerol as cryoprotectant. X-ray data were collected at 100 K at the SLS synchrotron (PXII and PXIII beamlines), with tuning of the wavelength at the Hg edge in the case of the EMP-containing crystals for SAD data collection. The data were processed with XDS⁵⁵. The crystals belong to a monoclinic space group (*P*21) with two molecules per asymmetric unit. We used PHENIX.autosol⁵⁶ for phasing and Buccaneer⁵⁷ for the initial automatic model building. We completed the model with iterative rounds of manual model building with Coot⁵⁸ and restrained refinement with PHENIX⁵⁶. The final model has 97.3% residues in the most-favored regions of the Ramachandran plot, as calculated with MolProbity⁵⁹.

Pulldown assays.

For the experiments in Figure 2f, 50 pmol of bait (GST or GST-Not1c) were incubated with 100 pmol of prey (Not2–Not5c, Not2-ΔN–Not5c and Not2–Not5-ΔN) for 1 h at 4 °C in 40 mM Tris, pH 7.5, 150 mM NaCl, 2 mM DTT, 12.5% (v/v)

glycerol and 0.1% (w/v) NP-40 (binding buffer). The protein mix was incubated with 20 μ l of GSH-agarose beads (Protino) for 1 h with gentle rocking at 4 °C. The resin was washed three times with the binding buffer, and the proteins were eluted in 15 μ l of binding buffer containing 100 mM glutathione. Input and precipitates were mixed with SDS loading dye, resolved on 4–12% Bis-Tris NuPAGE gel (Invitrogen) with MES as running buffer, and visualized by Coomassie-blue staining. A similar protocol was used for the GST pulldown assays in Figure 4d with 40 mM Tris, pH 8.5, 100 mM NaCl, 2 mM DTT and 12.5% (v/v) glycerol as binding buffer.

Sequence alignments and superpositions.

All the sequence alignments were done with ClustalW⁶⁰ and ALINE⁶¹, and the structural superpositions were done with SSM in Coot⁵⁸. The r.m.s. deviations reported are from the output of Coot. Structure-based sequence alignment was done in STRAP⁶² with the Aligner3D method and manually edited in ALINE.

Yeast strains.

Yeast strains used in this study are all derivatives of W303 (*ade2-1, can1-100, leu2-3,112, his3-11,15, ura3-1, trp1-1*). Genes differing from W303 are as follows: T26N28 (*MATa, Δ trp1, Δ NOT1::HIS3 pFL 38 (NOT1)*), BSY1110 (*MATa, Δ trp1, not2::HISMX6*), BSY1111 (*MATa, Δ trp1, not3::HISMX6*), BSY1230 (*MATa, Δ trp1, NOT3-VSV, NOT2-3HA::hisMX6*), BSY1231 (*MATa, Δ trp1, POP2-VSV, NOT2-3HA::hisMX6*), BSY1240 (*MATa, Δ trp1, NOT3-TAP::TRP1Kl, POP2-VSV, NOT4-HA::hisMX6*) and BSY1242 (*MATa, Δ trp1, NOT3-TAP::TRP1Kl, POP2-VSV, NOT2-HA::hisMX6*).

Coprecipitation assays.

Protein extract preparation and coimmunoprecipitation were performed as described previously²⁸.

Mutant analyses.

Mutations in Not1, Not2 and Not3 were constructed in plasmids pBS4806 (ref. 28) (Not1-TAP), pBS4968 (Not2-VSV) and pBS4975 (Not3-VSV) by one or multiple rounds of site-directed mutagenesis. The presence of the desired mutation was ascertained by sequencing. The resulting plasmids were introduced into yeast strains with the lithium acetate transformation procedure. Plasmid shuffling, growth assays, protein extraction and western blot analyses were performed with standard procedures as previously described²⁸.

Electrophoretic mobility shift assays.

Proteins at 3, 10, 13 and 20 μ M concentration (30, 100, 130 and 200 pmol, respectively) were incubated with 50 nM (0.5 pmol) of 5'-labeled RNA (A15 or U15) at 4 °C overnight in 20 mM HEPES, pH 7.5, 5 mM EDTA, 0.1% NP-40 and 2 mM DTT (EMSA buffer). The reaction mixtures were complemented with gel-filtration buffer to a final NaCl concentration of 54 mM, resolved on a 6% (w/v) native PAGE and visualized by phosphorimaging.

RNA cross-linking.

200 pmol (20 μ M) of Not1–Not2–Not5c complex were incubated with 2.5 pmol (250 nM) of body-labeled U20 RNA overnight in EMSA buffer at 4 °C. The cross-linking was performed by irradiation of the mix at a wavelength of 254 nm for 5 min on ice. The mixture was then treated with 1% SDS and 0.5 μ l of RNase A/T1 mixture at 37 °C for 5 min. The samples were heated with SDS loading dye at 70 °C for 2 min, separated on 13.5% (w/v) SDS-PAGE gel and visualized by phosphorimaging and Coomassie-blue staining.

Mass spectrometry.

UV-induced protein-RNA cross-linking and enrichment of cross-linked peptides. UV cross-linking and enrichment of cross-linked peptides was performed according to the established protocols described in ref. 63. Briefly, 1 nmol of the single-stranded U15 RNA oligonucleotide and 1 nmol of Not1c–Not2–Not5c complex were mixed in a 1:1 molar ratio, and the total reaction volume was brought to 100 μ l in 20 mM HEPES, pH 7.5, 50 mM NaCl, 2 mM DTT and 5 mM EDTA. The mixture was incubated on ice overnight for complex formation. The samples were then transferred to black polypropylene microplates (Greiner Bio-One) and irradiated at 254 nm for 10 min. After ethanol precipitation, the samples were denatured in 4 M urea and 50 mM Tris-HCl, pH 7.9, and digested for 2 h at 52 °C with 1 μ g RNase A (Ambion, Applied Biosystems). After RNA digestion, proteolysis with trypsin (Promega) was performed overnight at 37 °C. The sample was desalted on an in-house-prepared C18 (Dr. Maisch GmbH) column, and the cross-linked peptides were enriched on an in-house-prepared TiO₂ (GL Sciences) column with the protocol described in ref. 63. The samples

were dried and then resuspended in 10 μ l sample solvent (5% v/v ACN and 1% v/v FA) for MS analysis.

Nano-liquid chromatography and MS analysis. 5 μ L of the above sample was injected onto a nano-liquid chromatography system (Agilent 1100 series, Agilent Technologies) including a C18 trapping column of length \sim 2 cm and inner diameter 150 μ m, in line with a C18 analytical column of length \sim 15 cm and inner diameter 75 μ m (both packed in house; C18 AQ 120 \AA 5 μ m, Dr. Maisch GmbH). Analytes were loaded on the trapping column at a flow rate of 10 μ L/min in buffer A (0.1% v/v FA) and subsequently eluted and separated on the analytical column with a gradient of 7–38% buffer B (95% v/v acetonitrile and 0.1% v/v FA) with an elution time of 33 min (0.87%/min) and a flow rate of 300 nL/min. Online ESI-MS was performed with an LTQ-Orbitrap Velos instrument (Thermo Scientific), operated in data-dependent mode with a TOP10 method. MS scans were recorded in the m/z range of 350–1,600 and for subsequent MS/MS the top ten most-intense ions were selected. Both precursor ions as well as fragment ions were scanned in the Orbitrap. Fragment ions were generated by higher-energy collision dissociation (HCD) activation (normalized collision energy = 40) and recorded from $m/z = 100$. As precursor ions as well as fragment ions were scanned in the Orbitrap, the resulting spectra were measured with high accuracy (<5 p.p.m.), both in the MS and MS/MS level.

Data analysis.

The MS .raw files were converted into the .mzML format with msconvert⁶⁴. Protein-RNA cross-links were analyzed with OpenMS^{65,66} and OMSSA⁶⁷ as a search engine. Data-analysis workflows were assembled as described¹¹. The high-scoring cross-linked peptides were manually annotated for confirmation. Protein-RNA interactions between the complex and poly(U) RNA were analyzed with UV-induced protein-RNA cross-linking followed by MS. Peptide RCGNDFVYNEEDFEKL in Not5 (position 545–560) was observed carrying an additional mass of 476.0338 Da corresponding to U nucleoside with an adduct of 152. The γ -ion series could be observed from 1 to 10, unshifted. In contrast, b ions from b3 until b8 were observed with a mass shift corresponding to U-H3PO4 and 152 adduct (Fig. 5c). Also, the b ions from b5 until b8 were observed with a mass shift corresponding to U and 152. We have always observed that the 152 adduct is observed as a shift associated with cysteine, which could be the amino acid that is cross-linked. In the corresponding figure (Fig. 5c), the b ions that were observed with a mass shift corresponding to U-H3PO4 + 152 and to U + 152 are shown with an asterisk (*) and hash (#), respectively, and the immonium ions with IM.

RNase protection assays.

100, 150 and 200 pmol (10, 15 and 20 μ M) of Not1c–Not2–Not5c complex were incubated with 0.5 pmol (50 nM) of U20 RNA in the EMSA buffer overnight at 4 $^{\circ}$ C. The reaction mixtures were treated with 0.5 μ l of RNase A/T1 mix for 30 min at 4 $^{\circ}$ C. RNA was purified with phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. The purified RNA was 5' labeled with [γ -³²P]ATP with T4 polynucleotide kinase, repurified by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation, separated on 22% (w/v) denaturing PAGE with 5 M urea and visualized by phosphorimaging.

Fluorescence anisotropy.

5'-6-carboxy-fluorescein (6-FAM)-labeled U15 RNA was used in fluorescence anisotropy measurements at 20 $^{\circ}$ C with Genios Pro (Tecan). RNA was 9.1 nM at final concentration and was incubated with varying concentrations of Not1c–Not2–Not5c complex in the gel-filtration buffer supplemented with 10 mM EDTA. We used 250 mM NaCl in the buffer for the measurement because the protein was not stable at 100 or 150 mM salt in such high concentration at 20 $^{\circ}$ C. The excitation and emission wavelengths were 485 nm and 535 nm, respectively. Each titration point was measured three times with ten reads with an integration time of 40 μ sec. The data were analyzed by nonlinear regression fitting with Origin (OriginLab; <http://www.originlab.com/>).

Accession codes.

Coordinates and structural factors have been deposited in the Protein Data Bank under accession code 4BY6.

Caractérisation biochimique et structurale du
rôle du complexe CCR4-NOT et ATPase DDX6
dans la répression des microARN

E.3. Caractérisation biochimique et structurale du rôle du complexe CCR4-NOT et ATPase DDX6 dans la répression des microARN.

E.3.1. Contexte et objectif de l'étude

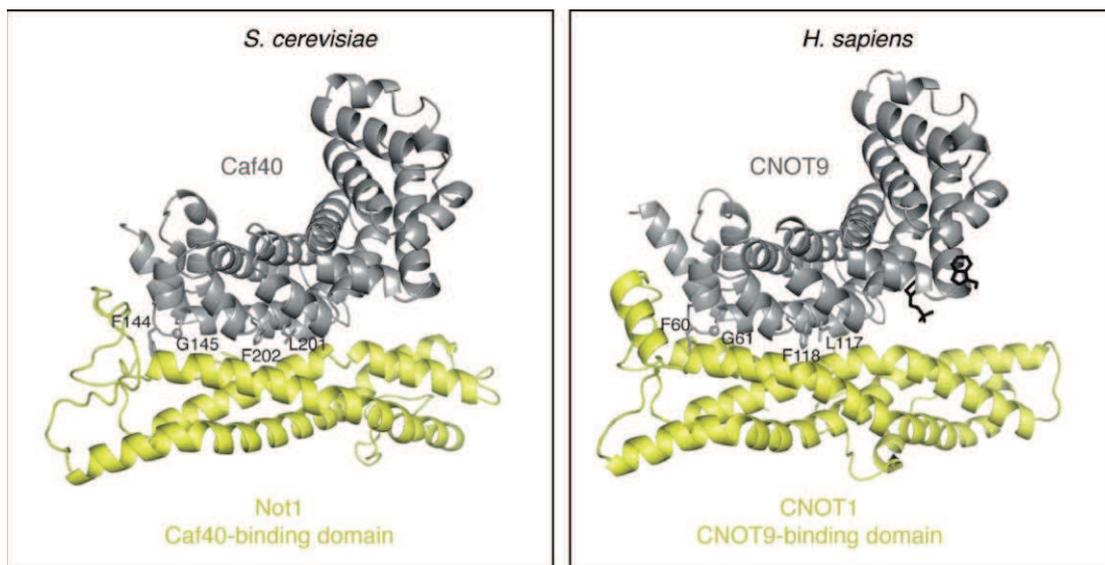
L'objectif de cette étude était de caractériser le rôle de deux régions de CNOT1, à savoir CN9BD et MIF4G, toutes deux impliquées dans l'activité inhibitrice des protéines GW182. Ces dernières sont des médiateurs importants de la répression des microARN. D'autre part le domaine MIF4G de CNOT1 recrute la protéine DEAD box DDX6, un inhibiteur de la traduction. Nous avons entrepris de comprendre le rôle de ces deux domaines de la protéine NOT1 en caractérisant ses interactions avec les protéines CNOT9 et DDX6.

E.3.2. Approches expérimentales

Notre laboratoire s'est concentré sur l'aspect structural de l'étude en combinant une nouvelle fois les approches biochimique et structurale. Il a été possible de mettre en évidence par des expériences d'interaction *in vitro* que le domaine CN9DB de NOT1 interagissait avec la protéine CNOT9. Nous avons focalisé nos efforts sur le complexe CNOT1 et CAF40 (homologue chez *S. cerevisiae* de CNOT9). Les protéines NOT1(1071-1282) et CAF40 ont été sous-clonées dans des vecteurs plasmidiques d'expression de type pET et sur-exprimées séparément afin de reconstituer le complexe. Une expérience de protéolyse ménagée avec la trypsine a permis d'isoler un complexe plus compact comprenant le domaine CN9BD de NOT1 et CAF40. Il a été possible de produire un échantillon suffisant pour les essais de cristallisation en réalisant cette procédure à plus grande échelle. Il a ainsi été permis de produire et d'optimiser des cristaux de ce complexe en vue de l'analyse par cristallographie aux rayons X. Plusieurs jeux de données ont été collectés à une résolution maximale de 3,8 Å. Un modèle de l'homologue humain de CAF40 (CNOT9) a été utilisé pour le remplacement moléculaire, ce qui a permis de résoudre une partie de la structure mais celle correspondant à CNOT1 n'était pas assez définie pour envisager une reconstruction à partir de cette carte de densité électronique. J'ai donc dérivatisé des cristaux avec un cluster de tantale de brome, collecté des données anormales et déterminé des phases expérimentales à basse résolution. Il m'a été ensuite possible de résoudre la structure en combinant remplacement

moléculaire et phases expérimentales. La carte de densité électronique ainsi obtenue était de qualité suffisante pour permettre le traçage de la plus grande partie du modèle. Cet aspect du projet a représenté la plus grande difficulté d'un point de vue cristallographique. En effet, l'interprétation de la carte et la construction du modèle ont nécessité un grand nombre de cycles de construction et d'affinement.

J'ai ensuite entrepris de produire les homologues humains CNOT1 et CNOT9 afin d'en résoudre également la structure. Cette démarche a été motivée par la possibilité d'obtenir une structure à plus haute résolution en changeant d'espèces. J'ai opté pour la même stratégie que celle utilisée pour le complexe de la levure en sur-exprimant une construction de CNOT1 (1054-1601) et CNOT9. Ce complexe a été protéolysé avec de la trypsine, puis purifié par chromatographie d'exclusion de taille et soumis à des essais de cristallisation. J'ai obtenu des cristaux de bonne qualité et la structure de ce complexe a été résolue par remplacement moléculaire à 2,1 Å en utilisant le modèle du complexe de la levure précédemment résolu. Comme les résidus tryptophanes du domaine CED de la protéine GW182 sont essentiels pour l'interaction avec CNOT9, nous avons réalisé une expérience de trempage de nos cristaux du complexe CNOT1 CNOT9 dans une solution contenant du tryptophane. Plusieurs jeux de données ont été collectés révélant deux sites d'interactions pour cet acide aminé ; l'un étant occupé par une molécule de tryptophane, l'autre étant occupé par une molécule de MES provenant du tampon de cristallisation. L'identification de ces sites nous a permis de générer des mutants de CNOT9 pour valider ces sites d'interactions.



Structure cristalline du complexe NOT1-CAF40
Résolution: 3,8 Å
Phasage expérimental avec un composé de bromure de Tantale
pdb: 4CV5

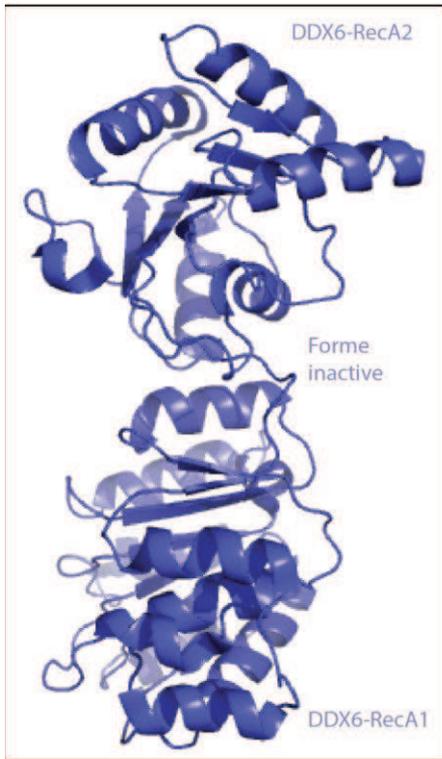
Structure cristalline du complexe CNOT1-CNOT9
Résolution: 2,0 Å
Remplacement moléculaire
pdb: 4CT7

Nous nous sommes ensuite aussi intéressés à l'ATPase DDX6. Des études d'interactions ont montré que cette protéine interagissait avec le domaine central MIF4G de la protéine NOT1. Il a été possible de surproduire la protéine DDX6 dans *E. coli* et de la purifier à une qualité compatible avec des essais de cristallisation. Il a été ensuite possible d'isoler un complexe stœchiométrique de la protéine DDX6 et du domaine MIF4G de la protéine NOT1 par chromatographie de perméation de gel, ainsi que de cristalliser à la fois la protéine DDX6 et le complexe DDX6-CNOT1(MIF4G). Les structures de ces deux complexes ont été réalisées par remplacement moléculaire à 3,0 et 2,3 Å respectivement.

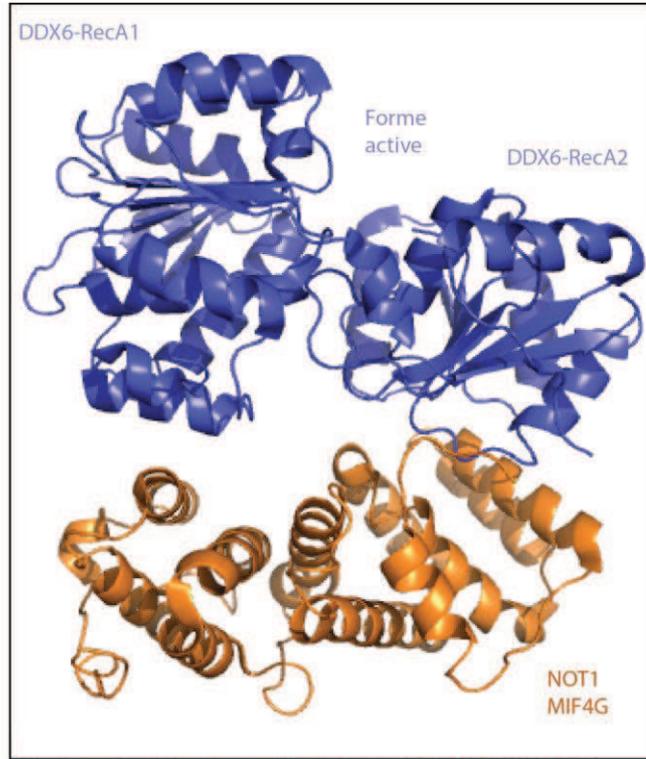
L'analyse de ces deux structures nous a montré que la protéine DDX6 subissait un large changement de conformation entre la forme libre et la forme complexée à la protéine CNOT1. Cette analyse nous a conduit à émettre l'hypothèse que l'activité ATPase de la protéine DDX6 pouvait être stimulée par ce changement de conformation. Des essais biochimiques d'ATPase ont permis de confirmer cette hypothèse.

E.3.3. Résumé de l'étude

Le domaine CN9BD de CNOT1 interagit directement avec la sous-unité CNOT9 du complexe CCR4-NOT, qui elle-même contacte la région répressive SD/CED de la protéine GW182, médiateur important de la répression des microARN, par un motif à tryptophane. Le domaine MIF4G de CNOT1 recrute la protéine DEAD box DDX6, un inhibiteur de la traduction, en stimulant son activité enzymatique. L'activité de DDX6 est importante pour la répression des microARN.



Structure cristalline de DDX6
 Résolution: 3,0 Å
 Remplacement moléculaire
 pdb: 4TC5



Structure cristalline du complexe binaire DDX6-CNOT1-MIF4G
 Résolution: 2,3 Å
 Remplacement moléculaire
 pdb: 4TC4

Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression

Hansruedi Mathys,^{1,2,7} Jérôme Basquin,^{3,7} Sevim Ozgur,³ Mariusz Czarnocki-Cieciura,^{4,5,6} Fabien Bonneau,³ Aafke Aartse,¹ Andrzej Dziembowski,^{4,5} Marcin Nowotny,⁶ Elena Conti,^{3,*} and Witold Filipowicz^{1,2,*}

¹Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland

²University of Basel, 4003 Basel, Switzerland

³Max Planck Institute of Biochemistry, Department of Structural Cell Biology, 82152 Martinsried/Munich, Germany

⁴Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-109 Warsaw, Poland

⁵Faculty of Biology, University of Warsaw, 02-109 Warsaw, Poland

⁶International Institute of Molecular and Cell Biology, 02-109 Warsaw, Poland

⁷Co-first authors

*Correspondence: conti@biochem.mpg.de (E.C.), witold.filipowicz@fmi.ch (W.F.)

<http://dx.doi.org/10.1016/j.molcel.2014.03.036>

SUMMARY

MicroRNAs (miRNAs) control gene expression by regulating mRNA translation and stability. The CCR4-NOT complex is a key effector of miRNA function acting downstream of GW182/TNRC6 proteins. We show that miRNA-mediated repression requires the central region of CNOT1, the scaffold protein of CCR4-NOT. A CNOT1 domain interacts with CNOT9, which in turn interacts with the silencing domain of TNRC6 in a tryptophan motif-dependent manner. These interactions are direct, as shown by the structure of a CNOT9-CNOT1 complex with bound tryptophan. Another domain of CNOT1 with an MIF4G fold recruits the DEAD-box ATPase DDX6, a known translational inhibitor. Structural and biochemical approaches revealed that CNOT1 modulates the conformation of DDX6 and stimulates ATPase activity. Structure-based mutations showed that the CNOT1 MIF4G-DDX6 interaction is important for miRNA-mediated repression. These findings provide insights into the repressive steps downstream of the GW182/TNRC6 proteins and the role of the CCR4-NOT complex in posttranscriptional regulation in general.

INTRODUCTION

MicroRNAs (miRNAs) control gene expression posttranscriptionally, generally by imperfectly base-pairing to the 3' UTR of target mRNAs and causing inhibition of their translation as well as deadenylation and destabilization. miRNAs function as part of ribonucleoprotein particles (miRNPs) with Argonaute (AGO) and GW182 family proteins as key components. The GW182s, known as TNRC6A, TNRC6B, and TNRC6C in mammals, are recruited to mRNAs by direct interaction with AGO proteins via their N-proximal Gly-Trp (GW) repeats (reviewed by Fabian

et al., 2010; Huntzinger and Izauralde, 2011). Although a direct role of AGOs in inducing repression by miRNAs has also been reported (Fukaya and Tomari, 2012), much evidence indicates GW182s to be mediators of both translational inhibition and deadenylation (Fabian et al., 2010; Huntzinger and Izauralde, 2011). Recent studies have revealed that GW182s induce mRNA deadenylation by recruiting two major cellular deadenylase complexes, PAN2-PAN3 and CCR4-NOT (Braun et al., 2011; Chekulaeva et al., 2011; Chen et al., 2009; Christie et al., 2013; Eulalio et al., 2009; Fabian et al., 2009, 2011; Piao et al., 2010).

Translational repression by miRNAs is less well understood, and several different mechanisms have been proposed. There is much evidence that early steps of cap-dependent initiation are targeted by miRNAs. However, targeting of other initiation steps, such as scanning by the initiation factor eIF4A, and of elongation has also been proposed (Fabian et al., 2010; Fukaya and Tomari, 2012; Huntzinger and Izauralde, 2011; Meijer et al., 2013). Apart from AGO and GW182 proteins, many other factors have been implicated in the regulation of repression by miRNAs, although the mechanisms remain largely unknown (Fabian et al., 2010). Importantly, kinetic studies have indicated that translational repression precedes mRNA deadenylation and decay, with the latter effects being generally dominant at steady state (Bazzini et al., 2012; Béthune et al., 2012; Djuranovic et al., 2012; Fabian et al., 2009).

Observations that the CCR4-NOT complex or its subunits act not only in mRNA deadenylation, but also in mediating translational repression, both miRNA controlled (Braun et al., 2011; Chekulaeva et al., 2011) and miRNA independent (e.g., see Cooke et al., 2010; Van Etten et al., 2012), indicate CCR4-NOT as a potential central player in posttranscriptional regulation. CCR4-NOT is an evolutionarily conserved complex of ~10 subunits, two of which (CCR4 and CAF1) have mRNA deadenylase activity (reviewed by Collart and Panasenko, 2012). Recently, it was shown that human TNRC6s interact directly with CCR4-NOT, likely by contacting its large scaffold subunit CNOT1. The TNRC6 C-terminal regions, referred to as C-terminal effector domain (CED) or silencing domain (SD), are responsible for the



CCR4-NOT recruitment, consistent with their activity as major effectors in miRNA repression (Braun et al., 2011; Chekulaeva et al., 2011; Fabian et al., 2011). Detailed analysis of the CED/SD regions of GW182/TNRC6 proteins, and also of a more N-proximal repressive region in *Drosophila* GW182, revealed that functional interaction with CCR4-NOT requires Trp (W) residues in the G/S/TW or WG/S/T context (W motifs), dispersed across the domains and acting in an additive manner (Chekulaeva et al., 2011; Chekulaeva et al., 2010). CIM-1 (CCR4-NOT interaction motif 1) and CIM-2 elements identified as important for functional TNRC6C interaction with CCR4-NOT also contain conserved W residues (Fabian et al., 2011).

While the importance of the W motif-bearing GW182/TNRC6 domains in the interaction with CCR4-NOT is well documented (Chekulaeva et al., 2011; Fabian et al., 2011; Huntzinger et al., 2013), little is known about their partners in the CCR4-NOT complex. Likewise, the mechanism of CCR4-NOT-mediated repression of mRNA translation is largely unknown. Here, we have used a combination of functional and structural approaches and revealed interactions and mechanisms related to the role of the CCR4-NOT complex in posttranscriptional regulation.

RESULTS

A Central Region of CNOT1 Mediates Repression of mRNA Function

We and others have identified the CCR4-NOT complex as a factor acting downstream of GW182 during miRNA-induced repression (Braun et al., 2011; Chekulaeva et al., 2011; Fabian et al., 2011). These studies have also suggested that CNOT1, the largest subunit of CCR4-NOT, plays a key role. We have now investigated which regions of human CNOT1 repress mRNA function. Different CNOT1 fragments expressed in human embryonic kidney 293 (HEK293) cells at similar levels (Figures 1A and S1A available online) were tested for their ability to repress protein synthesis when tethered to reporter mRNA (Figure S1B). Tethering NHA-tagged CNOT1 repressed activity of the *Renilla* luciferase (RL) 5BoxB reporter ~8-fold when compared to the protein lacking the λ phage N peptide. Terminal and internal deletions revealed that the central region of CNOT1 (residues 727–1,601; named CNOT1-R) was still able to repress mRNA function 4.5-fold (Figures 1A and 1B).

The CNOT1 fragments were also tested for repression of the RL-5BoxB reporter, which is devoid of the poly(A) tail. This reporter, RL-5BoxB-HSL+HhR (Figure S1B), is repressed by the tethered CNOT1 without a decrease in mRNA level (Chekulaeva et al., 2011). Effects of CNOT1 fragments on RL-5BoxB-HSL+HhR activity were qualitatively similar to those seen with the poly(A)⁺ reporter (Figure 1C). Both the full-length CNOT1 and the CNOT1-R fragment repressed activity of the poly(A) RNA approximately 4-fold, consistent with previous indications that the CCR4-NOT complex can mediate translational repression in a deadenylation-independent manner (Braun et al., 2011; Chekulaeva et al., 2011; Huntzinger et al., 2013).

The repressive region of CNOT1-R is centered at the MIF4G (middle of initiation factor 4G; Basquin et al., 2012; Petit et al., 2012) domain, which is flanked by the region that interacts with the RNA decay factor tristetraprolin (TTP) (Fabian et al., 2013)

and the region that binds the CCR4-NOT subunit CNOT9 (Bawankar et al., 2013), hereafter referred to as CNOT9-binding domain or CN9BD (Figure 1A). Individual deletion of each flanking domain had a limited effect on repression, but their simultaneous elimination (resulting in fragments bearing MIF4G alone) affected repression very strongly. The flanking domains had no appreciable repressive activity in isolation (Figures 1B and 1C).

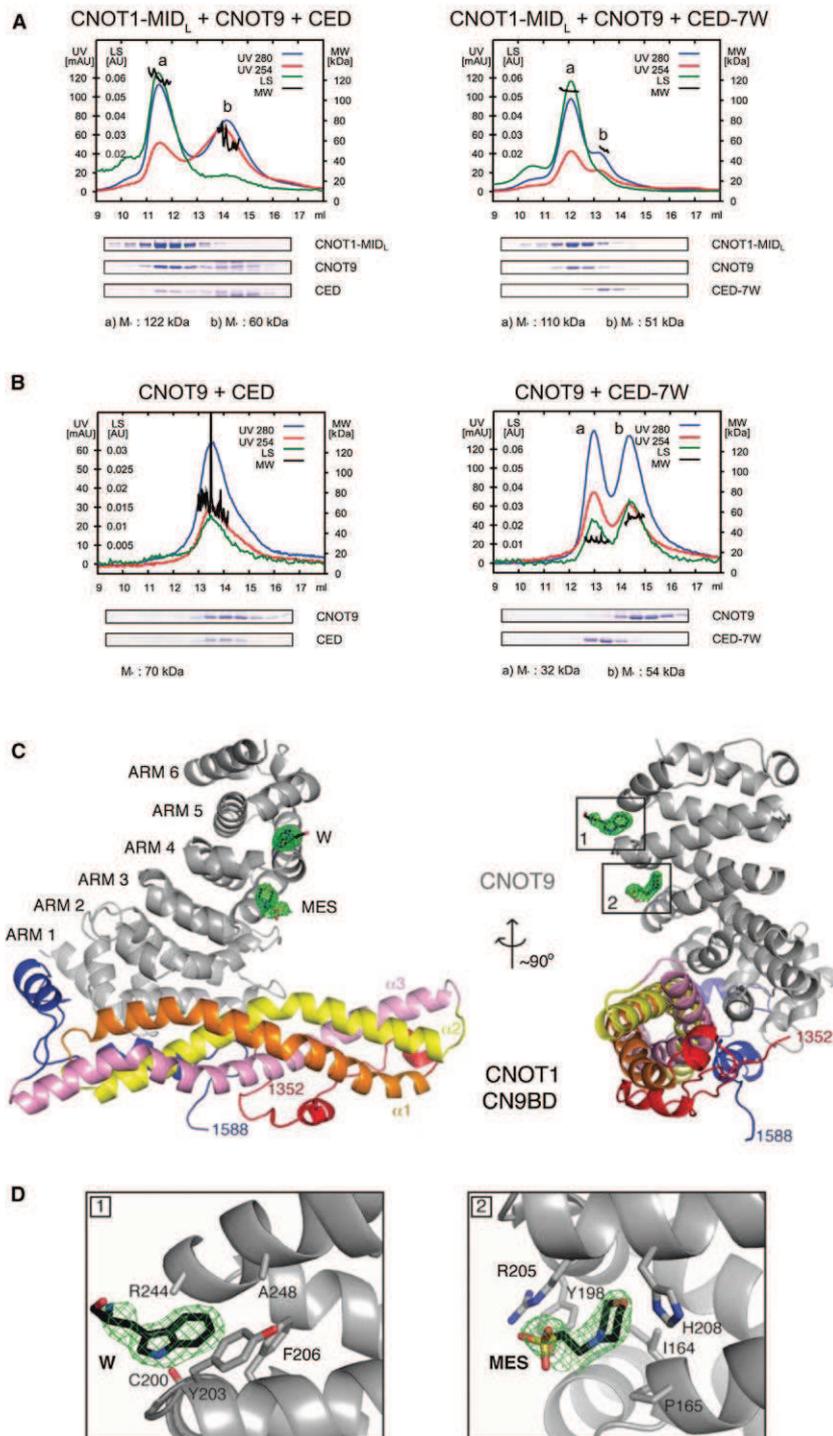
The CNOT1-R CN9BD Recruits TNRC6C via Interaction with CNOT9

We investigated which CNOT1 regions mediate the interaction with the TNRC6C C-terminal repressive regions SD (residues 1,260–1,690) or CED (residues 1,369–1,690) (Figure S1C). We tested CNOT1 fragments for their association with a glutathione S-transferase (GST) fusion of SD or CED coexpressed in HEK293 cells. The glutathione Sepharose pull-downs identified two different nonoverlapping CNOT1 fragments that strongly associate with SD in an RNA-independent manner, namely CN9BD and the C-terminal region CNOT1-C (residues 1,827–2,376) (Figures 1A, S1D, and S1E). Importantly, the interaction of both fragments with CED, tested in both directions, depended on Trp residues present in W motifs (Figures 1D and S1F). The weaker interaction with SD/CED of a third CNOT1 region, corresponding to residues 830–1,001 (TTP region; Figures 1A and S1D) was not investigated further. Since SD and CED regions showed no appreciable difference in their interaction with the CNOT1 fragments investigated, we will often refer to the silencing region as SD/CED.

We focused on CNOT1 CN9BD, which is part of the repressive CNOT1-R region. Experiments performed in HEK293 cells indicated that CNOT1 CN9BD is pulled down by CED more effectively in the presence of coexpressed CNOT9, and that CNOT9 expressed alone is pulled down by the CED in a W motif-dependent manner (Figure 1E). Testing CED subfragments, we found that substitution of just one or two Trp residues strongly affected the pull-down of CNOT9 (Figure S1G). Using size-exclusion chromatography with purified recombinant proteins, we assessed whether the interactions observed in HEK293 cells are direct (Figures 2A, 2B, and S2A–S2D). We identified a soluble fragment of CNOT1 encompassing both the MIF4G domain and CN9BD (CNOT1 MID_L; residues 1,092–1,827). Purified CNOT1 MID_L formed a trimeric complex with purified CNOT9 and CED (Figure 2A, left panel). Mutation of W residues in the CED (CED-7W) abrogated its interaction with the CNOT1 MID_L-CNOT9 binary complex (Figure 2A, right panel). Notably, CNOT9 alone (Figure 2B), but not CNOT1 MID_L alone (Figure S2E), was found to interact with the CED; the CNOT9-CED interaction was W motif dependent (Figure 2B). We concluded that CNOT9 bridges the interaction between CNOT1 CN9BD and the TNRC6C CED.

The CNOT1-CNOT9 Structure Reveals a Trp-Binding Pocket in CNOT9

We subjected the binary complex of CNOT9 with CNOT1 MID_S (residues 1,054–1,601) to limited proteolysis to cleave away the CNOT1 MIF4G domain and crystallized the resulting CNOT1 CN9BD-CNOT9 complex. The structure was solved to a 2 Å resolution and refined to an R_{free} of 24%, R_{factor} of 21%,



(legend on next page)

Table 1. Data Collection, Phasing, and Refinement Statistics

	CNOT1 CN9BD- CNOT9	CNOT1 CN9BD- CNOT9-W	Not1 CN9BD- Caf40-Ta	CNOT1 MIF4G- DDX6	DDX6
Data Collection					
Space group	C_{121}	C_{121}	P_{43212}	P_{21}	P_{212121}
Cell Dimensions (Å)					
<i>a</i>	153.67	155.031	188.341	71.54	113.12
<i>b</i>	66.141	66.984	188.341	120.03	113.14
<i>c</i>	72.436	72.653	126.689	78.48	77.18
Cell Dimensions (°)					
α	90.0	90.00	90.0	90.00	90.00
β	100.35	99.85	90.0	104.01	90.00
γ	90.0	90.00	90.0	90.00	90.00
Wavelength	1.00002	1.00002	1.25472	1.00000	1.00002
Resolution (Å)	47.90–2.00 (2.11–2.00)	48.27–2.0 (2.11–2.00)	48.29–3.81 (4.01–3.81)	76.15–2.29 (2.41–2.29)	80.00–3.00 (3.17–3.00)
R_{sym} or R_{merge}	3.7 (48.2)	5.0 (55.7)	13.5 (170.0)	7.0 (66.6)	7.6 (66.0)
<i>I</i> / σ <i>I</i>	14.6 (4.2)	15.6 (2.3)	18.0 (2.7)	11.6 (2.2)	11.7 (2.40)
Completeness (%)	98.9 (95.13)	99.4 (96.1)	99.8 (99.5)	98.9 (93)	99.4 (98.8)
Anomalous completeness (%)	–	–	99.7 (99.2)	–	–
Redundancy	4.6 (4.2)	4.6 (4.3)	25.5 (23.3)	4.7 (4.50)	5.2 (5.3)
Refinement					
Resolution (Å)	47.90–2.00 (2.11–2.00)	48.27–2.0 (2.11–2.00)	48.29–3.81 (4.01–3.81)	76.15–2.29 (2.41–2.29)	80.00–3.00 (3.17–3.00)
No. of reflections	41,743	49,511	24,269	66,414	27,769
$R_{\text{work}}/R_{\text{free}}$	0.1961/0.2342	0.1837/0.2212	0.292/0.316	0.1660/0.2198	0.2313/0.2624
No. of atoms	4,378	4,402	6,395	9,926	11,627
Protein	3,961	3,975	6,361	9,653	11,613
Ligand/ion	–	12	34	5	14
Water	417	415	–	268	0
B Factors	54.00	52.60	91.1	57.90	117
Protein	53.3	51.70	89.3	58.10	117
Ligand/ion	–	101.30	211	61.30	128
Water	60.50	60.30	–	50.60	–
Root-Mean-Square Deviation					
Bond lengths (Å)	0.009	0.008	0.005	0.008	0.004
Bond angles (°)	1.17	1.09	0.95	1.19	0.79

Two crystals (one native and one Hg derivative) were used for data collection. Values in parentheses are for the highest-resolution shell.

wraps around the other end of the rod and around the first ARM repeat of CNOT9.

Since Trp residues in the CED of TNRC6C are essential for the interaction with CNOT9, we used a strategy similar to that described for the TNRC6-binding protein AGO-2 (Schirle and

MacRae, 2012), and soaked free tryptophan in the CNOT1 CN9BD-CNOT9 crystals. The 2.0 Å data set (Table 1) revealed the presence of an unaccounted density peak on CNOT9, at the convex surface between ARM5 and ARM6, where a Trp moiety could be modeled unambiguously (Figure 2C). The Trp indole

Figure 2. CNOT9 Mediates TNRC6C-CED Interaction with CNOT1 CN9BD

(A and B) Analysis of complexes between CED, CNOT9, and the CN9BD-containing fragment of CNOT1 (CNOT1 MID_L; residues 1,092–1,827). Molecular weights (MW), measured by MALS, are given at the bottom below the gels showing protein distribution along the column. Theoretical MW values for CNOT9, CED, CED-7W, and CNOT1 MID_L are 30.5, 28.9, 28.2, and 80.5 kDa, respectively. LS, light scattering.

(C) Structure of CNOT9 (in gray, with the 6 ARM repeats labeled) and the CNOT1 CN9BD (with the different structural elements colored from red to blue). The structure is shown in two orientations, with the difference density corresponding to two small molecules (Trp and MES) (green, contoured at $\times\sigma$).

(D) Zoomed-in views of the two CNOT9-bound small molecules identified from the structural analysis. The two panels refer to the two insets in (C). Selected CNOT9 residues are shown and labeled.

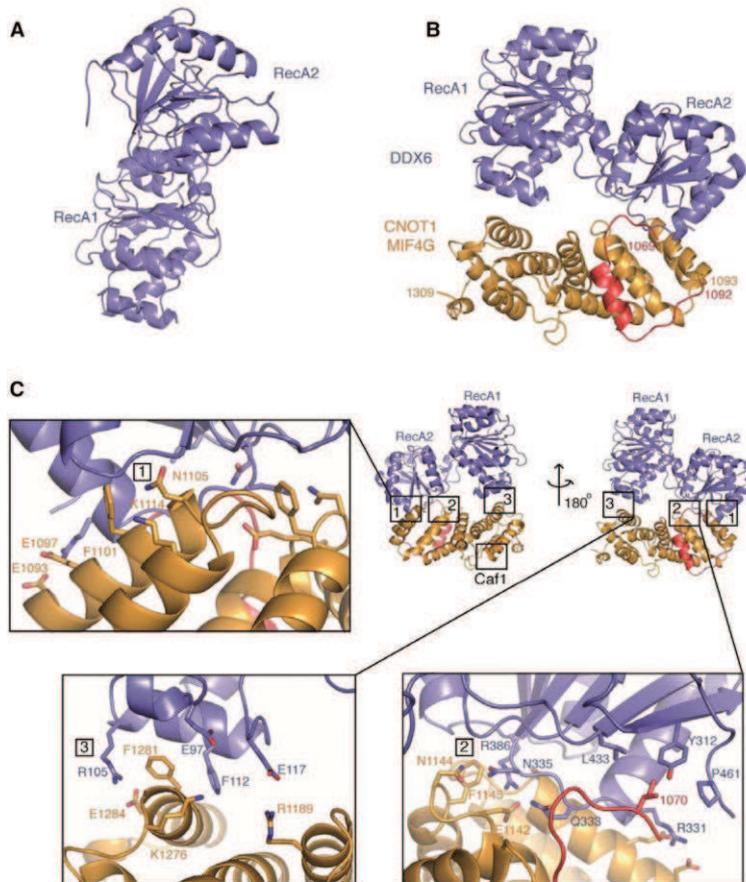


Figure 4. Structure of Human DDX6 in Complex with the CNOT1 MIF4G Domain

(A) Structure of DDX6 in isolation. The two RecA domains are indicated.

(B) Structure of DDX6 (blue) bound to CNOT1 MIF4G. The molecule is viewed with the RecA1 domain of DDX6 in the same orientation as the structure of unbound CNOT1 in (A). The five HEAT repeats of the CNOT1 MIF4G domain are in orange, and the N-terminal extension is in red.

(C) Hotspots of the interaction between MIF4G and DDX6 RecA2 (patches 1 and 2) or RecA1 (patch 3). The upper right panel shows the position of the patches in two views of the complex (rotated by 180° around a vertical axis). The zoomed-in panels highlight conserved interacting residues that have been subjected to mutagenesis for the experiments in Figure 6. CNOT1 mutations at patch 1 (F1101A, N1105A, K1114A), patch 2 (E1142R, N1144A, F1145A), and patch 3 (K1276A, F1281A, E1284A) correspond to mutants MIF4G_{hel1}, MIF4G_{hel2}, and MIF4G_{hel3}, respectively. The lower panels also show a subset of the DDX6 residues targeted for mutagenesis for the experiments in Figure 7 ([R105E, F112D in RecA1; mutant DDX6_{mif1}], [Q333E, N335E in RecA2; DDX6_{mif2}], and [R386E in RecA2; DDX6_{mif3}]). The CNOT1 surface targeted for disrupting CAF1 binding is indicated in the upper right panel.

protein phosphatase 2A, and the yeast kinase TOR1) repeats pack side by side and form a slightly curved molecule with characteristic concave and convex surfaces. The CNOT1 MIF4G interacts with CAF1 via a small patch of conserved residues on its convex surface (Basquin et al., 2012; Petit et al., 2012). However, the most extensive patch of evolutionarily conserved residues resides on the concave surface (Basquin et al., 2012; Petit et al., 2012). Previous structural studies have shown that the MIF4G domains of several proteins interact with RNA-dependent ATPases/helicases of the DEAD-box family via their concave surface, as in the case of eIF4A–eIF4G (Schütz et al., 2008), Gle1–Dbp5 (Montpetit et al., 2011), and CWC22–eIF4AIII (Buchwald et al., 2013). We reasoned that the CNOT1 MIF4G domain might also interact with a DEAD-box protein.

To investigate interacting proteins, we expressed GST-CNOT1-R in HEK293 cells, performed pull-downs, and analyzed the precipitates by mass spectrometry (MS). We identified the DEAD-box protein DDX6 and a member of a different ATPase/helicase family, the DEAA-box protein HELZ (Table S1). DDX6 (also known as Rck/p54 in mammals, Me31b in flies, and Dhh1

in yeast) is an established inhibitor of translation and activator of decapping (Parker and Sheth, 2007) and has been implicated in miRNA repression (Chu and Rana, 2006; Su et al., 2011) as well as interaction with the CCR4-NOT complex in yeast (Maillet and Collart, 2002). We tested the interaction of DDX6 with the CNOT1 MIF4G domain using purified

The Structure of the CNOT1 MIF4G-DDX6 Complex Explains the Basis of the Conserved and Specific Interaction

DDX6 has the typical domain architecture of DEAD-box proteins, with two RecA domains containing the ATP- and RNA-binding residues (reviewed by Henn et al., 2012). We first determined the 3.0 Å resolution crystal structure of a region of DDX6 encompassing both RecA domains (residues 95–469) (R_{free} of 26.2%, R_{factor} of 23.1%; Table 1). In the apo state of DDX6, the two RecA domains pack against each other in a close configuration with extensive intramolecular contacts (Figure 4A), although with an overall conformation different to that observed in the yeast ortholog Dhh1 (Cheng et al., 2005; Figure S4D). We then determined the crystal structure of DDX6_{95–469} in complex with a CNOT1 MIF4G fragment (residues 1,063–1,314) (Figure 4B). The final model of the complex was refined to a 2.3 Å resolution, with an R_{free} of 22.0%, R_{factor} of 16.6%, and good stereochemistry (Table 1). In the complex, the RecA domains of DDX6

show a conformation different from the apo state and contact the concave surface of the CNOT1 MIF4G domain (orange) (Figure 4B). In addition, an N-terminal extension of the MIF4G domain (red) contributes to the formation of the RecA2-binding site.

In the DDX6-CNOT1 structure, HEAT repeats 1 and 2 of MIF4G bind the RecA2 domain of DDX6, while HEAT 5 binds the RecA1 domain. We identified three patches of evolutionarily conserved interactions (Figure 4C). Patch 1 is centered at CNOT1 Phe1101, Asn1105, and Lys1114, and Patch 2 is centered at Glu1142, Asn1144, and Phe1145, both at the RecA2-binding interface. Patch 3 is at the RecA1-binding site and is centered at CNOT1 Phe1281, Glu1284, and more peripherally, Lys1276. The interactions involve DDX6 Arg105 and Phe112 in the RecA1 domain and Arg331, Gln333, and Asn335 in the RecA2 domain (Figure 4C).

Several CNOT1-binding residues are conserved not only in DDX6 orthologs, but also in eIF4A paralogs (Figure S4B). Recently, eIF4AII has been implicated in CNOT1 binding (Meijer et al., 2013). However, structural superimpositions suggested that eIF4AII binding would result in electrostatically unfavorable interactions (Figure S4B). We designed and purified a DDX6 mutant that would mimic eIF4AII at these positions. Mutating DDX6 Thr327 and Ser330 to the corresponding negatively charged residues of human eIF4AII (Asp266, Glu269) indeed abolished binding to CNOT1 in vitro (Figure S4A). We concluded that eIF4A proteins are unlikely to interact directly with the CNOT1 MIF4G domain (see below).

Structure-Based Mutations in the CNOT1 MIF4G Domain Disrupt DDX6 Binding In Vivo

We designed structure-based mutations on the concave surface of the CNOT1 MIF4G domain to specifically weaken the interaction with DDX6. Residues listed above for patches 1, 2, and 3 of CNOT1 (Figure 4C) were substituted by alanine (with the exception of Glu1142 changed to Arg), giving rise to MIF4G_{hel1}, MIF4G_{hel2}, and MIF4G_{hel3} mutant proteins and combinations thereof (Figure S5A). We verified that the mutations did not disrupt CNOT1-R interaction with CAF1 (Figure S5B), arguing that they do not significantly affect the MIF4G fold. We also introduced mutations in the CNOT1 convex surface to impair CAF1 binding (F1252E and P1257E [MIF4G_{caf1}]; Basquin et al., 2012; Petit et al., 2012) (Figure S5A). We expressed GST-CNOT1-R or its MIF4G_{hel2+3}, MIF4G_{caf1}, and combination MIF4G_{hel2+3/caf1} mutants in HEK293 cells and carried out pull-down assays. The DDX6 association with CNOT1-R or its MIF4G_{caf1} mutant was strongly affected by the MIF4G_{hel2+3} mutation (Figures 5A and 5B). Importantly, other DEAD-box proteins with established functions in translation initiation, including eIF4AI, eIF4AII, DDX3X, and DHX29 (reviewed by Hinnebusch, 2011), did not associate with CNOT1-R (Figures 5A and 5B). These findings confirmed the MS and in vitro data (Figure S4A). Western blotting also confirmed the interaction with the HELZ helicase. However, since the HELZ interaction was not affected by the MIF4G_{hel2+3} mutation and rather involved CNOT1 sequences flanking the MIF4G domain (Figures 5A–5C), it was not investigated further.

We found that the CNOT1-R subfragment corresponding to the MIF4G domain alone (residues 1,054–1,314; CNOT1

MIF4G) pulled down endogenous DDX6 more effectively than CNOT1-R; here again, the MIF4G_{hel2+3} mutation eliminated the interaction (Figure 5C). To further document the specificity of the interaction of CNOT1 MIF4G with DDX6, we also performed pull-downs with the MIF4G domain of human eIF4GI. The eIF4GI MIF4G domain effectively associated with endogenous eIF4AI and eIF4AII, but not with DDX6. In contrast, CNOT1 MIF4G associated with DDX6, but not eIF4AI or eIF4AII (Figure 5D).

Disruption of DDX6 Binding Affects the Repressive Function of CNOT1

We tested the importance of the DDX6-CNOT1 interaction for the repressive activity of CNOT1-R. Tethering experiments revealed that each of the three CNOT1-R MIF4G_{hel} mutations partially relieved repression of the poly(A)⁺ RL-5BoxB reporter (Figure 6A). The effect of combining two or three patch mutations was not stronger than that of single-patch mutants (Figure S6A). The MIF4G_{caf1} mutation relieved the repression more than MIF4G_{hel} mutations, and combination of MIF4G_{caf1} and MIF4G_{hel} mutations had an additive effect (Figures 6A and S6A).

To eliminate a potential contribution of deadenylation to the observed effects, we tested the influence of different CNOT1-R MIF4G mutations on the translation of poly(A)⁺ reporter and of poly(A)⁺ reporter in the presence of overexpressed dominant-negative catalytic mutants of PAN2 and CAF1 deadenylases, PAN2_{cat} and CAF1_{cat}. We and others have shown previously that these mutants effectively inhibit miRNA-mediated deadenylation of reporters (Béthune et al., 2012; Chen et al., 2009; Piao et al., 2010). For both of the aforementioned conditions, the effect of single MIF4G_{hel} mutations was markedly stronger than that seen with mRNA undergoing deadenylation. The MIF4G_{hel} mutations, either alone or when combined with MIF4G_{caf1}, generally alleviated repression 2.5- to 3.5-fold (Figures 6B and 6C). In contrast, the MIF4G_{caf1} mutation had no effect on CNOT1-R-induced repression of reporters not undergoing deadenylation (Figures 6B and 6C).

Northern analysis of RNA isolated from cells expressing the RL-5BoxB poly(A)⁺ reporter in the absence or presence of PAN2_{cat}/CAF1_{cat} revealed that the MIF4G_{hel} mutations had a minimal (PAN2_{cat}/CAF1_{cat} absent; Figure 6D) or moderate (~1.5-fold increase; PAN2_{cat}/CAF1_{cat} present; Figure 6E) effect on mRNA levels. In both conditions, the MIF4G_{caf1} mutation alone had no effect on the RL-5BoxB RNA level, but in combination with MIF4G_{hel} mutations it noticeably suppressed mRNA decay, particularly when PAN2_{cat}/CAF1_{cat} were present (Figures 6D, 6E, and S6B). Comparison of the protein and mRNA data (Figure S6B) indicated that individual MIF4G_{hel} mutations had a relieving effect on protein stronger than that on mRNA levels.

Importantly, the MIF4G_{hel} mutations, either alone or in combination with MIF4G_{caf1}, also had significant suppressive effects on repression induced by tethering of full-length CNOT1 to either poly(A)⁺ or poly(A) reporter (Figures S6C and S6D).

CAF1 Deadenylase Does Not Contribute to Translational Repression

The results presented above, particularly those obtained with reporters not undergoing deadenylation, indicated that the

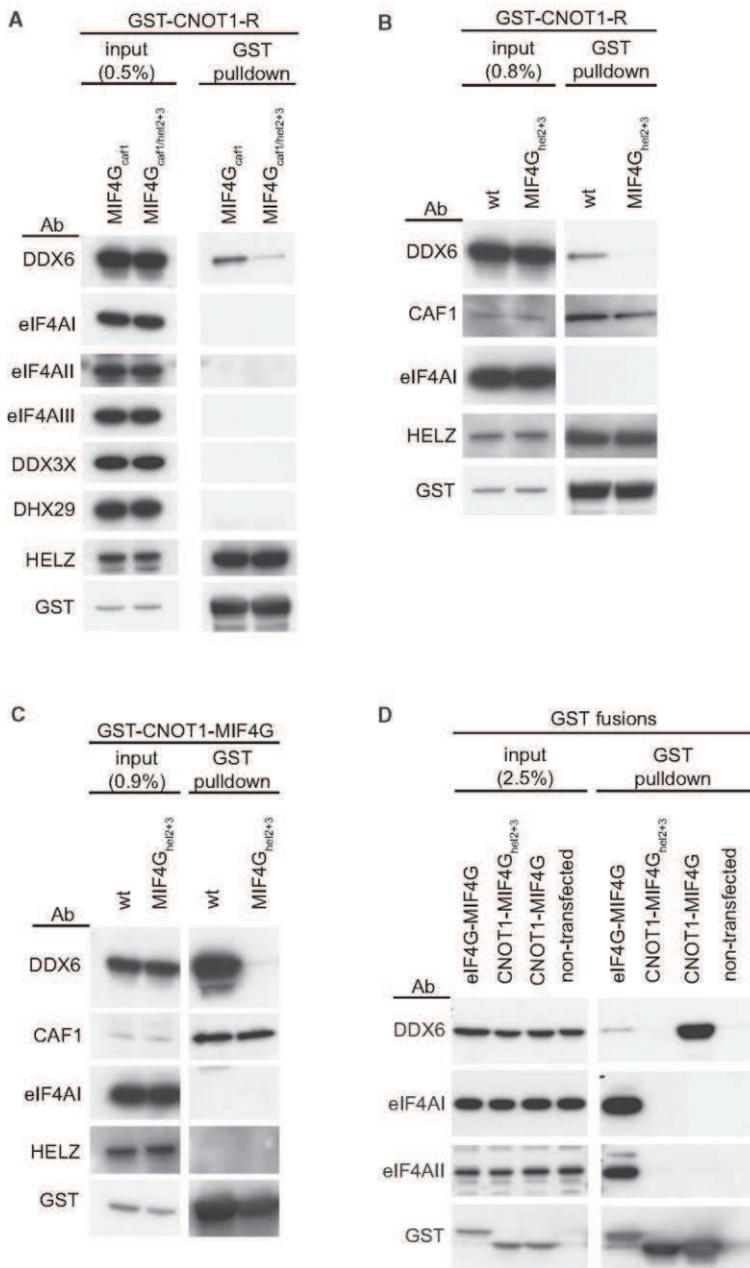


Figure 5. The CNOT1 MIF4G Specifically Interacts with Endogenous DDX6, but Not Other Tested ATPases/Helicases

(A and B) The interaction of CNOT1-R, either the MIF4G_{cat1} mutant form (A) or MIF4G WT form (B), with DDX6 is diminished by the MIF4G_{hel2-3} mutation.

(C) Isolated MIF4G domain (CNOT1 MIF4G) pulls down endogenous DDX6 specifically and more effectively than CNOT1-R.

(D) MIF4G domain of eIF4G interacts with endogenous eIF4AI and eIF4AII, but not DDX6, while CNOT1 MIF4G interacts only with DDX6.

tested the effect of direct tethering of CAF1 and its mutants on the activity of poly(A)⁺ and poly(A)⁻ RL reporters. CAF1 mutants unable to interact with either CNOT1 (CAF1_{not1}) or the deadenylase CCR4 (CAF1_{ccr4}), and combinations thereof with the CAF1 catalytic mutation (CAF1_{catal}), were found to eliminate only interactions with the expected protein partners (Figures S6E and S6F).

Consistent with previous results (Bawankar et al., 2013; Chekulaeva et al., 2011; Cooke et al., 2010), tethering of wild-type (WT) CAF1 or the CAF1_{catal} mutant repressed the activities of both poly(A)⁺ and poly(A)⁻ reporters. However, analysis of the remaining mutants (Figures S6G and S6H) strongly suggested that CAF1 per se is not a translation inhibitor but exerts its repressive activity by recruiting CNOT1, which acts as a potent inhibitor of translation of both poly(A)⁺ and poly(A)⁻ RNAs, probably by contacting DDX6 (see Discussion).

DDX6 Mutations Preventing the Interaction with CNOT1 MIF4G Mitigate miRNA Repression

Structural information on the DDX6-MIF4G complex also allowed us to design mutations in DDX6 that affect its association with CNOT1 MIF4G. Four mutants of DDX6 were generated: one in the RecA1 domain (R105E, F112D [DDX6_{mf1}]), two in the RecA2 domain (Q333E, N335E [DDX6_{mf2}] and R386E [DDX6_{mf3}]), and one representing a combination mutant

integrity of the concave surface of MIF4G is more important for the repression than the potential of MIF4G to associate with CAF1 deadenylase. Yet, Cooke et al. (2010) reported recently that the tethering of CAF1 to either poly(A)⁺ or poly(A)⁻ mRNA reporters leads to the cap-dependent repression of translation in *Xenopus* oocytes. To address this apparent inconsistency, we

(DDX6_{mf1+2}) (for more details, see Figures 4C and S7A). All failed to interact with the CNOT1 MIF4G domain expressed in HEK293 cells (Figure 7A).

Complementation assays were used to test whether CNOT1 MIF4G-mediated recruitment of DDX6 is required for repression of reporters responding to miRNAs; mutant reporters with

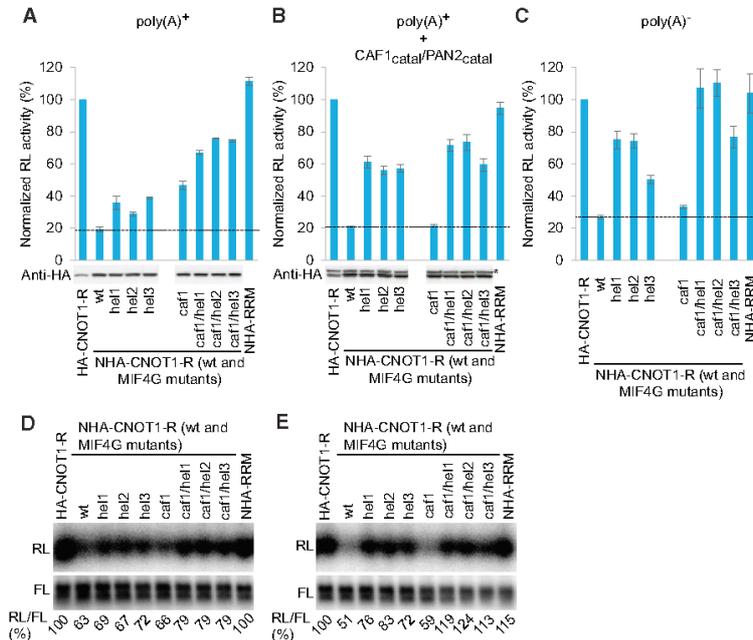


Figure 6. Importance of the DDX6-Interacting Surface of the CNOT1 MIF4G for Mediating Repression of mRNA Function by Tethered CNOT1-R

(A–C) Effect of mutations in the CNOT1-R MIF4G on repression of poly(A)⁺ (A), poly(A)⁺ (in HEK293 cells also expressing PAN2^{cat} and CAF1^{cat}) (B), and poly(A)⁻ (C) reporters. Values are means ± SEM (n = 3) with the value for HA-CNOT1-R at 100%. Western blots documenting similar expression are below (A) and (B). *HA-tagged PAN2^{cat}. NHA-RRM is not shown due to its faster migration.

(D and E) mRNA northern blots for (A) and (B) are shown in (D) and (E), respectively. Numbers are averages of two biological replicates corrected for coexpressed FL mRNA (see also Figure S6B).

nucleotide substitutions in seed regions were used as controls. Assays were performed in either the absence or presence of PAN2^{cat} and CAF1^{cat} to eliminate effects resulting from mRNA deadenylation. Following depletion of the endogenous DDX6 by RNAi (Figure S7B), cells were transfected with plasmids expressing RNAi-resistant DDX6, either WT or DDX6^{mif} mutant forms. We found that depletion of DDX6 partially relieved repression of each of the three tested reporters responding to miRNAs expressed either endogenously or exogenously in HEK293 cells. Importantly, WT DDX6, but not DDX6^{mif1}, DDX6^{mif3}, or DDX6^{mif1+2}, were able to rescue reporter repression, irrespective of whether or not PAN2^{cat} and CAF1^{cat} were coexpressed (Figures 7B, 7C, and S7C). The RecA2 mutant DDX6^{mif2} retained considerable repressive activity. All DDX6 mutants were expressed at comparable levels (Figure S7D). We also verified that DDX6 depletion had no effect on levels of CNOT1, AGO, and GW182/TNRC6 proteins (Figure S7F). Taken together, the data support a role for MIF4G-DDX6 interaction in miRNA repression.

CNOT1 MIF4G Binding Results in Activation of the DDX6 ATPase

Having established the importance of the CNOT1 MIF4G-DDX6 interaction in miRNA repression, we sought to determine the molecular mechanism. Comparison of DDX6 in the apo and CNOT1-bound states showed a large conformational change in the DEAD-box protein (Figures 4A and 4B). In the apo structure, the occurrence of tight intramolecular interactions between the RecA domains is reminiscent of the yeast ortholog Dhh1, which is in a constitutively inactive conformation (Dutta et al., 2011). In the CNOT1-bound structure, the two RecA

domains are oriented such that the RNA-binding and ATP-binding residues approach the positions expected for the active, ATPase-competent state. The conformation of DDX6 bound to CNOT1 is similar to that observed in the structures of eIF4A-eIF4G (Schütz et al., 2008) and Gle1-Dbbp5 (Montpetit et al., 2011), which represent activated states of these ATPases (Figure S4C).

The structural analysis prompted us to test whether CNOT1 binding impacts the ATPase activity of DDX6. Reactions containing purified proteins (Figure S7E), polyuridylic acid, and α-[³²P]ATP were analyzed by thin-layer chromatography (Figure 7D). For DDX6 alone, we detected very low accumulation of ADP. In contrast, incubation of DDX6 with the CNOT1 MIF4G domain resulted in increased hydrolysis of ATP. Importantly, the result of DDX6 incubation with the CNOT1 MIF4G^{hel2+3} mutant was similar to DDX6 in isolation. From these qualitative assays, we concluded that interaction with CNOT1 MIF4G stimulates the ATPase activity of DDX6 by releasing it from the constitutive inactive conformation of the unbound state and stabilizing it in an activated conformation.

DDX6 ATPase Activity Is Important for miRNA Repression

We used complementation assays described above to determine if ATPase activity is important for the DDX6 function as mediator of miRNA repression. We generated DDX6 mutants (DDX6^{cat1} through DDX6^{cat4}; Figure S7A) bearing amino acid substitutions in four different motifs known as important for ATP binding or hydrolysis and, consequently, for RNA or RNP remodeling (Henn et al., 2012). We found that WT DDX6, but none of the mutants, was able to rescue the RL-Hmg2 reporter repression, irrespective of whether or not PAN2^{cat} and CAF1^{cat} were coexpressed (Figure 7E; for western analysis, see Figure S7F). We concluded that DDX6 ATPase activity is essential for miRNA repression.

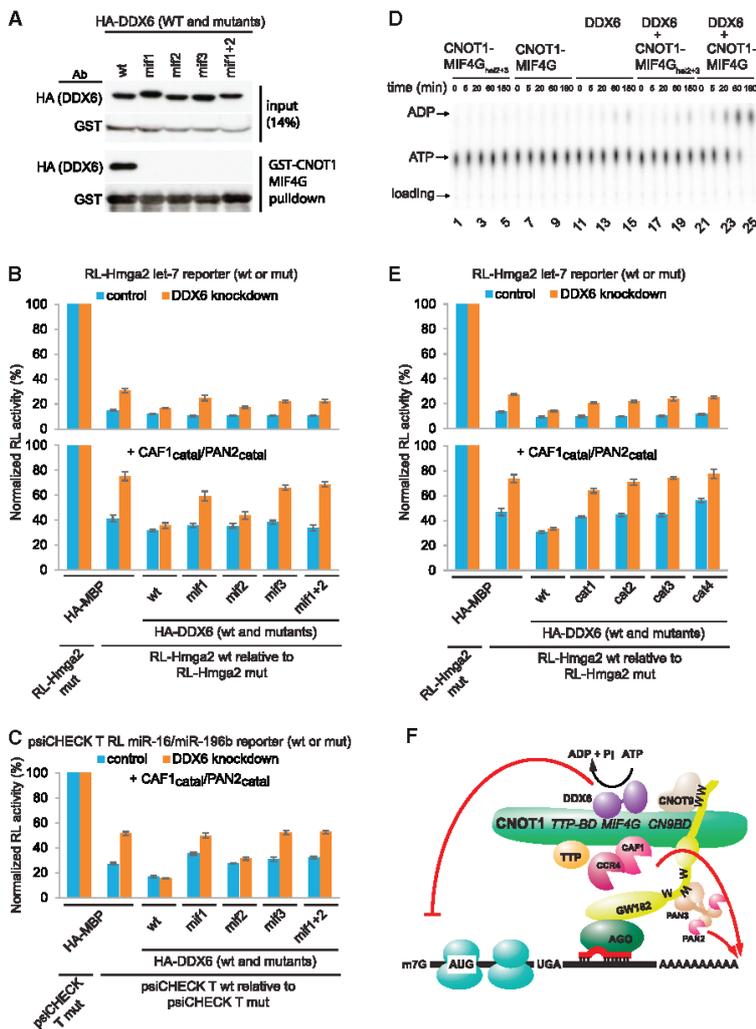


Figure 7. Importance of the CNOT1 MIF4G- DDX6 Interaction and ATPase Activity of DDX6 for miRNA-Mediated Repression

(A) Mutations in DDX6 affect its interaction with CNOT1 MIF4G. GST-CNOT1 MIF4G and HA- DDX6 (WT and mutants) were expressed in HEK293 cells and used for pull-downs. (B and C) DDX6_{mif} mutations mitigate repression of reporters responding to exogenous (B) or endogenous (C) miRNAs. In (B) (lower graph) and (E), HEK293 cells also expressed PAN2_{catal} and CAF1_{catal}. Values are means ± SEM (n = 3–7). (D) Effect of MIF4G and its MIF4G_{hel2+3} mutant on the ATPase activity of DDX6. (E) DDX6_{catal} mutations prevent repression of RL-Hmga2 reporter. In lower graph, HEK293 cells also expressed PAN2_{catal} and CAF1_{catal}. Values are means ± SEM (n = 4–5). (F) Central CNOT1 region as a hub of protein-protein interactions important for posttranscriptional regulation.

CCR4-NOT, and that CNOT9, through association with the CN9BD of CNOT1, acts as an adaptor mediating the recruitment of the CCR4-NOT complex to TNRC6C. Structural analysis identified two potential Trp-binding pockets in CNOT9. We found that integrity of each pocket is essential for binding of the CED, supporting a role for a combination of the CED W motifs in the interaction. In contrast, the TNRC6 SD/ CED interaction with the PAN3 subunit of the PAN2/PAN3 deadenylase appears to rely on a single Trp-binding pocket (Christie et al., 2013). Moreover, the PAN3 interaction was shown to depend on the M2 region, in distinction with the CNOT9, which preferentially bound to the CED Cterm (Figures S1G and S3A). We have shown previously (Chekulaeva et al., 2011) that the W motifs present in M2 and Cterm

CED regions function, in a context of full-length TNRC6 proteins, in a genuine miRNA-mediated repression. Present data provide molecular insights into the mechanism of the CCR4-NOT complex recruitment by the W motifs and also argue that these motifs may represent a collection of functionally different elements.

The CN9BD is part of a larger region of CNOT1 (CNOT1-R), which we have identified as a potent repressor of mRNA activity. The CN9BD has a helical bundle architecture in contrast to the HEAT repeat architecture of all other known CNOT1 domains (Basquin et al., 2012; Bhaskar et al., 2013; Boland et al., 2013; Fabian et al., 2013; Petit et al., 2012). The structure of the CNOT9-CN9BD complex reveals that the CNOT9 surface involved in the interaction with CN9BD is also responsible for CNOT9 dimerization (Garces et al., 2007). The two interactions appear to be mutually exclusive, suggesting that CNOT9 dimerization may have a regulatory role.

DISCUSSION

By combining biochemical and structural approaches, we have characterized the involvement of two CNOT1 regions, CN9BD and MIF4G, in the inhibitory activity of GW182/TNRC6 proteins as key mediators in miRNA repression. The CNOT1 CN9BD domain interacts directly with the CCR4-NOT subunit CNOT9, which in turn contacts the repressive SD/CED region of TNRC6C in a W motif-dependent manner. The CNOT1 MIF4G domain recruits the DEAD-box ATPase DDX6, an established inhibitor of translation, stimulating its enzymatic activity. The ATPase activity of DDX6 is important for miRNA repression.

CNOT9 Mediates TNRC6C-CNOT1 Interaction

One of the important findings of this work is that the SD/CED region of TNRC6C directly contacts the CNOT9 subunit of

The CNOT9-mediated mechanism is not the only way by which CCR4-NOT may be recruited by TNRC6C. We found that the TNRC6C SD/CED is also able to associate with the C terminus of CNOT1 (in a W motif-dependent manner) and, less effectively, with the TTP-binding CNOT1_{830–1,001} region. Interaction of the latter region with the TNRC6C SD is direct and depends on the integrity of CIM-1 and CIM-2 elements (M. Fabian and N. Sonenberg, personal communication). These observations point to the complexity and regulatory potential of the GW182/TNRC6 interactions with CCR4-NOT. They also provide a plausible explanation of why CNOT9 depletion and rescue experiments did not provide a clear answer as to the role of this protein in miRNA repression. It is likely that redundant TNRC6-CNOT1 interactions mediated by TTP and C-terminal regions of CNOT1 mask the effect of CNOT9 depletion.

The CNOT1 MIF4G Domain Recruits and Activates the DEAD-Box ATPase DDX6

In CNOT1-R, a centrally positioned MIF4G domain was found to be indispensable for the induction of mRNA repression and to specifically recruit the DEAD-box protein DDX6. The overall mode of interaction of CNOT1 MIF4G with DDX6, as revealed by crystallization of the complex, is similar to that observed in the structures of yeast eIF4A-eIF4G (Schütz et al., 2008) and Gle1-Dbp5 (Montpetit et al., 2011). Comparison with the structure of human DDX6 in isolation revealed a dramatic effect of CNOT1 MIF4G binding on the conformation of DDX6, with the relative position of the two RecA domains in the CNOT1-bound complex approaching the typical conformation of DEAD-box proteins in the active, ATPase-competent state (Henn et al., 2012). Consistently, while isolated DDX6 had no significant RNA-dependent ATPase activity, addition of the purified CNOT1 MIF4G markedly stimulated ATP hydrolysis.

Our identification of the CNOT1 MIF4G as the ligand of DDX6 that induces its ATPase activity fills an important gap in understanding the role of DDX6 proteins in posttranscriptional regulation. To date, attempts to measure activity of purified DDX6 have met with limited success (e.g., Tritschler et al., 2009). Only recently, Dutta et al. (2011) reported that yeast Dhh1 has weak RNA-dependent ATPase activity that can be enhanced by mutations disrupting intramolecular RecA1-RecA2 contacts. Our findings that mutations in ATP binding/hydrolysis motifs of DDX6 compromise its repressive activity strongly suggest that CNOT1-induced activation of the DDX6 ATPase is also important for its function in miRNA repression. Previous work has shown that similar motif mutations avert the repressive activity of DDX6 proteins in cell extracts (Coller and Parker, 2005) and tethering assays in *Xenopus* oocytes (Minshall et al., 2009; Minshall and Standart, 2004). These and other authors have found that DDX6 inhibits translation at the initiation step, in the absence of mRNA decay (Carroll et al., 2011; Coller and Parker, 2005; Minshall et al., 2009; Minshall and Standart, 2004). In contrast, Sweet et al. (2012) reported that Dhh1 affects elongation by slowing ribosome movement and targeting the mRNA to P bodies. Is it possible to reconcile these apparently conflicting observations? Generally, DDX6 proteins are very abundant (~10-fold molar excess over mRNAs in vertebrate cells) and have the potential to oligomerize along RNA in a sequence-

and ATP-independent manner (Ernault-Lange et al., 2012; Minshall and Standart, 2004). This form of RNA binding differs from the ATP-dependent mode generally required for RNA or RNP remodeling by DEAD-box ATPases/helicases (Henn et al., 2012). In yeast Dhh1, the two modes of RNA binding appear to involve different, though partially overlapping, surfaces, and RNA binding to the ATP-independent surface is mutually exclusive with Dhh1 binding to other proteins involved in mRNA translational repression and/or decapping (Sharif et al., 2013). These observations underscore the complexity of the DDX6 function in posttranscriptional regulation. They also raise the possibility that multiple DDX6 molecules can be recruited to mRNA, perhaps via sequential interactions with the CNOT1 MIF4G and cycles of ATP binding and hydrolysis, affecting rates of initiation or elongation. Such a scenario would explain observations that not only initiation, but also elongation, can be targeted by miRNAs (Fabian et al., 2010; Huntzinger and Izaurralde, 2011).

Consistent with the findings of others (Carroll et al., 2011; Coller and Parker, 2005; Minshall et al., 2009) and the postulated role of DDX6 in germ cells (reviewed by Minshall et al., 2009; Weston and Sommerville, 2006), our data support the DDX6 role as a translational repressor. In both tethering and miRNA reporter assays, repression mediated by DDX6 also applied to mRNAs not undergoing deadenylation, and analysis of mRNA levels showed that the DDX6 recruitment-dependent repression by CNOT1-R has a stronger effect on protein than on mRNA levels. Previous research has suggested that the repressive activity of DDX6 involves targeting eIF4E or eIF4E-transporter (eIF4E-T; eIF4E interacting protein, which blocks its association with eIF4G) (reviewed by Minshall et al., 2009; Weston and Sommerville, 2006). In fact, two-hybrid screens with human proteins revealed that DDX6 may interact with eIF4E-T (<http://interactome.dfci.harvard.edu/>). Moreover, depletion of eIF4E-T was recently found to partially alleviate miRNA repression (Kamenska et al., 2014).

By establishing the molecular mechanism of DDX6 recruitment and activation by CCR4-NOT, our data significantly extend previous observations that depletion of DDX6 in mammalian cells partially alleviates miRNA-mediated inhibition (Chu and Rana, 2006; Su et al., 2011). Moreover, our data argue against the recently proposed model that the repressive activity of CCR4-NOT1 occurs by the CNOT1 MIF4G domain specifically recruiting the initiation factor eIF4AII, which potentially interferes with the activity of its paralog, eIF4AI, in mRNA scanning for initiator AUG (Meijer et al., 2013). We found that the CNOT1 MIF4G domain does not interact with either eIF4AI or eIF4AII, in contrast to the MIF4G of eIF4G, which effectively associated with endogenous eIF4AI and eIF4AII but not with DDX6 (Figure 5D). In addition, the effect of mutating DDX6 to make its CNOT1 MIF4G-interacting surface similar to that of either eIF4AI or eIF4AII indicated that eIF4A proteins are incompatible with CNOT1 association.

The Central CNOT1 Region as a Hub of Protein-Protein Interactions

Our finding that the central CNOT1 region, CNOT1-R, associates directly with CNOT9 and DDX6 via the CN9BD and MIF4G domains, respectively, makes this region a hub of protein-protein interactions relevant to posttranscriptional regulation (Figure 7F).

MIF4G also associates directly with CAF1, and indirectly through CAF1, with CCR4 and TOB (Basquin et al., 2012; Petit et al., 2012 and references therein), while the region N-proximal to MIF4G interacts with TTP (Fabian et al., 2013). Our observation that CNOT1-R, but not the MIF4G domain alone, is able to induce repression of tethered mRNA suggests that additional CNOT1-R-mediated interactions may contribute to the repression. Regions flanking MIF4G may help recruit hPat1, Edc3, and RAP55/Tral, which are repressive proteins known to directly interact, in a mutually exclusive manner, with DDX6 (Sharif et al., 2013; Tritschler et al., 2009). Moreover, TTP was recently proposed to induce translational repression, possibly via recruitment of CNOT1 and DDX6 (Qi et al., 2012). DDX6 was also identified among proteins associating with EDD, a protein interacting with GW182 and implicated in miRNA repression (Su et al., 2011). In this context, it is also important to discuss the potential role of CAF1. Cooke et al. (2010) reported that tethering of CAF1 to mRNA reporters leads to cap-dependent, but deadenylation-independent, repression of translation in *Xenopus* oocytes. However, our findings (Figure S6E–S6H), together with results obtained with *Drosophila* S2 cells (Bawankar et al., 2013), indicate that CAF1 itself is unlikely to act as a translational inhibitor but exerts its effect by interacting with CNOT1, which in turn recruits DDX6 and possibly other inhibitory proteins. Similarly, the reported repression of translation by tethered CNOT9 (Bawankar et al., 2013) is at least partially due to its recruitment of CNOT1 (Figure 3D).

In summary, our experiments identify a set of interactions involved in miRNA repression and document their importance with structural information. They also reveal the complexity and redundancy of reactions leading to the repression. This is consistent with the general importance of the miRNA- and CCR4-NOT-mediated regulation of gene expression.

In this issue, Chen et al. (2014) report findings that are similar to the data described in our paper.

EXPERIMENTAL PROCEDURES

Pull-Down Assays and Western Blotting

GST pull-down assays were performed as described previously (Chekulaeva et al., 2011), using lysates from HEK293 cells transfected with plasmids expressing the indicated GST-tagged proteins. Lysates were treated with micrococcal nuclease (10 ng/ μ l) for 25 min at 25°C and then incubated with glutathione Sepharose beads (GE Healthcare) for 2 hr at 4°C. For western blotting analysis, proteins were separated by SDS-PAGE and detected using ECL (GE Healthcare) or SuperSignal West Femto Chemiluminescent Substrate (Thermo Scientific).

Northern Blotting

Northern blotting was performed as described previously (Chekulaeva et al., 2011). Briefly, 7 μ g of total RNA from transfected HEK293T cells were resolved in a denaturing formaldehyde agarose gel and transferred to Hybond-N+ membrane (GE Healthcare). RL- or FL-specific probes internally labeled with α -³²P]UTP were sequentially hybridized to the RNA on the membrane, and signals were detected using a GE Typhoon 9400 scanner. We have verified, by preparing calibration curves, that northern signals were within the linear range of quantification.

ATPase Assays

Assays were carried out in 20 μ l reactions containing 50 mM HEPES (pH 7.0), 50 mM KCl, 5 mM Mg acetate, 0.2 mg/ml polyuridylic acid (Sigma), 2 mM

dithiothreitol (DTT), and 0.1 mg/ml BSA. Final protein concentrations were 3 μ M for DDX6 and 4.5 μ M for the CNOT1 MIF4G fragments. α -[³²P]ATP was added to the final ATP concentration of 1 mM (final specific activity 50 μ Ci/ μ mol). Reaction mixtures were incubated at 37°C, and aliquots were collected at the indicated time points. Samples were resolved on Polyethylenimine Cellulose TLC Plates (Merck), in 300 mM of KH₂PO₄/K₂HPO₄ (pH 7.6).

miRNA Reporter and Tethering Assays

These assays were carried out as described (Chekulaeva et al., 2011) and as detailed in the Supplemental Experimental Procedures.

Protein Purification and Structure Determination

CNOT9 and DDX6 proteins, and CNOT1 and TNRC6C protein fragments, were obtained in a recombinant form by overexpression in *E. coli* and purification protocols detailed in the Supplemental Experimental Procedures. Crystallization was carried out by vapor diffusion using a nanoliter setup. All diffraction data were collected at 100 K at the Swiss Light Source (SLS), and the structure was determined using standard phasing experiments and software. The data collection and refinement statistics are summarized in Table 1.

ACCESSION NUMBERS

Structural data are deposited in the Protein Data Bank (PDB) under accession numbers 4CT4 (CNOT1 MIF4G-DDX6), 4CT5 (DDX6), 4CT6 (CNOT1 CN9BD-CNOT9), 4CT7 (CNOT1 CN9BD-CNOT9-W), and 4CV5 (CNOT1 CN9BD-Caf40).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2014.03.036>.

AUTHOR CONTRIBUTIONS

All cell-based and mutagenesis experiments were carried out by H.M. with help from A.A. Structural studies were done by J.B. (Figures 2C and 3B) or S.O. and J.B. (Figure 4). Experiments shown in Figures 2A and 2B were performed by M.C.-C., A.D., and M.N., and those in Figure 7D were performed by F.B. W.F., H.M., and E.C. coordinated the project and wrote the paper.

ACKNOWLEDGMENTS

We would like to thank the protein structure (Heinz Gut) and MS (Ragna Sack) facilities of the Friedrich Miescher Institute (FMI); Nancy Hynes and Francisca Maurer for help and discussions; the Max Planck Institute (MPI) Crystallization Facility for screening and optimization; the MPI Core Facility for MS analysis; and the beamline scientists at the SLS for excellent assistance with data collection. We also thank our colleagues for providing plasmids and antibodies. This work was supported by the FMI (to W.F.), Howard Hughes Medical Institute IECs (to M.N.), a European Research Council (ERC) Starting Grant (to A.D.), the Max-Planck Gesellschaft, the EU (ERC Advanced Investigator Grant and Marie Curie Initial Training Network RNPnet), and the Deutsche Forschungsgemeinschaft (SFB646, SFB1035, GRK1721, FOR1680, and ClPSM; to E.C.).

Received: December 26, 2013

Revised: March 4, 2014

Accepted: March 18, 2014

Published: April 24, 2014

REFERENCES

Basquin, J., Roudko, V.V., Rode, M., Basquin, C., Séraphin, B., and Conti, E. (2012). Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Mol. Cell* 48, 207–218.

- Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S., and Izaurralde, E. (2013). NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol.* **10**, 228–244.
- Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* **336**, 233–237.
- Béthune, J., Artus-Revel, C.G., and Filipowicz, W. (2012). Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Rep.* **13**, 716–723.
- Bhaskar, V., Roudko, V., Basquin, J., Sharma, K., Urlaub, H., Séraphin, B., and Conti, E. (2013). Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex. *Nat. Struct. Mol. Biol.* **20**, 1281–1288.
- Boland, A., Chen, Y., Raisch, T., Jonas, S., Kuzuoğlu-Öztürk, D., Wohlbold, L., Weichenrieder, O., and Izaurralde, E. (2013). Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat. Struct. Mol. Biol.* **20**, 1289–1297.
- Braun, J.E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44**, 120–133.
- Buchwald, G., Schüssler, S., Basquin, C., Le Hir, H., and Conti, E. (2013). Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. *Proc. Natl. Acad. Sci. USA* **110**, E4611–E4618.
- Carroll, J.S., Munchel, S.E., and Weis, K. (2011). The DEXD/H box ATPase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics. *J. Cell Biol.* **194**, 527–537.
- Chekulaeva, M., Parker, R., and Filipowicz, W. (2010). The GW/WG repeats of *Drosophila* GW182 function as effector motifs for miRNA-mediated repression. *Nucleic Acids Res.* **38**, 6673–6683.
- Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226.
- Chen, C.Y., Zheng, D., Xia, Z., and Shyu, A.B. (2009). Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nat. Struct. Mol. Biol.* **16**, 1160–1166.
- Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing. *Mol. Cell* **54**. Published online April 24, 2014. <http://dx.doi.org/10.1016/j.molcel.2014.03.034>.
- Cheng, Z., Collier, J., Parker, R., and Song, H. (2005). Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* **11**, 1258–1270.
- Christie, M., Boland, A., Huntzinger, E., Weichenrieder, O., and Izaurralde, E. (2013). Structure of the PAN3 pseudokinase reveals the basis for interactions with the PAN2 deadenylase and the GW182 proteins. *Mol. Cell* **51**, 360–373.
- Chu, C.Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol.* **4**, e210.
- Collart, M.A., and Panasenko, O.O. (2012). The Ccr4-not complex. *Gene* **492**, 42–53.
- Coller, J., and Parker, R. (2005). General translational repression by activators of mRNA decapping. *Cell* **122**, 875–886.
- Cooke, A., Prigge, A., and Wickens, M. (2010). Translational repression by deadenylases. *J. Biol. Chem.* **285**, 28506–28513.
- Djuranovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* **336**, 237–240.
- Dutta, A., Zheng, S., Jain, D., Cameron, C.E., and Reese, J.C. (2011). Intermolecular interactions within the abundant DEAD-box protein Dhh1 regulate its activity in vivo. *J. Biol. Chem.* **286**, 27454–27470.
- Emoult-Lange, M., Baconnais, S., Harper, M., Minshall, N., Souquere, S., Boudier, T., Bénard, M., Andrey, P., Pierron, G., Kress, M., et al. (2012). Multiple binding of repressed mRNAs by the P-body protein Rck/p54. *RNA* **18**, 1702–1715.
- Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., and Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. *RNA* **15**, 21–32.
- Fabian, M.R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J.T., Svitkin, Y.V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J.A., et al. (2009). Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Mol. Cell* **35**, 868–880.
- Fabian, M.R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* **79**, 351–379.
- Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., and Sonenberg, N. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217.
- Fabian, M.R., Frank, F., Rouya, C., Siddiqui, N., Lai, W.S., Karetnikov, A., Blackshear, P.J., Nagar, B., and Sonenberg, N. (2013). Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat. Struct. Mol. Biol.* **20**, 735–739.
- Fukaya, T., and Tomari, Y. (2012). MicroRNAs mediate gene silencing via multiple different pathways in *Drosophila*. *Mol. Cell* **48**, 825–836.
- Garces, R.G., Gillon, W., and Pai, E.F. (2007). Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci.* **16**, 176–188.
- Henn, A., Bradley, M.J., and De La Cruz, E.M. (2012). ATP utilization and RNA conformational rearrangement by DEAD-box proteins. *Annu Rev Biophys* **41**, 247–267.
- Hinnebusch, A.G. (2011). Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol. Mol. Biol. Rev.* **75**, 434–467.
- Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat. Rev. Genet.* **12**, 99–110.
- Huntzinger, E., Kuzuoğlu-Öztürk, D., Braun, J.E., Eulalio, A., Wohlbold, L., and Izaurralde, E. (2013). The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets. *Nucleic Acids Res.* **41**, 978–994.
- Kamenska, A., Lu, W.T., Kubacka, D., Broomhead, H., Minshall, N., Bushell, M., and Standart, N. (2014). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. *Nucleic Acids Res.* **42**, 3298–3313.
- Maillet, L., and Collart, M.A. (2002). Interaction between Not1p, a component of the Ccr4-not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. *J. Biol. Chem.* **277**, 2835–2842.
- Meijer, H.A., Kong, Y.W., Lu, W.T., Wilczynska, A., Spriggs, R.V., Robinson, S.W., Godfrey, J.D., Willis, A.E., and Bushell, M. (2013). Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. *Science* **340**, 82–85.
- Minshall, N., and Standart, N. (2004). The active form of Xp54 RNA helicase in translational repression is an RNA-mediated oligomer. *Nucleic Acids Res.* **32**, 1325–1334.
- 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.
- Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* **472**, 238–242.
- Parker, R., and Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635–646.
- Petit, A.P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2012). The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res.* **40**, 11058–11072.

- Piao, X., Zhang, X., Wu, L., and Belasco, J.G. (2010). CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells. *Mol. Cell. Biol.* *30*, 1486–1494.
- Qi, M.Y., Wang, Z.Z., Zhang, Z., Shao, Q., Zeng, A., Li, X.Q., Li, W.Q., Wang, C., Tian, F.J., Li, Q., et al. (2012). AU-rich-element-dependent translation repression requires the cooperation of tristetraprolin and RCK/P54. *Mol. Cell. Biol.* *32*, 913–928.
- Schirle, N.T., and MacRae, I.J. (2012). The crystal structure of human Argonaute2. *Science* *336*, 1037–1040.
- Schütz, P., Bumann, M., Oberholzer, A.E., Bieniossek, C., Trachsel, H., Altmann, M., and Baumann, U. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc. Natl. Acad. Sci. USA* *105*, 9564–9569.
- Sharif, H., Ozgur, S., Sharma, K., Basquin, C., Urlaub, H., and Conti, E. (2013). Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Res.* *41*, 8377–8390.
- Su, H., Meng, S., Lu, Y., Trombly, M.I., Chen, J., Lin, C., Turk, A., and Wang, X. (2011). Mammalian hyperplastic discs homolog EDD regulates miRNA-mediated gene silencing. *Mol. Cell* *43*, 97–109.
- Sweet, T., Kovalak, C., and Collier, J. (2012). The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement. *PLoS Biol.* *10*, e1001342.
- Tritschler, F., Braun, J.E., Eulalio, A., Truffault, V., Izaurralde, E., and Weichenrieder, O. (2009). Structural basis for the mutually exclusive anchoring of P body components EDC3 and Tral to the DEAD box protein DDX6/Me31B. *Mol. Cell* *33*, 661–668.
- Van Etten, J., Schagat, T.L., Hrit, J., Weidmann, C.A., Brumbaugh, J., Coon, J.J., and Goldstrohm, A.C. (2012). Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs. *J. Biol. Chem.* *287*, 36370–36383.
- Weston, A., and Sommerville, J. (2006). Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation. *Nucleic Acids Res.* *34*, 3082–3094.

Molecular Cell, Volume 54

Supplemental Information

**Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in
MicroRNA Repression**

Hansruedi Mathys, Jérôme Basquin, Sevim Ozgur, Mariusz Czarnocki-Cieciura, Fabien Bonneau, Aafke Aartse, Andrzej Dziembowski, Marcin Nowotny, Elena Conti, and Witold Filipowicz

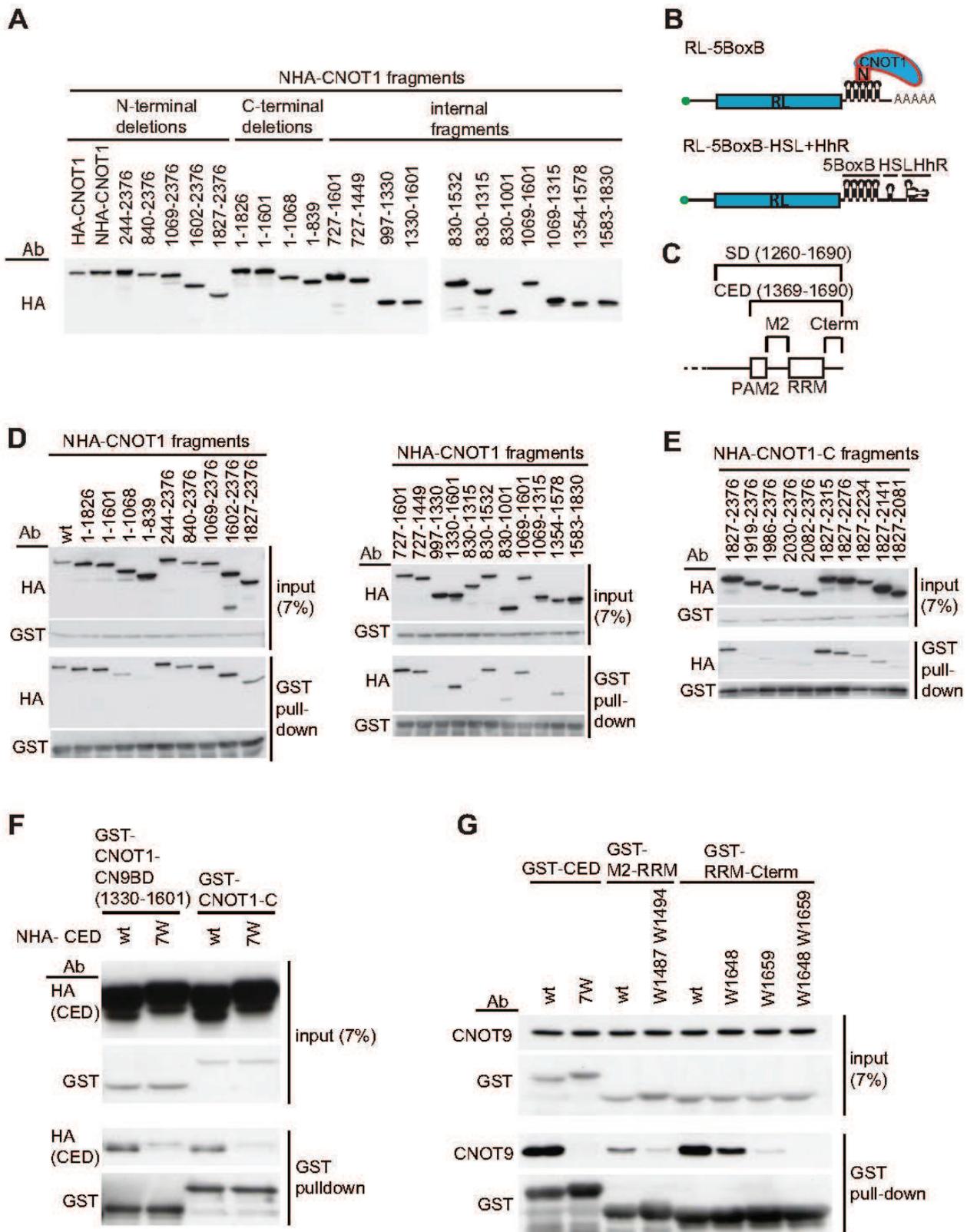


Figure S1

Figure S1. Representative western analyses for pull-downs summarized in Figure 1A and activity assays summarized in Figure 1B and C, related to Figure 1

(A) Western analysis of different NHA-CNOT1 fragments (schematically shown in Figure 1A) expressed in HEK293 cells and tested for repression of mRNA function in Figures 1B and C. To obtain approximately equal expression of individual protein fragments, different concentrations of plasmids were used for transfections.

(B) Scheme of poly(A)⁺ and poly(A)⁻ reporters used in tethering assays. HSL, histone stem-loop; HhR, hammer-head ribozyme; BoxB, an RNA hairpin recognized by the phage λ N peptide.

(C) Scheme of the TNRC6C repressive regions SD and CED used for experiments shown in other panels. M2 and C-term, W-motifs-containing non-structured repressive regions of the CED (Chekulaeva et al., 2011). RRM, RNA recognition motif; PAM2, the poly(A) binding protein (PABP)-interacting motif 2.

(D) Representative western analyses of pull-downs performed with GST fusions of either TNRC6C SD (shown in the figure) or TNRC6C CED (see panel C) and specified NHA-tagged fragments of CNOT1. Identity of pulled-down CNOT1 fragments, representing either terminal deletions (left panel) or internal fragments (right panel) is specified at the top. Lower western panels in both input and pull-down pictures visualize expression levels of the GST-SD fusion. The pull-down data are summarized in Figure 1A. Prior to performing the pull-downs, all samples used in this and other experiments were treated with micrococcal nuclease (see Supplemental Experimental Procedures). Regarding the SD interaction with the CNOT1 TTP region [residues 830-1001 (right part of panel D)], note that specific deletion of the TTP region from the CNOT1 fragment 1-1068 eliminates its interaction with SD (compare lanes 1-1068 and 1-839 in left part of panel D), further supporting role of the TTP region in the SD binding (see also Discussion).

(E) Pull-down experiments performed with additional N- and C-terminal deletions of the CNOT1-C fragment (residues 1827-2376; see Figure 1A) and the GST-CED, documenting that integrity of this fragment is important for optimal interaction with the TNRC6C SD/CED.

(F) The CNOT1 CN9BD and CNOT1-C interactions with the TNRC6 CED, measured in direction opposite to that shown in panel D, depend on W-motifs. wt and 7W, NHA fusions of wt CED and CED bearing seven W→A mutations, respectively. Pull-downs were performed with GST fusions of CNOT1 CN9BD and CNOT1-C. Antibodies used for westerns are indicated.

(G) Interaction of the CED sub-fragments, M2-RRM and RRM-Cterm (see panel C, and Chekulaeva et al., 2011) with CNOT9 is affected by mutation of even one or two Trp residues. wt, fragments with wild-type sequence.

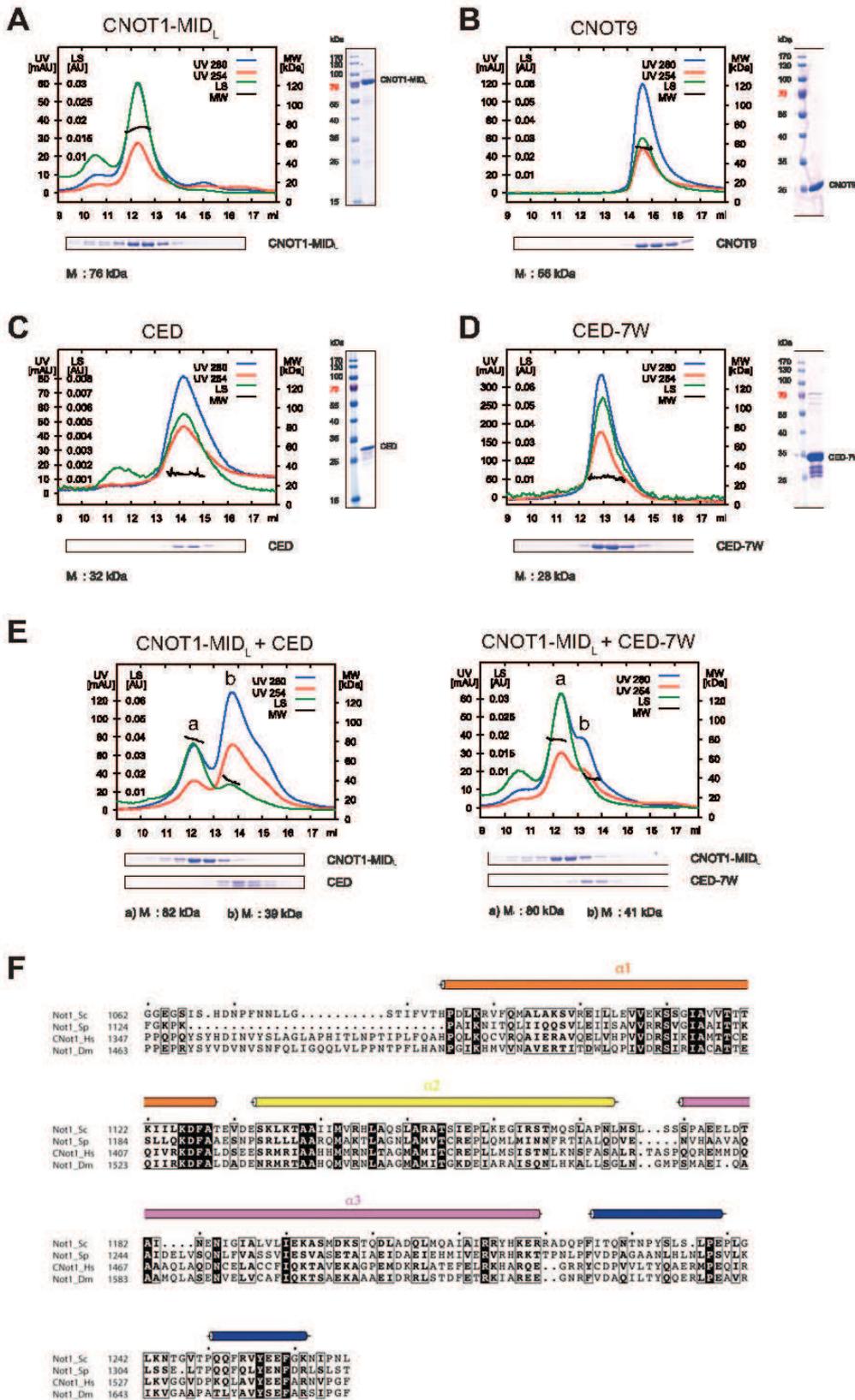


Figure S2

Figure S2. CNOT9 mediates the TNRC6C-CED interaction with CNOT1-CN9BD, related to Figure 2

(A-D) Characterization of recombinant purified proteins used for analysis of complexes shown in Figures 2A, B and S2E. Shown are the gel filtration and electrophoresis profiles assessing purity and molecular weights, as estimated by multi-angle static light scattering (MALS), of recombinant proteins or their fragments: CNOT1-MID_L (residues 1092-1827; panel A), CNOT9 (residues 18-285; panel B), CED (panel C), and CED-7W mutant (panel D). Theoretical M_w for CNOT9, CED, CED-7W, and CNOT1-MID_L are 30.5, 28.9, 28.2, and 80.5 kDa, respectively. The CED and CED-7W used in experiments shown in this figure and also in Figure 2A,B correspond to TNRC6C residues 1417-1690 encompassing M2-RRM-Cterm region (see Figure S1C). Note abnormal elution of the CED-7W during gel filtration and the dimerization of CNOT9.

(E) Size-exclusion chromatography demonstrating inability of the TNRC6C CED (or its 7W mutant) to form a complex with the CN9CD-containing middle fragment of CNOT1 (CNOT1-MID_L).

Analyzed proteins are indicated at the top of each panel, and molecular weights of complexes or individual proteins are at the bottom, below the gels showing distribution of individual proteins along the column. Gel electrophoresis profiles of analyzed input proteins (A-D) are shown on the right, with molecular weights of size markers (in kDa) indicated. LS, light scattering.

(F) Structure-based sequence alignment of CNOT1-CN9BD. Secondary-structure elements are shown above the sequences (with the different structural elements colored from orange to blue as in Figure 2C), which include orthologs from *S. cerevisiae* (*Sc*), *Schizosaccharomyces pombe* (*Sp*), *Homo sapiens* (*Hs*) and *D. melanogaster* (*Dm*). Two short red helices depicted in Figure 2C are not shown due to low sequence conservation among species.

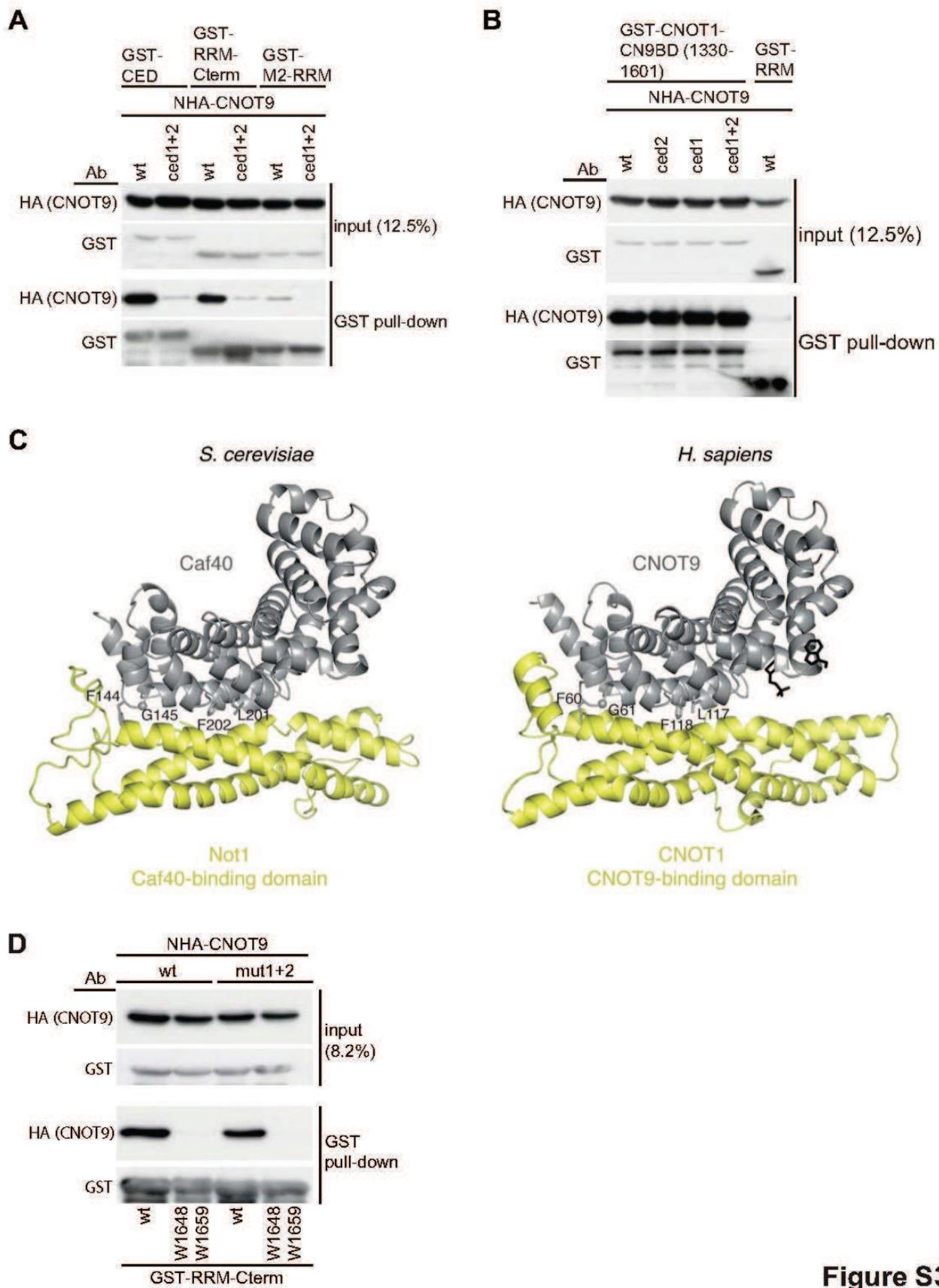


Figure S3

Figure S3. Characterization of the CNOT9 Trp/MES-binding pockets and structural analysis of the CNOT1 CN9BD in a complex with CNOT9/Caf40, related to Figure 3

(A) CNOT9 interacts with the RRM-Cterm CED subfragment more strongly than with M2-RRM (see also Figure S1G). Both interactions are affected by mutations in Trp and MES pockets (mutant CNOT9_{ced1+2}).

(B) Mutations in individual pockets, either Trp (R244E, A248L; mutant CNOT9_{ced1}) or MES (R205D, H208D; CNOT9_{ced2}), and also the combined mutation CNOT9_{ced1+2} have no effect on the CNOT1 CN9BD - CNOT9 interaction. GST-RRM was used as a control.

(C) Comparison of structures of the yeast NOT1 CN9BD – Caf40 (orthologue of human CNOT9) complex (left panel) and the human CNOT1 CN9BD – CNOT9 complex (right panel). Caf40/CNOT9 is in grey and the CN9BD in yellow). Conserved residues at protein interfaces, which have been mutated in CNOT9 are indicated.

(D) The CNOT9_{mut1+2} mutant showing compromised interaction with CNOT1 retains the W-motif dependent interaction with the RRM-Cterm fragment of the CED. NHA-CNOT9, either wt or mutant form, was co-expressed in HEK293 cells with GST fusions of RRM-Cterm, either wt or W-motifs mutant.

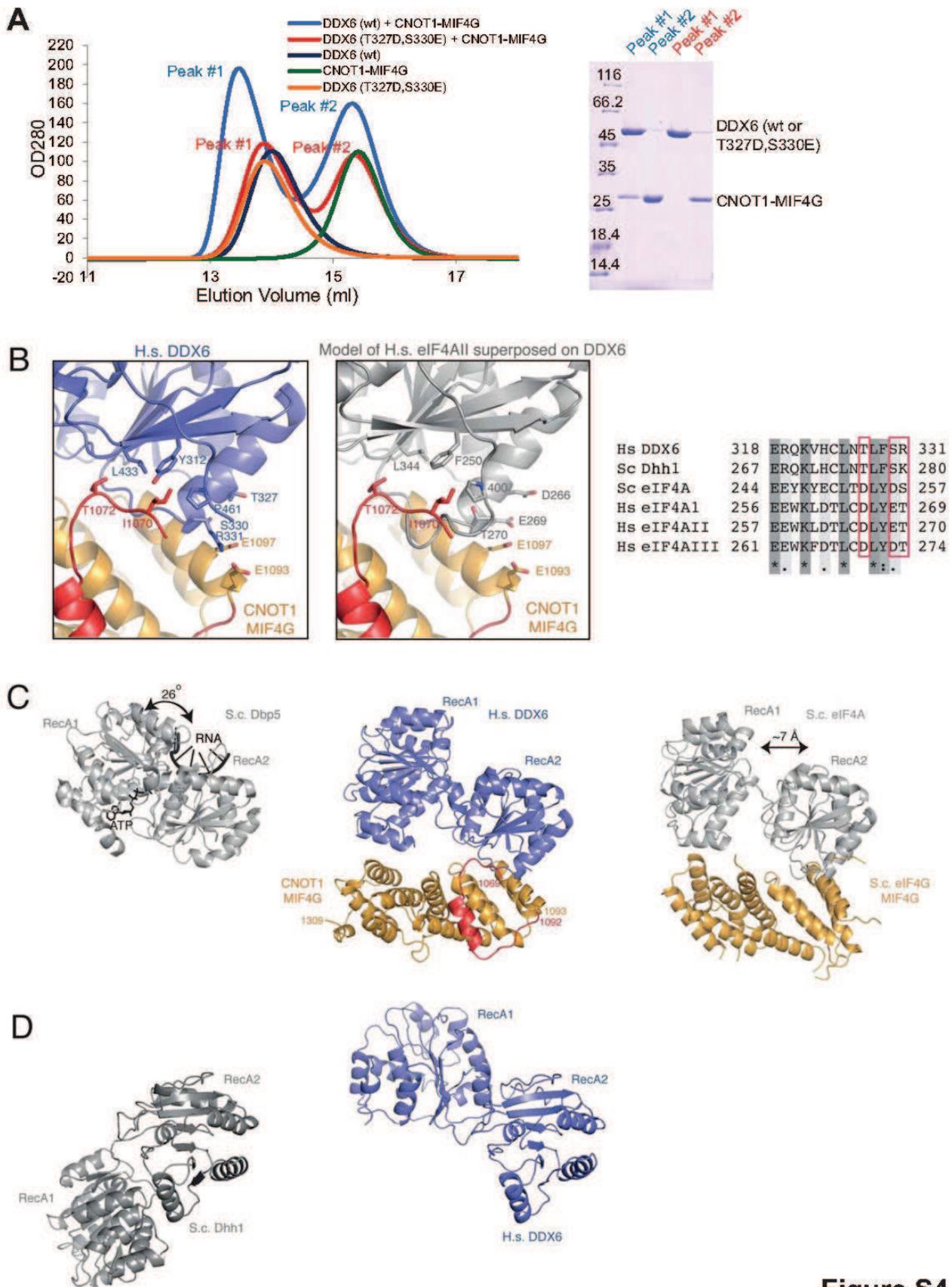


Figure S4

Figure S4. Structural analysis of the interaction between DDX6 and the CNOT1 MIF4G domain, related to Figure 4

(A) DDX6-MIF4G complex formation. Left panel shows size-exclusion chromatography profiles (from a Superdex75 column), with different colors corresponding to the different input proteins (indicated). The right panel is a 12% SDS-PAGE gel with the peak fractions of the DDX6 + CNOT1 MIF4G mixture (in blue) and the peak fractions of the DDX6 (T327D, S330E) + CNOT1 MIF4G mixture (in red). This DDX6 mutant is detailed in panel B. The CNOT1 MIF4G domain used for experiments shown in this figure corresponds to residues 1063-1314.

(B) Specificity of the DDX6-CNOT1 interaction. The structure of the CNOT1 MIF4G bound to DDX6 is shown on the left, zoomed in at the RecA2-binding site and colored as in Figure 4. In the middle panel, the structure of human eIF4AII has been modeled on DDX6 in the complex. The modeling shows that eIF4AII would approach a negatively-charged surface of CNOT1 (Glu1093 and Glu1097) with negatively-charged residues (Asp266 and Glu269). In the case of DDX6-CNOT1 (left panel), DDX6 approaches this CNOT1 surface with Thr327 and Ser330, and in addition with a positively-charged residue (Arg331). The right panel shows a sequence alignment of these interacting residues of DDX6 and how they differ in eIF4A paralogues (highlighted in red).

(C) Difference in the conformation of the two RecA domains of DDX6 in the CNOT1 complex (middle panel) as compared to the yeast eIF4A-eIF4G structure ((Schutz et al., 2008); right panel) and to a typical DEAD-box protein in the RNA-ATP-bound state (Dbp5; (von Moeller et al., 2009); left panel). The structures are viewed in the same orientation after superposition of their RecA2 domains. The arrows indicate the relative rotations or transitions of the RecA1 domains with respect to the DDX6-CNOT1 structure in the middle panel.

(D) Difference in the conformation of the two RecA domains of apo DDX6 (right panel, in blue) and apo Dhh1 ((Cheng et al., 2005); left panel, in gray). The structures are viewed in the same orientation after superposition of their RecA2 domains (which are rotated by 90° around a horizontal axis with respect to the views in panel C).

A**Mutations introduced to CNOT1 MIF4G**

MIF4G_{hel1} F1101A, N1105A, K1114A
MIF4G_{hel2} E1142R, N1144A, F1145A
MIF4G_{hel3} K1276A, F1281A, E1284A
MIF4G_{caf1} F1252E, P1257Y

Combination of MIF4G_{hel} mutations: e.g., MIF4G_{hel1+2} when combining MIF4G_{hel1} and MIF4G_{hel2} mutations

Combination of MIF4G_{caf1} and MIF4G_{hel} mutations: e.g., MIF4G_{hel1/caf1} when combining MIF4G_{hel1} and MIF4G_{caf1} mutations

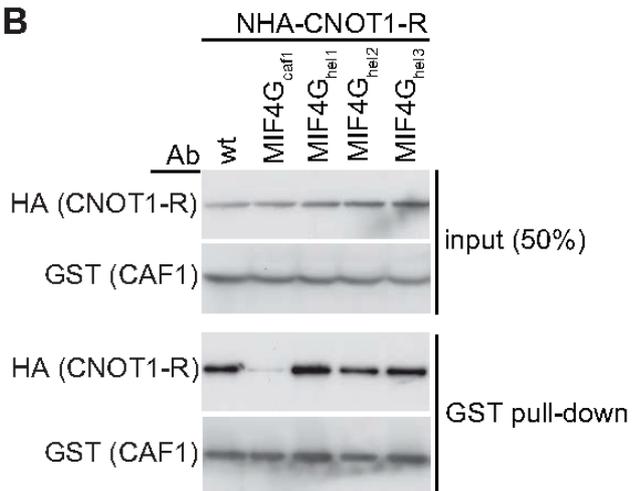
B**Figure S5**

Figure S5. Mutations introduced to the CNOT1 MIF4G domain do not affect its interaction with CAF1, related to Figure 5

(A) Listing of all investigated mutations introduced to the MIF4G domain of CNOT1.

(B) Mutations introduced to the concave surface of MIF4G (listed in panel A) have no effect on the CNOT1-R interaction with CAF1. CNOT-R MIF4G wt and its mutants were expressed in HEK293 cells as HA fusions and CAF1 as a GST fusion.

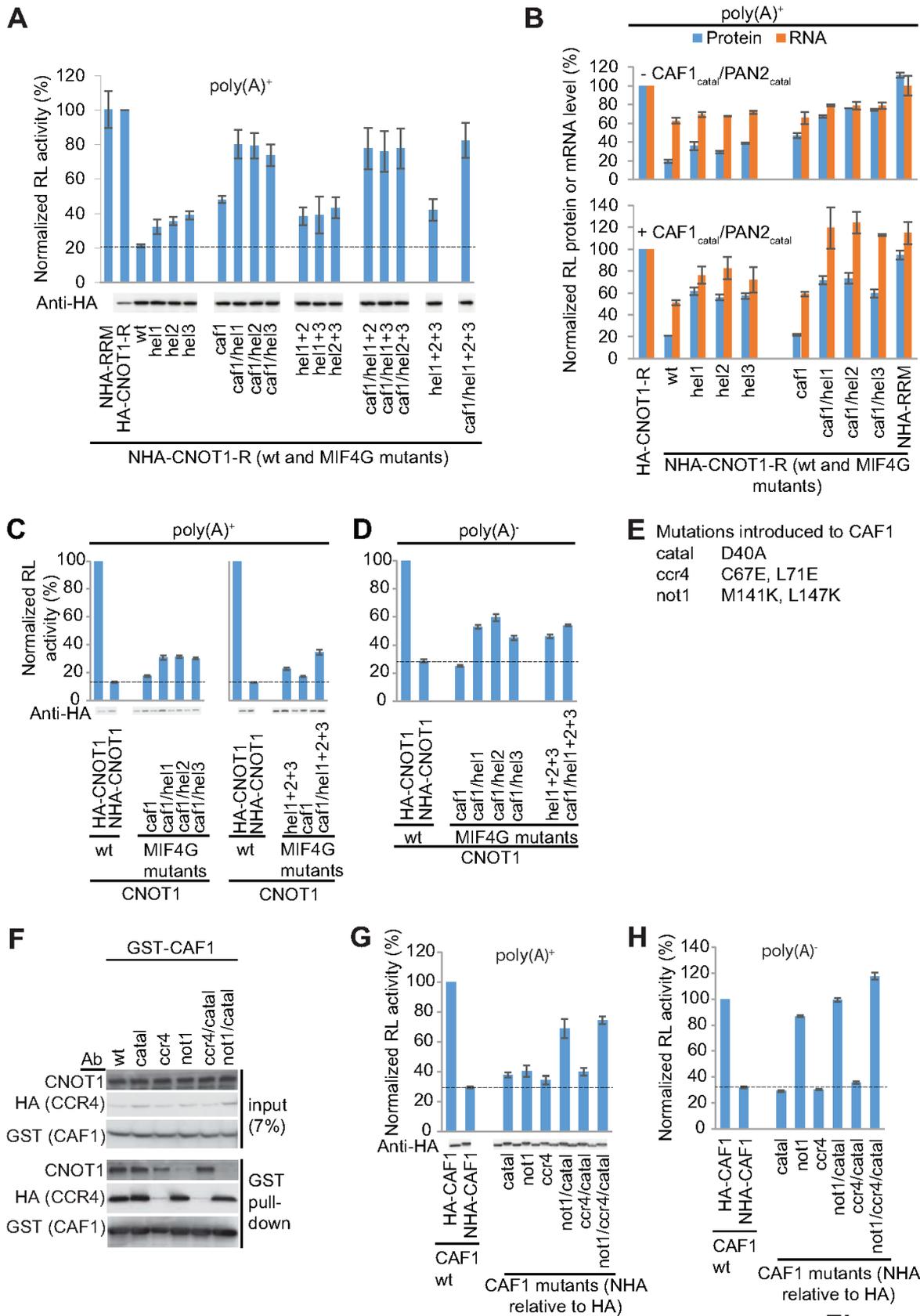


Figure S6

Figure S6. Mutational analysis of the role of the CNOT1 MIF4G domain and CAF1 in repression of mRNA function, related to Figure 6

(A) Effect of mutations in the CNOT1-R MIF4G domain on repression of poly(A)⁺ RL-5BoxB reporter. HEK293 cells were transfected with RL and control FL reporter plasmids, and plasmids expressing indicated CNOT1-R MIF4G mutants (mutated amino acids are listed in Figure S5A). Values represent means +/- SEM (n = 3-6) with a value for HA-CNOT1-R taken as 100%. Western blots documenting similar expression of all mutants are shown below. Bands originate from the same gel but are aligned with individual groups of mutants. Broken line marks activity in the presence of wt CNOT1-R.

(B) Summary of the data presented in Figures 6A, B, D, and E, showing comparison of the effects of different mutations in the CNOT1-R MIF4G domain on protein expression (shown in blue) and mRNA levels [poly(A)⁺ RL-5BoxB reporter; shown in orange]. In lower part, HEK293 cells also expressed PAN2_{catal} and CAF1_{catal}. Protein values are means +/- SEM (n = 3) with the value for HA-CNOT1-R at 100%. RNA values, corrected for co-expressed FL mRNA, represent means +/- range (n = 2).

(C and D) Mutations in the CNOT1 MIF4G domain affect repression of poly(A)⁺ (panel C) and poly(A)⁻ (panel D) reporters also when present in the context of full length CNOT1 protein. Activity of wt or mutated CNOT1 to repress mRNA function was measured in tethering assay. Values represent means +/- SEM (n = 3-5) with a value for HA-CNOT1 taken as 100%. Activity of each NHA-fusion mutant was corrected for that of the HA-fusion mutant. Western analysis measuring expression of each CNOT1 mutant in either HA (left protein band) or NHA (right band) version is shown below the activity bars in panel C. Note that expression of HA-tagged proteins is slightly lower than NHA-tagged ones. For poly(A)⁺ reporter (C) MIF4G_{hel1+2+3} mutation showed ~1.8-fold relieve of repression; its combination with MIF4G_{caf1} mutation had a stronger (~2.5-fold) effect. Effect of individual MIF4G_{hel} mutations and triple MIF4G_{hel1+2+3} mutant was similar, consistent with the data obtained for the CNOT1-R protein fragment (see panel A). Findings for the poly(A)⁻ reporter (panel D) are similar for those seen with the poly(A)⁺ reporter, with the exception that the MIF4G_{caf1} mutation had no contribution to the effect.

(E) Summary of mutations introduced to CAF1.

(F) The CAF1 protein-protein interaction mutations or combinations thereof with the D40A catalytic mutation (CAF1_{catal}) eliminate only interactions with the expected protein partners.

(G and H) Analysis of different CAF1 mutants in tethering assays with poly(A)⁺ (panel G) and poly(A)⁻ (panel H) reporters suggests that CAF1 *per se* is not a translation inhibitor but exerts its repressive function by recruiting CNOT1. Tethering of wt CAF1 or CAF1_{catal} repressed activity of

both poly(A)⁺ and poly(A)⁻ reporters. The CAF1_{ccr4} mutation, either alone or in combination with CAF1_{catal} (mutant CAF1_{ccr4/catal}) also repressed both types of reporters. The CAF1_{not1} mutation alone relieved the repression of poly(A)⁺ RNA only marginally but its combination with CAF1_{catal} (CAF1_{not1/catal}) or CAF1_{catal} and CAF1_{ccr4} (CAF1_{not1/ccr4/catal}) resulted in a marked relief of the reporter repression. In the case of the poly(A)⁻ reporter, the CAF1_{not1} mutation alone was sufficient to alleviate the repression, the effect slightly exacerbated by combining it with other mutations. Values are means +/- SEM (n = 3-6).

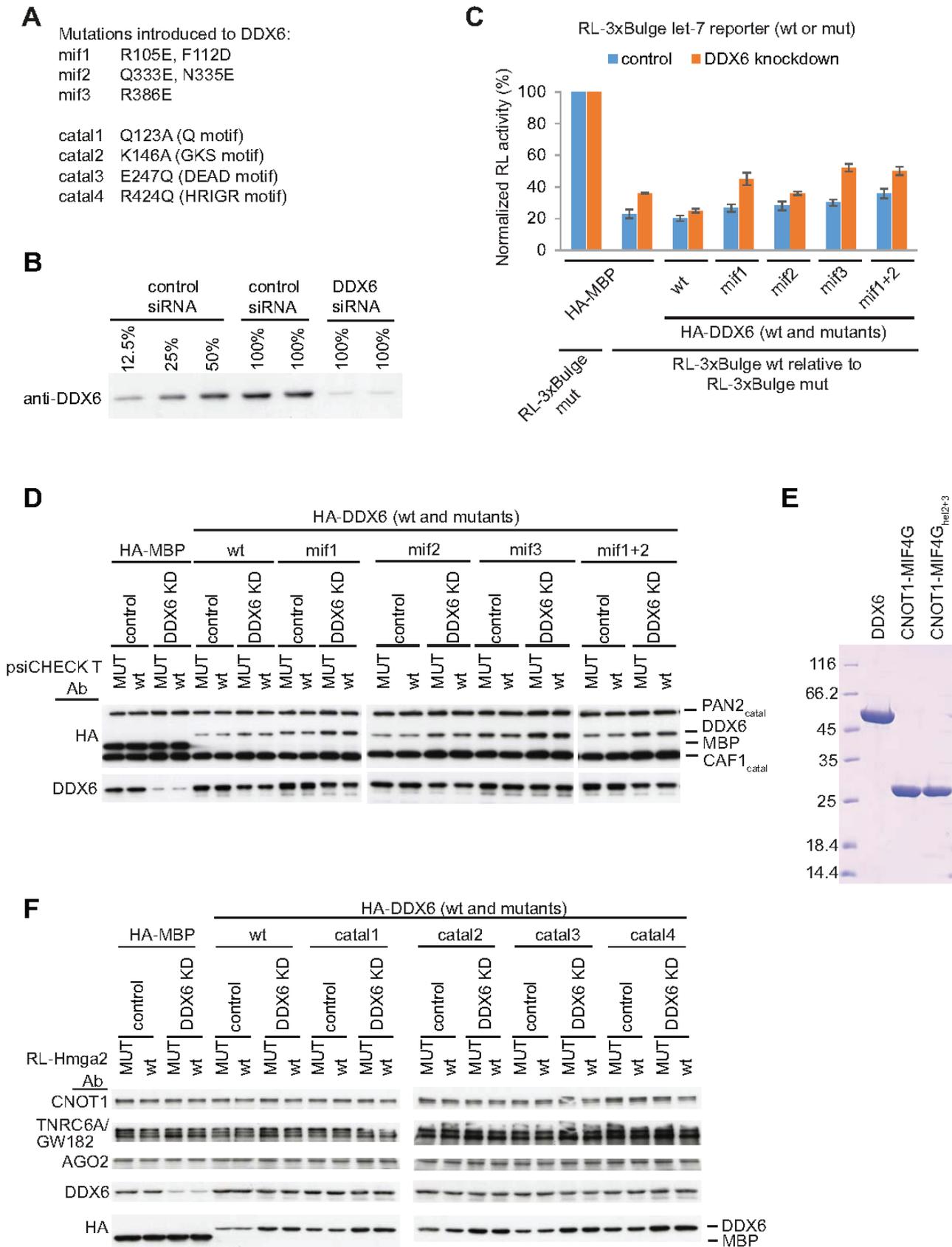


Figure S7

Table S1 Mass spectrometry analysis of GST-CNOT1-R pull-down, related to Figure 5
 Identified proteins are listed along with the number of assigned spectra (Exclusive Spectrum Count).

#	Accession Number	Molecular Weight	Exclusive Spectrum Count
1	CNOT1_HUMAN	267 kDa	742
2	EF1A1_HUMAN/EF1A3_HUMAN	50 kDa	299
3	EF1G_HUMAN	50 kDa	248
4	HSP7C_HUMAN	71 kDa	128
5	HSP71_HUMAN	70 kDa	119
6	GRP78_HUMAN	72 kDa	90
7	TBB5_HUMAN	50 kDa	80
8	TBA1B_HUMAN	50 kDa	80
9	HELZ_HUMAN	219 kDa	64
10	CBR1_HUMAN	30 kDa	52
11	IMMT_HUMAN	84 kDa	36
12	ADT2_HUMAN	33 kDa	36
13	EF1B_HUMAN	25 kDa	31
14	EF1A2_HUMAN	50 kDa	29
15	RN219_HUMAN	81 kDa	28
16	PYR1_HUMAN	243 kDa	28
17	PRKDC_HUMAN	469 kDa	27
18	NUCL_HUMAN	77 kDa	22
19	CLH1_HUMAN	192 kDa	19
20	SYCC_HUMAN	85 kDa	19
21	DDX6_HUMAN	54 kDa	18
22	BASI_HUMAN	42 kDa	16
23	RLA2_HUMAN	12 kDa	16
24	DYHC1_HUMAN	532 kDa	15
25	HS74L_HUMAN	95 kDa	15
26	FAS_HUMAN	273 kDa	14
27	AT1A1_HUMAN	113 kDa	14
28	HS105_HUMAN	97 kDa	14
29	RLA0_HUMAN	34 kDa	14
30	ACTB_HUMAN/ACTG_HUMAN	42 kDa	13
31	HS90B_HUMAN	83 kDa	13
32	ATPA_HUMAN	60 kDa	12
33	SYEP_HUMAN	171 kDa	12
34	HNRH1_HUMAN	49 kDa	11
35	ROA1_HUMAN	39 kDa	11
36	HSP74_HUMAN	94 kDa	11
37	SAM50_HUMAN	52 kDa	11
38	RL3_HUMAN	46 kDa	10
39	ATPB_HUMAN	57 kDa	10
40	RS15A_HUMAN	15 kDa	10
41	ILF2_HUMAN	43 kDa	10
42	RS19_HUMAN	16 kDa	10
43	LANC1_HUMAN	45 kDa	9
44	RL12_HUMAN	18 kDa	9
45	RCD1_HUMAN	34 kDa	9
46	RS18_HUMAN	18 kDa	9
47	CH60_HUMAN	61 kDa	8
48	PABP1_HUMAN	71 kDa	8
49	HAX1_HUMAN	32 kDa	7
50	TIM50_HUMAN	40 kDa	7
51	RL28_HUMAN	16 kDa	6
52	RUVB2_HUMAN	51 kDa	6
53	GRP75_HUMAN	74 kDa	5
54	RUVB1_HUMAN	50 kDa	5

54 proteins were identified with a peptide FDR of 1.75% and a protein FDR of 0%.

Figure S7. Mutations in DDX6 which prevent its interaction with the CNOT1 MIF4G domain compromise the DDX6 inhibitory function in miRNA repression, related to Figure 7

(A) Listing of all investigated mutations introduced to DDX6.

(B) Western blot analysis documenting DDX6 knock-down (~90%) efficiency. 100% lanes represent duplicates.

(C) DDX6_{mif} mutations mitigate repression of the RL-3xBulge reporter responding to let-7 miRNA. Values are means +/- SEM (n = 3-4).

(D) Western blot performed with the cell lysate samples analyzed in Figure 7C. The western analysis documents comparable expression of different HA-tagged DDX6 mutants (visualized with HA antibody) and their resistance (in contrast to endogenous DDX6 visualized with the DDX6-antibody; bottom lane) to RNAi. Note that the DDX6 antibody visualizes both the endogenous and the HA-tagged DDX6. Blots also visualize expression of PAN2_{catal} and CAF1_{catal}, and of maltose binding protein (MBP) used as a control. Identity of reporters (psiCHECK T wt or mut), siRNAs (control or DDX6 KD), and DDX6 mutants is indicated at the top.

(E) Purity of proteins used in the ATPase assay as determined by SDS-PAGE.

(F) Western blot performed with the cell lysate samples analyzed in Figure 7E. The western analysis documents comparable expression of different HA-tagged DDX6 mutants (visualized with HA antibody) and that the DDX6 depletion had no effect on levels of CNOT1, AGO2, and TNRC6A/GW182 proteins.

Table S1 Mass spectrometry analysis of GST-CNOT1-R pull-down, related to Figure 5
 Identified proteins are listed along with the number of assigned spectra (Exclusive Spectrum Count).

#	Accession Number	Molecular Weight	Exclusive Spectrum Count
1	CNOT1_HUMAN	267 kDa	742
2	EF1A1_HUMAN/EF1A3_HUMAN	50 kDa	299
3	EF1G_HUMAN	50 kDa	248
4	HSP7C_HUMAN	71 kDa	128
5	HSP71_HUMAN	70 kDa	119
6	GRP78_HUMAN	72 kDa	90
7	TBB5_HUMAN	50 kDa	80
8	TBA1B_HUMAN	50 kDa	80
9	HELZ_HUMAN	219 kDa	64
10	CBR1_HUMAN	30 kDa	52
11	IMMT_HUMAN	84 kDa	36
12	ADT2_HUMAN	33 kDa	36
13	EF1B_HUMAN	25 kDa	31
14	EF1A2_HUMAN	50 kDa	29
15	RN219_HUMAN	81 kDa	28
16	PYR1_HUMAN	243 kDa	28
17	PRKDC_HUMAN	469 kDa	27
18	NUCL_HUMAN	77 kDa	22
19	CLH1_HUMAN	192 kDa	19
20	SYCC_HUMAN	85 kDa	19
21	DDX6_HUMAN	54 kDa	18
22	BASI_HUMAN	42 kDa	16
23	RLA2_HUMAN	12 kDa	16
24	DYHC1_HUMAN	532 kDa	15
25	HS74L_HUMAN	95 kDa	15
26	FAS_HUMAN	273 kDa	14
27	AT1A1_HUMAN	113 kDa	14
28	HS105_HUMAN	97 kDa	14
29	RLA0_HUMAN	34 kDa	14
30	ACTB_HUMAN/ACTG_HUMAN	42 kDa	13
31	HS90B_HUMAN	83 kDa	13
32	ATPA_HUMAN	60 kDa	12
33	SYEP_HUMAN	171 kDa	12
34	HNRH1_HUMAN	49 kDa	11
35	ROA1_HUMAN	39 kDa	11
36	HSP74_HUMAN	94 kDa	11
37	SAM50_HUMAN	52 kDa	11
38	RL3_HUMAN	46 kDa	10
39	ATPB_HUMAN	57 kDa	10
40	RS15A_HUMAN	15 kDa	10
41	ILF2_HUMAN	43 kDa	10
42	RS19_HUMAN	16 kDa	10
43	LANC1_HUMAN	45 kDa	9
44	RL12_HUMAN	18 kDa	9
45	RCD1_HUMAN	34 kDa	9
46	RS18_HUMAN	18 kDa	9
47	CH60_HUMAN	61 kDa	8
48	PABP1_HUMAN	71 kDa	8
49	HAX1_HUMAN	32 kDa	7
50	TIM50_HUMAN	40 kDa	7
51	RL28_HUMAN	16 kDa	6
52	RUVB2_HUMAN	51 kDa	6
53	GRP75_HUMAN	74 kDa	5
54	RUVB1_HUMAN	50 kDa	5

54 proteins were identified with a peptide FDR of 1.75% and a protein FDR of 0%.

Supplemental Experimental Procedures

Cell culture, transfections, RNAi and luciferase assays.

Human HEK293T cells (referred to as HEK293) were grown in DMEM (GIBCO BRL) supplemented with 2 mM L-glutamine and 10% (v/v) FCS. In tethering experiments, cells were transfected with 25 ng RL-5BoxB (or 50 ng RL-5BoxB-HSL+HhR), 150 ng FL-Con and, if not indicated otherwise, 200 ng HA- or NHA-fusion constructs per well in a 12-well plate using nanofectin (PAA Laboratories) according to the manufacturer's instructions. Cells were lysed 24-48 h after transfection.

For DDX6 knockdown and rescue experiments with miRNA reporters, 300,000 cells per well of a 6-well plate were first transfected with a final concentration of 23 nM siRNA specific for DDX6 mRNA [5'-GCAGAAACCCUAUGAGAUUdTdT-3'; (Ozgun and Stoecklin, 2013; Qi et al., 2012)] or control siRNA [5'-GCAUUCACUUGGAUAGUAAdTdT-3'; (Sandler et al., 2011)], using HiPerFect Transfection Reagent (QIAGEN). A second transfection was performed 48 h later, using 80,000 cells per well of a 24-well plate and Attractene Transfection Reagent (QIAGEN). In the assays of let-7 reporters the second transfection mixtures contained 0.25 ng of pSF3 CMV-RL-hmga2-wt or -mut [these plasmids also encode FL; (Bethune et al., 2012)] or pRL-3xBulge-wt or -mut (Pillai et al., 2005) reporters, 100 ng of plasmid encoding pri-let-7a (Wu et al., 2006), 100 ng of plasmid encoding either HA-MBP (Maltose Binding Protein; used as control) or HA-DDX6 (wt or mutant forms; all siRNA-resistant), and either DDX6-specific or control siRNA at a final concentration of 50 nM. In experiments testing pRL-3xBulge-wt or -mut, transfections contained in addition 1.5 ng of FL-Con. In experiments with the miR-16/miR-196b reporter, second transfection mixtures contained, instead of let-7 reporters and a plasmid encoding pri-let7a, 0.25 ng of miR-16/miR-196b reporter psiCHECK T (Landthaler et al., 2008), either wt or mut form. Cells were lysed 48 h after the second transfection.

When indicated, transfection mixtures in both tethering and miRNA reporter experiments also contained 100-200 ng of plasmids encoding PAN2^{catal} and CAF1^{catal}.

Luciferase activities were measured with the Dual-Luciferase Reporter Assay System (Promega). RL activity was normalized to the activity of co-expressed FL and the

normalized RL values are shown relative to the indicated control either as percentage of the control (control at 100%) or as fold repression relative to the control. In experiments with miRNA reporters the normalized RL activity of the “wild-type” (wt) reporter is shown relative to the normalized activity of the mutant (mut) reporter tested in parallel. In tethering experiments (Figures 1B,C; 6A-C; and S6A), the normalized RL activity in cells expressing the indicated NHA-fusion proteins is shown relative to the normalized RL activity in cells expressing either HA-CNOT1 wt (Figure 1B,C) or HA-CNOT1-R wt (Figures 6A-C and S6A). In other tethering experiments (Figures 3D and S6C,D,G,H), the normalized RL activity in cells expressing the indicated NHA-fusion protein is shown relative to the normalized RL activity in cells expressing the corresponding HA-fusion protein. In all luciferase assays, values represent means \pm SEM from 3 to 7 independent biological replicate experiments.

DNA constructs

Reporter plasmids RL-5BoxB, FL-Con (Pillai et al., 2004), RL-5BoxB-HSL+HhR (Chekulaeva et al., 2011), pRL-3xBulge-wt, pRL-3xBulge-mut (Pillai et al., 2005), pSF3 CMV-RL-hmga2-wt and -mut (Bethune et al., 2012) have been described previously. Plasmids encoding NHA-CED, NHA-CED-7W, NHA-RRM, HA-CNOT1, NHA-CNOT1, HA-CAF1, NHA-CAF1, HA-CAF1 D40A (CAF1 D40A is referred to as CAF1_{catal} in the text), NHA-CAF1_{catal}, and plasmids expressing GST-tagged TNRC6C CED, TNRC6C M2-RRM, TNRC6C RRM-Cterm and their W-motif mutants have also been described (Chekulaeva et al., 2011; Zipprich et al., 2009). Plasmids encoding dominant negative catalytic mutants of PAN2 [PAN2_{catal}; pSR-HA-Pan2mut (Yamashita et al., 2005)] and CAF1 [CAF1_{catal}; pSR-HA-Caf1 mut (Zheng et al., 2008)] and the plasmid pSR-HA-CCR4 (Chang et al., 2004) were kind gifts of A.B. Shyu, University of Texas Medical School, Houston.

Plasmids encoding the HA or NHA-tagged sub-fragments of CNOT1 were generated by PCR-amplification of the corresponding fragments of pCI-NHA-CNOT1 and cloning into pCI-neo vector bearing an HA- or NHA-tag (Pillai et al., 2004). The CNOT1 numbering corresponds to isoform 1 of the protein (NCBI Reference Sequence: NP_057368.3). To obtain pCI-HA-CNOT9, pCI-NHA-CNOT9, and pCI-NHA-CNOT2, the

Pull-down assays and western blotting

GST pull-down assays were performed as described previously (Chekulaeva et al., 2011). Briefly, HEK293 cells were grown in 10-cm dishes and transfected with 3 to 6 µg of plasmid expressing the indicated GST-tagged proteins. When indicated the transfection mixes in addition contained 0.5 to 2 µg of plasmid expressing the investigated HA-tagged protein. Cells were lysed 24 h post-transfection in buffer A (50 mM Tris-HCl, pH 7.5, 150 mM KCl, 0.5% (v/v) Triton X-100, 1x complete EDTA-free protease inhibitor cocktail (Roche)). Lysates were cleared twice and treated with micrococcal nuclease (10 ng/µl) for 25 min at 25 °C. This nuclease treatment was shown to largely eliminate RNA-dependent interactions (Fabian et al. 2009 and data not shown). The lysates were then incubated with glutathione (GSH)-Sepharose beads (GE Healthcare) for 2 h at 4°C, beads were washed 3x with buffer B (50 mM Tris-HCl, pH 7.5, 150 mM KCl, 0.1% (v/v) Triton X-100, 1x complete protease inhibitor cocktail), and GST-tagged proteins were eluted with 50 mM GSH in buffer B.

For western blotting analysis, proteins were separated by SDS-PAGE using 10% linear polyacrylamide gels or NuPAGE Novex 4-12% Bis-Tris Gels (Life Technologies). Proteins were detected using ECL (GE Healthcare) or SuperSignal West Femto Chemiluminescent Substrate (Thermo Scientific). The following primary antibodies were used for western blotting: anti-CAF1 (CNOT7) (Abnova, clone 2F6), dilution 1:1,000; anti-CNOT1, 1:250 - 1:1,000 (kindly provided by M. Collart, University of Geneva); anti-CNOT1, (Proteintech 14276-1-AP), 1:1,000; anti-CNOT9 (RQCD1) (Proteintech 22503-1-AP), 1:1,000; anti-DDX3X (EMD Millipore), 1:3,000; anti-DDX6 (Bethyl Laboratories A300-461A), 1:10,000; anti-DHX29 (Bethyl laboratories A300-751A), 1:3,000; anti-EIF2C2 (AGO2) (Abnova, clone 2E12-1C9), 1:1,000; anti-eIF4AI (Abcam ab31217), 1:700 - 1:3,000; anti-eIF4AII (Abcam ab31218), 1:700 - 1:1,000; anti-eIF4AIII (Abcam ab32485), 1:3,000; anti-GST (GE Healthcare), 1:30,000 - 1:50,000; anti-GW182/TNRC6A (Bethyl A302-329A), 1:2,000; anti-HELZ (Abnova, clone 5B2), 1:1,000; and anti-HA tag (Roche, 3F10), 1:3,000. The following secondary antibodies were used: ECL Anti-Mouse IgG, HRP-linked whole Ab (from sheep) (GE Healthcare), dilution 1:10,000; ECL Anti-Rabbit IgG, HRP-linked whole Ab (from donkey) (GE Healthcare),

CNOT9 and CNOT2 coding regions were PCR-amplified from plasmids pSGCNOT9 (Zwartjes et al., 2004) and B42-hNOT2 (kind gifts of H. T. Marc Timmers, University of Utrecht), respectively, and cloned into pCI-HA or pCI-NHA vector (Pillai et al., 2004).

To obtain the plasmid encoding GST-TNRC6C SD (residues 1260-1690) and GST-TNRC6C CED Δ PAM2 (residues 1417-1690), the corresponding region of TNRC6C was PCR-amplified and cloned into the pEBG vector (Addgene). Similarly, to generate plasmids encoding GST-CNOT1-R (residues 727-1601), GST-CNOT1-MIF4G (residues 1054-1314), GST-CNOT1-CN9BD (1330-1601), and GST-CNOT1-C (1827-2376), the corresponding fragments of pCI-NHA-CNOT1 were PCR-amplified and cloned into the pEBG vector. To generate the plasmid encoding GST-eIF4G-MIF4G, the region encoding amino acids 674-1079 of eIF4GI (isoform 1) was amplified by PCR using the plasmid pcDNA3-HA-eIF4GI (1-1599) (Martineau et al., 2008) (a kind gift of Nahum Sonenberg, McGill University, Montreal) as template and cloned into the pEBG vector. To obtain the plasmid encoding GST-CAF1, the sequence encoding the ORF of CAF1 was PCR-amplified using the plasmid NHA-CAF1 as template and cloned into the pEBG vector.

To generate a plasmid encoding a siRNA-resistant version of HA-DDX6, the EcoR1-Blp1 fragment (encompassing the region of the DDX6 ORF targeted by the siRNA) of the plasmid Myc-RCK (Qi et al., 2012) (encoding a siRNA resistant version of DDX6) (a kind gift of Qing Jing, Shanghai Institutes for Biological Sciences) was cloned into the plasmid pcDNA3-HA-DDX6-R1r (Ozgun and Stoecklin, 2013) (a kind gift of Georg Stoecklin, DKFZ-ZMBH Alliance, Heidelberg). To obtain the plasmid encoding HA-MBP, the ORF of MBP was amplified by PCR using the plasmid pMAL-TRBP2 (Dorin et al., 2003) as template and cloned into pcDNA3-HA-DDX6, replacing the ORF of DDX6.

Plasmid psiCHECK T (a kind gift of Markus Landthaler, Max-Delbrück-Center for Molecular Medicine, Berlin) has been described previously (Landthaler et al., 2008). To obtain the control psiCHECK T mut plasmid, 5-6 point mutations, mostly targeting seed sequences, were introduced into each of the two miR-16 and two miR-196b target sites present in psiCHECK T.

Amino acid mutations in CNOT1, CAF1, CNOT9, and DDX6 were introduced by site-directed mutagenesis according to (Zheng et al., 2004).

1:10,000; ECL Anti-Rat IgG, HRP-linked whole antibody (from goat) (GE Healthcare), 1:10,000; donkey anti-goat IgG-HRP (Santa Cruz Biotechnology), 1:10,000.

Purification of the human CNOT1 - CNOT9 core complex and its *S. cerevisiae* ortholog

Human CNOT1-MID_s (residues 1054-1601), including the MIF4G domain and the CN9BD, was subcloned as an N-terminal His-GST-tag fusion protein cleavable with 3C protease. Human CNOT9 was cloned as an N-terminal His-SUMO (6xHistidine – Small Ubiquitin-like Modifier) tag fusion protein cleavable with the SUMO-specific protease SENP2. Proteins were expressed in *Escherichia coli* BL21(DE3) STAR pRARE (Stratagene) in Terrific Broth (TB) medium. Cells were resuspended in lysis buffer (50 mM NaH₂PO₄ pH 7.5, 250 mM NaCl, 25 mM imidazole) supplemented with 5 µg/ml DNase, 20 mM MgSO₄ and lysed by sonication. Proteins were purified using nickel-based affinity chromatography. After tag cleavage and reverse affinity chromatography, proteins were further purified by size-exclusion chromatography (Superdex 200, GE Healthcare in size-exclusion buffer (20 mM Tris pH 7.4, 250 mM NaCl, 1 mM DTT). The CNOT1-MID_s – CNOT9 complex was assembled by mixing the proteins in 1 : 1.2 molar ratio followed by size-exclusion chromatography (Superdex 200, GE Healthcare). The CNOT1-MID_s – CNOT9 complex was subjected to limited proteolysis with trypsin (0.2 mg protease for 10 mg of complex) at 4°C for 30 min. MS analysis showed that trypsin proteolyzed a portion of CNOT9 and cleaved CNOT1-MID_s into two fragments. The proteolyzed complex was further purified by Superdex 200 size-exclusion chromatography. One of the two fragments of CNOT1 corresponding to the MIF4G domain eluted separately in size-exclusion chromatography from a core binary complex that we characterized using MS analysis as encompassing CNOT1₁₃₅₁₋₁₅₈₈ (referred to as CNOT1 C9BD) and CNOT9₁₆₋₂₈₄ (CNOT9). A similar strategy was used to purify the corresponding *S. cerevisiae* complex of Not1₁₀₇₁₋₁₂₈₂ and Caf40₁₀₂₋₃₇₀ (the yeast orthologue of CNOT9). These human and yeast CNOT1 C9BD - CNOT9 core complexes were used for crystallization trials.

Crystallization and structure determination of the human and yeast CNOT1 C9BD - CNOT9 complexes

Crystallization was carried out at 18°C using the vapor diffusion method by mixing equal volumes of protein complex and of crystallization buffer. For the human CNOT1 C9BD – CNOT9 complex, a protein solution at 18.5 mg/ml was used and the best diffracting crystals were obtained in 50 mM MES pH 6.0, 200 mM ammonium sulfate after 2 days. Crystals were flash-frozen after soaking in a cryo-protection buffer containing 50 mM MES pH 6.0, 200 mM ammonium sulfate and 25 % glycerol and diffracted to 2.0 Å resolution. For the yeast complex, a protein solution at 20 mg/ml was used and the best diffracting crystals were obtained in 50 mM Hepes pH 7.0, 3 M NaCl. Crystals were flash-frozen after soaking in a cryo-protection buffer containing 50 mM Hepes pH 7.0, 3 M Sodium Chloride and 25 % glycerol and diffracted to 3.8 Å resolution. All diffraction data were collected at 100K at the beamline PXIII of the Swiss Light Source (SLS) synchrotron and processed using XDS (Kabsch, 2010) (Table 1).

The structure of the yeast complex was solved using a combination of single anomalous dispersion using a tantalum bromide soak and molecular replacement with the structure of the human orthologue of CNOT9 as search model (Garces et al., 2007) using the program PHASER (McCoy et al., 2007). Model building was carried out manually with the program Coot (Emsley et al., 2010) and refined with PHENIX (Adams et al., 2010). The structure was refined to 3.8 Å resolution to an R_{free} of 31.6% and R_{factor} of 29.2%. The atomic model of the yeast complex was used as search model to solve the structure of the orthologous human CNOT1 C9BD - CNOT9 complex by molecular replacement using the program PHASER (McCoy et al., 2007). The structure was refined to 2Å resolution to an R_{free} of 23.4% and R_{factor} of 19.6%. Crystal of the human CNOT1 C9BD - CNOT9 complex were soaked with free tryptophan at a concentration of 10 mM for 15 min. The structure of the ternary complex was refined to 2 Å resolution with an R_{free} of 22.2% and R_{factor} of 18.4% (Table 1).

Purification of proteins used for size-exclusion chromatography

The recombinant proteins or their fragments used for size-exclusion chromatography and MALS (multiangle static light scattering) measurements shown in Figure 2A,B and Figure

5 mM β -Mercaptoethanol) supplemented with 5 mg/ml DNase, 20 mM MgSO₄ and 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication. Proteins were purified using Nickel-based affinity chromatography and ion-exchange chromatography at pH 8.0 (Heparin, GE healthcare), followed by tag cleavage with SENP2 and size-exclusion chromatography (Superdex 200) in size-exclusion buffer at pH 8.0. The DDX6₉₅₋₄₆₉ construct was identified by limited proteolysis of wt DDX6 followed by N-terminal sequencing and MS analysis, and corresponds to a region spanning the two RecA domains. DDX6₉₅₋₄₆₉ crystallized in 17% (w/v) PEG 10000, 100 mM bis-Tris pH 5.5 and 100 mM ammonium acetate. DDX6₉₅₋₄₆₉ crystals were flash-frozen and diffracted to 3.0 Å resolution at the beamline PXII of the Swiss Light Source (SLS) synchrotron. Data were processed using XDS (Kabsch, 2010). The structure was solved by molecular replacement with the program PHASER (McCoy et al., 2007) using the DDX6 RecA2 domain structure (Tritschler et al., 2009). The model was built with Coot (Emsley et al., 2010) and refined with PHENIX (Adams et al., 2010) to an R_{free} of 26.2% and R_{factor} of 23.1% (Table 1).

Human CNOT1₁₀₆₃₋₁₃₁₄ (CNOT1 MIF4G) and its MIF4G_{hel2+3} mutant form were cloned as N-terminal His-GST-tag fusion proteins cleavable with 3C protease. CNOT1 MIF4G proteins were expressed and purified with similar protocols described above for DDX6. After tag cleavage and reverse affinity chromatography, CNOT1 MIF4G proteins were further purified by Superdex 75 size-exclusion chromatography. The CNOT1 MIF4G - DDX6₉₅₋₄₆₉ complex was reconstituted by mixing the proteins in 1.5 : 1 molar ratio followed by Superdex 200 size-exclusion chromatography in 20 mM bis-Tris pH 6.5, 150 mM NaCl, 1 mM DTT. Crystallization was carried out at 18°C using the vapor diffusion method by mixing equal volumes of protein complex at 5.4 mg/ml and of crystallization buffer (100 mM MES pH 6.5, 300 mM MgCl₂). Crystals were flash-frozen after soaking in cryo medium containing 100 mM MES pH 6.5, 300 mM MgCl₂, and 25 % glycerol, and diffracted to 2.3 Å resolution. The structure was solved by molecular replacement and refined to R_{free} of 22.0% and R_{factor} of 16.6% (Table 1). We found that the conformation of DDX6 bound to CNOT1 is similar to that observed in the structures of eIF4A-eIF4G (Schutz et al., 2008) and of Gle1-Dbp5 (Montpetit et al., 2011), which represent activated states of these ATPases. In all these cases, the relatively small interaction surfaces

between MIF4G and the RecA domains appear sufficient to drive the conformational change from the unbound form of the ATPase.

Ability of purified DDX6 proteins (wt and DDX6^{T327D,S330E} mutant forms) to form a complex with the purified CNOT1 MIF4G was also analyzed by size-exclusion chromatography as shown in Figure S4A.

ATPase assays

Assays were carried out in 20 μ l reaction volume containing 50 mM HEPES pH 7.0, 50 mM KCl, 5 mM Mg acetate, 0.2 mg/ml polyuridylic acid (Sigma), 2 mM DTT and 0.1 mg/ml BSA. Final protein concentrations were 3 μ M for DDX6 and 4.5 μ M for the CNOT1 MIF4G fragments. [α -³²P]ATP (sp. act. 3,000 Ci/mmol, Perkin-Elmer) was added to final ATP concentration of 1 mM (final sp. act. 50 μ Ci/ μ mol) to start the reaction. A 3- μ l aliquot was taken immediately for the initial time point. Reaction mixtures were then incubated at 37°C and further 3- μ l aliquots were collected after 5, 20, 60 and 180 min and quenched by adding 3 μ l of a buffer containing 50 mM EDTA and 0.1 % SDS. 0.5 μ l of the collected samples were spotted onto a Polyethyleneimine-Cellulose TLC plate (Merck), and eluted with 300 mM of KH₂PO₄/K₂HPO₄, pH 7.6. Plates were exposed overnight at room temperature to Fuji image plates and visualized using a Typhoon FLA 7000 phosphorimager (GE Healthcare).

Northern blotting

Northern blotting was performed as described previously (Chekulaeva et al., 2011). Briefly, 7 μ g of total RNA from transfected HEK293T cells isolated using Trizol Reagent (GE Healthcare) was resolved in a denaturing formaldehyde 1% (w/v) agarose gel and transferred to Hybond-N+ membrane (GE Healthcare) using 10X SSC. RL- or FL-specific probes internally labelled with [α -³²P]UTP were sequentially hybridized to the RNA on the membrane in ULTRAhyb Ultrasensitive Hybridization Buffer (Life Technologies) at 68 °C. Hybridization with FL-specific probe was performed ~2.5 months after that with RL probe to allow for decay of RL-specific signal. After washing the membrane with 0.2X SSC containing 0.1% (v/v) SDS at 68 °C, the signal was detected using a PhosphorImager screen and a GE TyphoonTM 9400 scanner. The bands were quantified with ImageQuant

TL software. We have verified, by preparing calibration curves, that Northern signals shown in Figure 6D,E were within the linear range of quantification.

Mass spectrometry analysis

For MS analysis, cysteine residues of proteins from GST pull-downs were reduced and alkylated prior to SDS-PAGE separation. The Coomassie-stained gel fragments were digested with trypsin. The peptides were analyzed by LC-MS/MS on an LTQ Orbitrap Velos mass spectrometer equipped with an Easy-nLC 1000 pump using an Acclaim PepMap 100 trap column (C18, 75 μm x 2 cm, 3 μm) and an Easy-Spray column (C18, 75 μm x 15 cm, 3 μm) at 40°C (all Thermo Fischer Scientific). Peptides were identified searching SwissProt database (version 2012_09) using Mascot (Matrix Science, version 2.3.02), considering carbamidomethylation of cysteine as fixed modification and deamidation of asparagine and glutamine, oxidation of methionine, acetylation of the protein N-terminus and phosphorylation of serine and threonine as variable modifications. Database search results were evaluated in Scaffold (Proteome Software, version 4.2.0). The peptide as well as the protein probability threshold was set to 80% and proteins with five or more identified peptides were accepted. Contaminants were manually excluded from the list of identified proteins. In total, 54 proteins (Table S1) were identified with a peptide FDR of 1.75% and a protein FDR of 0%.

Supplemental references

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* *66*, 213-221.
- Bethune, J., Artus-Revel, C.G., and Filipowicz, W. (2012). Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Rep* *13*, 716-723.
- Chang, T.C., Yamashita, A., Chen, C.Y., Yamashita, Y., Zhu, W., Durdan, S., Kahvejian, A., Sonenberg, N., and Shyu, A.B. (2004). UNR, a new partner of poly(A)-binding protein, plays a key role in translationally coupled mRNA turnover mediated by the c-fos major coding-region determinant. *Genes Dev* *18*, 2010-2023.
- Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat Struct Mol Biol* *18*, 1218-1226.

Cheng, Z., Collier, J., Parker, R., and Song, H. (2005). Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* *11*, 1258-1270.

Dorin, D., Bonnet, M.C., Bannwarth, S., Gatignol, A., Meurs, E.F., and Vaquero, C. (2003). The TAR RNA-binding protein, TRBP, stimulates the expression of TAR-containing RNAs in vitro and in vivo independently of its ability to inhibit the dsRNA-dependent kinase PKR. *The Journal of biological chemistry* *278*, 4440-4448.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* *66*, 486-501.

Garces, R.G., Gillon, W., and Pai, E.F. (2007). Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* *16*, 176-188.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* *66*, 133-144.

Landthaler, M., Gaidatzis, D., Rothbauer, A., Chen, P.Y., Soll, S.J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M., and Tuschl, T. (2008). Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* *14*, 2580-2596.

Martineau, Y., Derry, M.C., Wang, X., Yanagiya, A., Berlanga, J.J., Shyu, A.B., Imataka, H., Gehring, K., and Sonenberg, N. (2008). Poly(A)-binding protein-interacting protein 1 binds to eukaryotic translation initiation factor 3 to stimulate translation. *Mol Cell Biol* *28*, 6658-6667.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *Journal of applied crystallography* *40*, 658-674.

Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* *472*, 238-242.

Ozgun, S., and Stoecklin, G. (2013). Role of Rck-Pat1b binding in assembly of processing-bodies. *RNA Biol* *10*, 528-539.

Pillai, R.S., Artus, C.G., and Filipowicz, W. (2004). Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *Rna* *10*, 1518-1525.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of Translational Initiation by *let-7* MicroRNA in Human Cells. *Science* *309*, 1573-1576.

Qi, M.Y., Wang, Z.Z., Zhang, Z., Shao, Q., Zeng, A., Li, X.Q., Li, W.Q., Wang, C., Tian, F.J., Li, Q., *et al.* (2012). AU-rich-element-dependent translation repression requires the cooperation of tristetraprolin and RCK/P54. *Mol Cell Biol* *32*, 913-928.

Sandler, H., Kreth, J., Timmers, H.T., and Stoecklin, G. (2011). Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res* *39*, 4373-4386.

Schutz, P., Bumann, M., Oberholzer, A.E., Bieniossek, C., Trachsel, H., Altmann, M., and Baumann, U. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci U S A* *105*, 9564-9569.

Tritschler, F., Braun, J.E., Eulalio, A., Truffault, V., Izaurralde, E., and Weichenrieder, O. (2009). Structural basis for the mutually exclusive anchoring of P body components EDC3 and Tral to the DEAD box protein DDX6/Me31B. *Mol Cell* *33*, 661-668.

von Moeller, H., Basquin, C., and Conti, E. (2009). The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat Struct Mol Biol* *16*, 247-254.

- Wu, L., Fan, J., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. *Proc Natl Acad Sci U S A* *103*, 4034-4039.
- Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.Y., and Shyu, A.B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol* *12*, 1054-1063.
- Zheng, D., Ezzeddine, N., Chen, C.Y., Zhu, W., He, X., and Shyu, A.B. (2008). Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *The Journal of cell biology* *182*, 89-101.
- Zheng, L., Baumann, U., and Reymond, J.L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res* *32*, e115.
- Zipprich, J.T., Bhattacharyya, S., Mathys, H., and Filipowicz, W. (2009). Importance of the C-terminal domain of the human GW182 protein TNRC6C for translational repression. *RNA* *15*, 781-793.
- Zwartjes, C.G., Jayne, S., van den Berg, D.L., and Timmers, H.T. (2004). Repression of promoter activity by CNOT2, a subunit of the transcription regulatory Ccr4-not complex. *The Journal of biological chemistry* *279*, 10848-10854.

Architecture du module d'ubiquitination du complexe CCR4-NOT

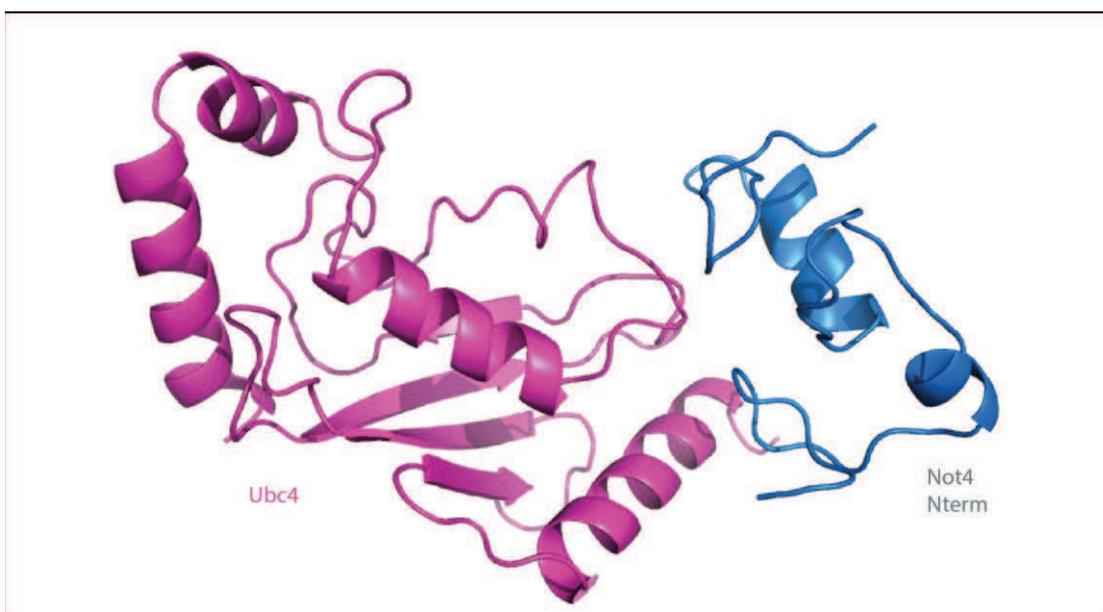
E.4. Architecture du module d'ubiquitination du complexe CCR4-NOT

E.4.1. Contexte et objectif de l'étude

Dans cette étude, nous avons étudié le module d'ubiquitination du complexe CCR4-NOT. L'objectif était de comprendre le mécanisme de recrutement des ligases E2 et E3 sur le complexe Ccr4-Not et de déterminer si cette interaction était compatible avec les autres modules du complexe.

E.4.2. Approches expérimentales

Le domaine N-terminal RING de NOT4 (30-83) chez *S. cerevisiae* peut interagir directement avec Ubc4. Afin d'obtenir des cristaux de ce complexe, nous avons décidé de suivre une procédure précédemment décrite pour un complexe E3-E2 ligase. Ces deux protéines ont été clonées dans un même vecteur d'expression *E. coli* et ont été fusionnées via un segment de 10 résidus. Il a été possible de surproduire la protéine chimérique et de la purifier à un degré de pureté compatible avec des essais de cristallisation. Un scan de fluorescence multi-longueurs d'onde réalisé au synchrotron a révélé la présence de zinc dans les cristaux. Nous avons donc collecté des données au pic d'absorption du zinc. La structure a été résolue en combinant une approche de remplacement moléculaire et de diffusion anormale à une résolution de 2,8 Å.



Structure cristalline du complexe binaire Ubc4-NOT4
Résolution: 2,8 Å
Remplacement moléculaire + diffusion anormale
pdb: 5AIE

Il a été montré que le domaine C-terminal de *S. cerevisiae* NOT4 interagissait avec le domaine C-terminal de NOT1 par des expériences de double hybride et de co-immunoprécipitation. Dans le but de produire un complexe NOT1-NOT4 et de s'assurer d'inclure les extrémités nécessaires à l'interaction nous avons aussi effectué une analyse de prédiction de structure secondaire afin de produire des fragments plus larges que ceux prédit par les études de co-immunoprécipitation. Nous avons réussi à produire un complexe de NOT1 (1348-2093) et NOT4 (418-587). L'approche de protéolyse ménagée (voir chapitre 2) a permis de générer un complexe plus compact comprenant les résidus 1541 à 2093 de NOT1, et les résidus 418 à 477 de NOT4. Ce complexe a ensuite été cloné dans des vecteurs d'expressions plasmidiques et il a été possible de surproduire ces deux nouvelles constructions. Le complexe a été formé par chromatographie d'exclusion de taille. Les cristaux obtenus à partir de ce complexe ont diffracté à 3,6 Å et la structure a été résolue par remplacement moléculaire en utilisant comme modèle la partie C-terminale de NOT1 précédemment résolue. Cette structure a mis en évidence une région étendue de NOT4 qui s'enroulait autour d'une des répétitions HEAT de NOT1 et prouve que NOT1 peut incorporer simultanément le module d'ubiquitination et le module NOT2 et NOT5 au sein du complexe CCR4-NOT.

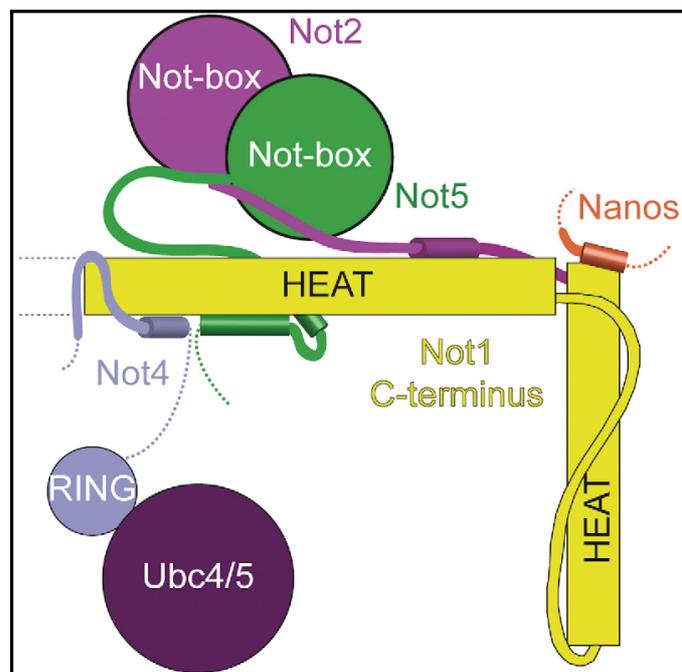
E.4.3. Résumé de l'étude

Dans cette étude, nous avons étudié le module d'ubiquitination du complexe CCR4-NOT et ainsi révélé, comment la ligase E3 contactait la ligase E2 et était recrutée sur le complexe CCR4-NOT. Nous avons déterminé la structure du domaine RING de NOT4 en complexe avec Ubc4. Cette structure nous a permis de déterminer les interactions des domaines E2-E3 de ce complexe. Nous avons également déterminé la structure du domaine C-terminal de NOT4 en complexe avec la protéine NOT1. Cette seconde structure a mis en évidence une région étendue de NOT4 qui s'enroulait autour d'une des répétitions HEAT de NOT1. Cette partie C-terminale n'est que partiellement conservée chez les métazoaires, ce qui peut expliquer son interaction faible avec NOT1. Cette étude structurale prouve en outre que NOT1 peut incorporer simultanément le module d'ubiquitination et le module NOT2 et NOT5 au sein du complexe CCR4-NOT.

Structure

Architecture of the Ubiquitylation Module of the Yeast Ccr4-Not Complex

Graphical Abstract



Authors

Varun Bhaskar, Jérôme Basquin, Elena Conti

Correspondence

conti@biochem.mpg.de

In Brief

Bhaskar et al. have determined the crystal structures of the complexes of the Not4 C-terminal domain with Not1 and the Not4 N-terminal RING domain with Ubc4. The structural and biochemical data highlight the underlying specificity of their interaction and rationalize the concomitant binding of Not4 and Not2-Not3/5 modules to Not1.

Highlights

- The Not1 C-terminal domain tethers the Not4 ubiquitylation module to yeast Ccr4-Not
- A low-complexity region of Not4 wraps around the C-terminal HEAT repeats of Not1
- In metazoans, Not4 lacks residues that confer high affinity binding to Not1 in yeast
- Not1_C can recruit Not4 and Not2-Not5 concomitantly to the Ccr4-Not complex

Accession Numbers

5AIE
5AJD

Architecture of the Ubiquitylation Module of the Yeast Ccr4-Not Complex

Varun Bhaskar,¹ Jérôme Basquin,¹ and Elena Conti^{1,*}

¹Department of Structural Cell Biology, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Munich, Germany

*Correspondence: conti@biochem.mpg.de

<http://dx.doi.org/10.1016/j.str.2015.03.011>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

SUMMARY

The Ccr4-Not complex regulates eukaryotic gene expression at multiple levels, including mRNA turnover, translational repression, and transcription. We have studied the ubiquitylation module of the yeast Ccr4-Not complex and addressed how E3 ligase binds cognate E2 and how it is tethered to the complex. The 2.8-Å resolution crystal structure of the N-terminal RING domain of Not4 in complex with Ubc4 shows the detailed interactions of this E3-E2 complex. The 3.6-Å resolution crystal structure of the C-terminal domain of the yeast Not4 in complex with the C-terminal domain of Not1 reveals how a largely extended region at the C-terminus of Not4 wraps around a HEAT-repeat region of Not1. This C-terminal region of Not4 is only partly conserved in metazoans, rationalizing its weaker Not1-binding properties. The structural and biochemical data show how Not1 can incorporate both the ubiquitylation module and the Not2-Not3/5 module concomitantly in the Ccr4-Not complex.

INTRODUCTION

The Ccr4-Not complex is a crucial player in the regulation of eukaryotic gene expression (reviewed in [Wahle and Winkler, 2013](#)). Ccr4-Not was originally discovered as a transcriptional regulator in yeast ([Collart and Struhl, 1994](#); [Draper et al., 1994](#)). Subsequent experiments revealed its fundamental function in cytoplasmic mRNA turnover, as a deadenylase that shortens the poly(A) tail at the 3' end of mRNAs ([Daugeron et al., 2001](#); [Tucker et al., 2002](#)). More recently, Ccr4-Not was also shown to act as a translational repressor (reviewed in [Chapat and Corbo, 2014](#)) and to be implicated in co-translational quality control ([Panassenko, 2014](#); [Matsuda et al., 2014](#)).

Purification of the Ccr4-Not core complex from endogenous sources has revealed the presence of a large macromolecular assembly containing several evolutionary conserved proteins and a few proteins that are instead species specific ([Chen et al., 2001](#); [Lau et al., 2009](#); [Temme et al., 2010](#); [Erben et al., 2014](#)). Ccr4-Not is assembled around Not1, a ~240 kDa protein that is built by consecutive helical domains. The individual domains of Not1 recruit the other core components of the complex

forming structurally and functionally distinct modules. The Not1 N-terminal domain is an elongated HEAT-repeat fold ([Basquin et al., 2012](#)) and appears to bind species-specific subunits (CNOT10-CNOT11 in metazoans, Caf130 in yeast) ([Chen et al., 2001](#); [Mauxion et al., 2013](#); [Bawankar et al., 2013](#)). The Not1 central MIF4G domain is next and recruits Caf1 (also known as Pop2 in yeast) and Ccr4, forming the deadenylase module of the complex ([Draper et al., 1994](#); [Bai et al., 1999](#)). This is followed by the Not1 helical bundle domain, which binds Caf40 ([Bawankar et al., 2013](#)). Last is the Not1 C-terminal domain, an elongated HEAT-repeat fold that binds Not2 and Not5 (and in yeast also the paralog Not3), forming the Not module of the complex ([Bai et al., 1999](#)). The C-terminal domain of Not1 also binds Not4, another core component of the yeast Ccr4-Not complex ([Bai et al., 1999](#)). Finally, several peripheral proteins are recruited to the core complex, such as DDX6, Nanos, tristetraprolin, and GW182 in metazoans ([Maillet and Collart, 2002](#); [Suzuki et al., 2010](#); [Sandler et al., 2011](#); [Braun et al., 2011](#); [Chekulaeva et al., 2011](#); [Fabian et al., 2011](#)).

In the past few years, most of the conserved interactions of the core complex as well as the interactions with several peripheral factors have been elucidated at the structural level ([Basquin et al., 2012](#); [Petit et al., 2012](#); [Fabian et al., 2013](#); [Boland et al., 2013](#); [Bhaskar et al., 2013](#); [Bhandari et al., 2014](#); [Chen et al., 2014](#); [Mathys et al., 2014](#)), with the exception of Not4. Not4 is an evolutionarily conserved protein that contains an N-terminal RING domain, a central RRM domain and a C-terminal domain predicted to be unstructured. As shown for both the yeast and human orthologs, the Not4 RING domain harbors an E3 ubiquitin ligase activity ([Albert et al., 2002](#); [Mulder et al., 2007a](#)). Consistently, Not4 has been reported to ubiquitylate a wide range of substrates ([Larabee et al., 2007](#); [Mulder et al., 2007b](#); [Mersman et al., 2009](#); [Cooper et al., 2012](#); [Gulshan et al., 2012](#)), including ribosome-associated factors ([Panassenko et al., 2006](#); [Panassenko and Collart, 2012](#)). Although the exact function is currently debated, the enzymatic activity of Not4 has been linked to proteasomal degradation in particular in the context of mRNA quality control pathways that respond to halted translation ([Dimitrova et al., 2009](#); [Matsuda et al., 2014](#)). The activity of the Not4 E3 ligase depends on its interaction with a specific E2, which has been identified as Ubc4/5 in yeast and the ortholog UbcH5B in humans ([Albert et al., 2002](#); [Mulder et al., 2007a](#)). Structural studies have shown how the RING domain of human CNOT4 folds via an unusual C4C4 motif whereby eight cysteine residues coordinate two zinc ions ([Hanzawa et al., 2001](#)). A model of the human CNOT4-UbcH5B complex has been proposed based on chemical shift nuclear magnetic resonance restraints,



computational docking approaches, and mutational analysis (Dominguez et al., 2004) but no crystal structure has been reported as yet.

Binding of yeast Not4 to the Ccr4-Not complex does not require the N-terminal RING domain but rather the C-terminal domain (Panasencko and Collart, 2011). The C-terminal domain of Not4, however, is the least conserved portion of the molecule. In addition, although Not4 is a bona fide Ccr4-Not subunit in yeast, it is not stably associated with the complex in human and *Drosophila* cells (Lau et al., 2009; Temme et al., 2010). The molecular basis for the Not1-Not4 interaction in yeast and the reason for the weaker association in higher eukaryotes are currently unknown. Also unknown is whether Not4 can bind Not1 in the context of the Not module, as Not2, Not3, and Not5 also dock to the same domain of Not1. Here, we report a structural and biochemical study that sheds light on how the E3 ligase of Not4 binds specifically its cognate E2 and how it is recruited to the Ccr4-Not complex.

RESULTS AND DISCUSSION

Overall Structure of *Saccharomyces cerevisiae* Not4_N Bound to Ubc4

The N-terminal RING domain of Not4 (Not4_N, residues 30–83 in *S. cerevisiae*) has been shown to interact with Ubc4 (Figure 1A) (Albert et al., 2002; Mulder et al., 2007a). To obtain crystals of the complex, we used a strategy that had been reported for another E3-E2 complex (Hodson et al., 2014) and connected the two proteins covalently via a 10-residue linker. The structure of the Not4_N-Ubc4 fusion protein was determined by a combination of zinc-based single-wavelength anomalous dispersion (SAD) and molecular replacement, and was refined to 2.8-Å resolution with R_{free} of 27.1%, R_{factor} of 22.0%, and good stereochemistry (Table 1). The final model of Not4_N-Ubc4 has well-defined electron density for most of the polypeptide, except for the connecting linker (Figure 1B).

The structure of yeast Not4_N bound to Ubc4 is very similar to that of the human CNOT4 ortholog in isolation (Hanzawa et al., 2001). The RING domain of Not4 contains two α -helices (the short $\alpha 1$ helix and the long $\alpha 2$ helix) and two zinc ions (Figure 1B). The zinc ions are coordinated in cross bracing fashion by cysteine residues that protrude from helix $\alpha 2$ and from the three loops regions L1, L2, and L3. The structure of yeast Ubc4 bound to Not4_N is very similar to a previously determined structure of Ubc4 in isolation (Cook et al., 1993). Briefly, Ubc4 is centered at a four-stranded antiparallel β -sheet flanked by an N-terminal α -helix ($\alpha 1$) and by three C-terminal α -helices ($\alpha 2$, $\alpha 3$, $\alpha 4$) (Figure 1B). When compared with the previously proposed model of the human CNOT4-UbcH5B complex (Dominguez et al., 2004), the experimentally determined structure of yeast Not4_N-Ubc4 shows localized differences (Figure S1A).

Specific Interaction Network between the Not4_N RING E3 and the Ubc4 E2

In the crystal structure, the Not4 helix $\alpha 2$ and the zinc-binding loops L1, L2, and L3 interact with two loops of Ubc4 that precede and follow the fourth strand of the β -sheet (L4 and L5) (Figure 1B). The central hotspot of the interaction is formed by Phe63 of Ubc4, which wedges into a hydrophobic pocket formed by

Leu35, Ile56, Cys60, Asn63, Leu70, and Pro75 of Not4 and by Pro62 and Pro96 of Ubc4 (Figure 1C). In addition, Ile37 of Not4 is involved in hydrophobic interactions with the aliphatic portion of the side chain of Lys5 and Lys9 in the helix $\alpha 1$ of Ubc4. This hydrophobic hotspot is surrounded by polar and electrostatic contacts: a hydrogen-bond interaction involving Not4 Arg78 and Ubc4 Gln93 and two salt bridge interactions between Not4 Glu38 and Ubc4 Lys5 and between Not4 Glu69 and Ubc4 Lys64 (Figure 1C). In addition, Ubc4 Lys64 is engaged in an intra-molecular salt bridge with Ubc4 Asp60. The Glu69-Lys64-Asp60 network effectively pulls the L3 loop of Not4 toward Ubc4, closing the hydrophobic core. The interaction interface is formed by evolutionary conserved residues (Figures 1D and 1E) and is consistent with the effects of mutations previously reported (Mulder et al., 2007a).

To understand the specificity of yeast Not4 toward Ubc4/5 enzymes, we structurally aligned the known yeast E2 proteins on Ubc4 and analyzed if residues at the Not4-binding interface are conserved (Figure S1B). The Ubc2 and Ubc9 E2 proteins lack a hydrophobic residue at the corresponding position of Ubc4 Phe63. Ubc3, Ubc7, Ubc10, Ubc12, and Ubc13 lack a positively charged residue at the corresponding position of Ubc4 Lys64. Ubc1, Ubc6, Ubc8, and Ubc11 lack the equivalent of Ubc4 Gln93. These subtle differences appear to weaken the interaction network observed in the Not4_N-Ubc4 structure, driving the specificity of Not4_N toward Ubc4/5 (Albert et al., 2002; Mulder et al., 2007a).

Overall Structure of Not4_C Bound to Not1_C

The C-terminal domain of *S. cerevisiae* Not4 (specifically residues 430–480) has been shown to interact with the C-terminal domain of Not1 by yeast two-hybrid and co-immunoprecipitation (co-IP) studies (Albert et al., 2002; Panasencko and Collart, 2011). To ensure the identification of the correct domain boundaries that would include all the determinants of the interaction, we used secondary structure predictions to engineer larger regions of the interacting proteins than those mapped from the co-IP experiment. We purified a complex encompassing Not1 residues 1348–2093 and Not4 residues 418–587 (Not4 $\Delta 417$). Limited proteolysis of this complex and subsequent gel filtration resulted in stable fragments that were characterized by N-terminal sequencing and mass spectrometry analysis as encompassing residues 418–477 of Not4 (Not4_C) and residues 1541–2093 of Not1 (the C-terminal domain of Not1, or Not1_C) (Figures 2A and S2). Consistent with the proteolysis results, GST-tagged Not4_C was able to precipitate Not1_C in pull-down assays (Figure 2B, lane 3). We purified the Not1_C-Not4_C minimal complex and obtained crystals diffracting to 3.6-Å resolution containing six copies of the complex in the asymmetric unit. We determined the structure by molecular replacement, using the previously determined structure of Not1_C as the search model (Bhaskar et al., 2013). The model was built and refined to R_{free} of 31.9%, R_{factor} of 26.6%, and good stereochemistry (Table 1). The six independent copies of the complex in the crystals are essentially identical and include residues 1568–2078 of Not1 (with the major exception of two loops between 1791–1800 and 2065–2071) and residues 420–469 of Not4 (Figure 2C).

The HEAT-repeat structure of Not1_C in the Not4_C-bound complex is similar to that in the Not2-Not5_C-bound complex (Bhaskar

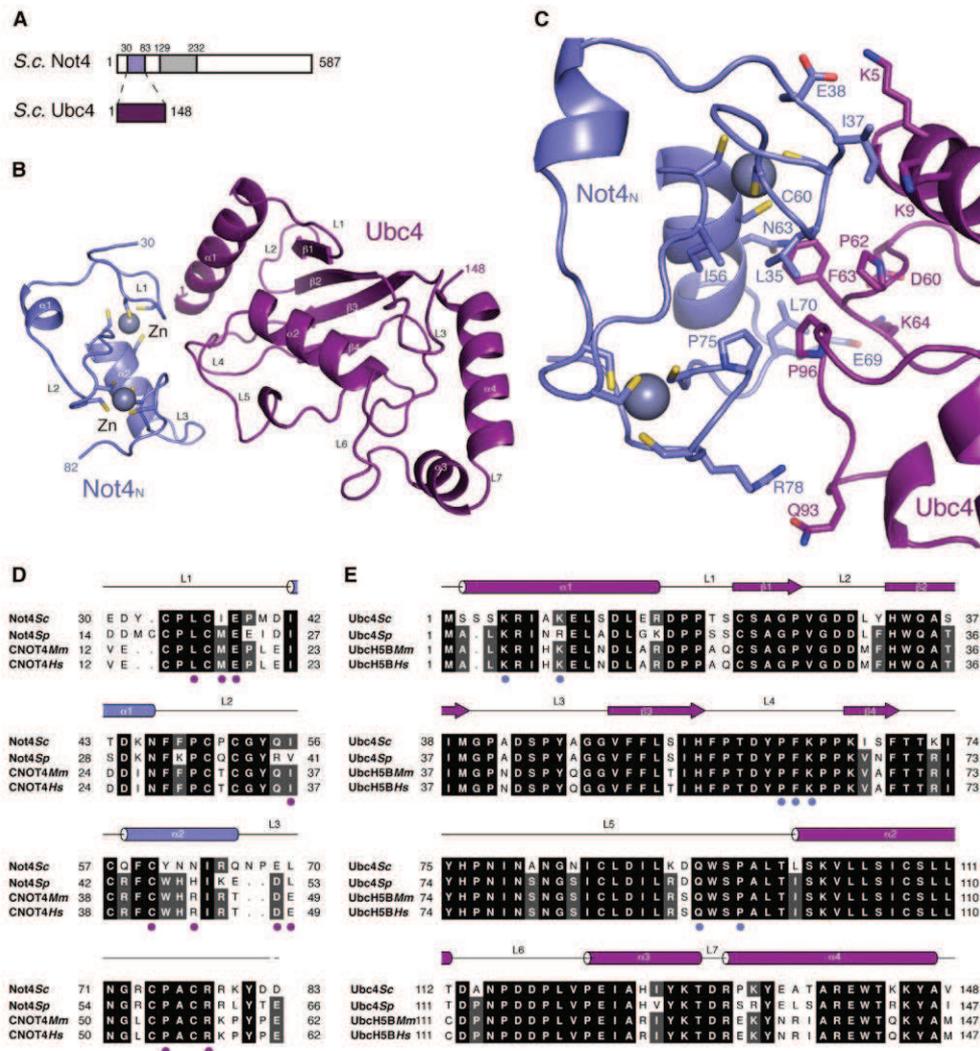


Figure 1. Structure of the Complex between the Not4 RING E3 and the Ubc4 E2

(A) A schematic diagram of the domain architecture of *S. cerevisiae* Not4 and Ubc4. The colored rectangles indicate the regions present in the crystal structure. The gray rectangle represents another folded domain, while the empty boxes represent low-complexity regions.

(B) Cartoon representation of the structure of yeast Not4_N (in blue) bound to Ubc4 (in purple). The N- and C-terminal residues of the two proteins ordered in the electron density are indicated. The secondary structure elements are labeled. The two zinc ions are shown as spheres and the cysteine residues that coordinate them are shown in stick representation. This structure figure and all others in the article were generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.2.3pre, Schrödinger, LLC).

(C) Close-up view of the interaction interface between Not4_N and Ubc4. Interacting residues are shown and labeled.

(D and E) Structure-based sequence alignment of Not4_N and Ubc4 from different species, including *S. cerevisiae* (Sc), *Mus musculus* (Mm) and *Homo sapiens* (Hs), highlighting the interacting residues. The secondary structure elements are shown above the sequence. See also Figure S1.

et al., 2013). HEAT repeats consist of two antiparallel α -helices (termed A and B) and pack side by side in a regular fashion. The ten HEAT repeats of Not1_C are organized in two units. The first unit is made of six HEAT repeats and is arranged in a perpendicular fashion with respect to the second unit, which is composed

of the four C-terminal repeats (Figure 2C). The loop connecting HEATs 7 and 8 is in an extended conformation, likely due to crystal contacts. Not4_C folds into an α -helix (residues 426–439) that is flanked by extended regions lacking defined secondary structure elements (residues 420–425 and 440–469) (Figure 2C).

Table 1. Data Collection and Refinement Statistics

	Not4 _N -Ubc4		
	Zinc SAD	Native	Not1 _C -Not4 _C
Wavelength (Å)	1.2819	1	1
Resolution range (Å) ^a	37.33–2.48	53.56–2.80 (2.90–2.80)	75.66–3.62 (3.75–3.62)
Space group	<i>P1</i>	<i>R3H</i>	<i>P3₂21</i>
Unit cell	62.45, 62.96, 65.43	107.11, 107.11, 62.20	173.66, 173.66, 262.61
α, β, γ (°)	108.50, 107.40, 108.09	90, 90, 120	90, 90, 120
Total reflections ^a	95,124 (13,582)	67,241 (6,289)	352,483 (34,067)
Unique reflections ^a	52,704 (7,634)	6,541 (628)	52,676 (5,095)
Multiplicity ^a	1.84 (1.77)	10.30 (10.00)	6.70 (6.70)
Completeness (%) ^a	92.5 (83.2)	99.4 (95.2)	99.7 (98.6)
Mean I/σ(I) ^a	14.66 (4.78)	29.35 (2.10)	11.26 (1.37)
SigAno ^a	1.31 (0.95)		
CC _{1/2} ^a	0.997 (0.949)	1 (0.911)	1 (0.733)
R _{merge} (%) ^a		5.6 (97.1)	11.2 (149.8)
R _{work} (%) ^a		22.0 (45.1)	26.6 (38.7)
R _{free} (%) ^a		27.1 (50.8)	31.9 (42.7)
Number of non-hydrogen atoms		1,575	24,555
Macromolecules		1,573	24,555
Ligands		2	0
Protein residues		204	3,223
RMS (bonds)		0.002	0.010
RMS (angles)		0.52	0.64
Ramachandran favored (%)		97	94
Ramachandran outliers (%)		0	0.064
Average B-factor		103.00	124.80
Macromolecules		103.10	124.80
Ligands		94.60	

SigAno = mean anomalous difference in units of its estimated standard deviation. (|F(+)_h - F(-)_h|/σ). F(+), F(-) are structure factor estimates obtained from the merged intensity observations in each parity class.

^aStatistics for the highest-resolution shell are shown in parentheses.

Extensive Interaction Network between Yeast Not1_C and Not4_C

Not4_C binds on the surface of the first three HEAT repeats of Not1_C, extending about 100 Å in length and burying a surface area of approximately 1500 Å² (Figure 2C). The contacts between Not4_C and Not1_C can be described as divided into three segments. In the first segment, the α-helix of Not4_C packs against the A helices of HEAT 2 and 3 of Not1_C. This interface is mainly dominated by hydrophobic interactions between Leu430, Leu434, and Leu437 of Not4_C and Leu1613, Val1671, and Val1675 of Not1_C (Figures 3A and S3A). In the second segment, residues 442–452 of Not4 interact extensively with

two loops of Not1_C connecting HEAT 1 to 2 and HEAT 2 to 3. The interactions are mediated by a salt bridge and few hydrophobic contacts (Figure 3B). In the third segment, residues 462–469 of Not4 are in extended conformation and pack between the A and B helices of the first HEAT repeat of Not1_C. This interface involves hydrophobic contacts between Leu463, Phe464, and Trp466 of Not4 and Val1575, Leu1582, Ile1592, Phe1596, Leu1600, and Val1605 of Not1 (Figures 3C and S3B).

To test the relevance of the interacting regions, we engineered deletion mutants of Not4_C and carried out pull-down assays. As the second segment of the Not1_C-Not4_C interface appeared the weakest from an analysis using the PISA server (Krissinel and Henrick, 2007), we constructed versions of Not4_C lacking either the first hydrophobic segment (Not4_C-ΔN) or the C-terminal hydrophobic segment (Not4_C-ΔC). GST-tagged Not4_C-ΔN precipitated Not1_C to a similar extent as GST-Not4_C (Figure 3D, lanes 2 and 3). In contrast, GST-tagged Not4_C-ΔC failed to interact with Not1_C in the pull-down assay (Figure 3D, lane 4). Next, we introduced specific mutations in the C-terminal segment of Not4_C and tested them for their ability to interact with Not1_C in GST pull-down assays. Mutations of Not4_C either at Leu463 and Phe464 (L463E F464E) or at Phe464 and Trp466 (F464E W466E) failed to precipitate Not1_C (Figure 3E, lanes 3 and 4). Altogether, these results suggest that the C-terminal segment of Not4_C makes the most significant contribution to the Not1-Not4 interaction while the first and second segments of Not4_C have a minor role.

Not4 Binding to Not1 Is Partially Conserved in Metazoa

To date, *S. cerevisiae* is the only species in which a stable association of Not4 within the Ccr4-Not core complex has been detected. This raises the question as to whether the interactions observed in the Not1_C-Not4_C crystal structure are likely to occur in other species, particularly as in metazoa the incorporation of Not4 in the endogenous Ccr4-Not core complex has been barely detectable (Lau et al., 2009; Temme et al., 2010). In the case of Not1, many Not4-binding residues are evolutionarily conserved in higher eukaryotes (Figure S3C). In Not4, the first hydrophobic segment of the Not1-binding region is conserved. Human CNOT4, for example, features Ile419, Leu423, and Gln426 at the equivalent positions of *S. cerevisiae* Leu430, Leu434, and Leu437, respectively (Figure 3F). However, the third Not1-binding segment of Not4 is not present in human CNOT4. Since the third segment is essential for stable binding of Not4 to Not1 in yeast (Figures 3D and 3E), such differences rationalize the weaker in vivo association in higher eukaryotes (Lau et al., 2009; Temme et al., 2010).

Not4 Binding to Not1 Is Independent of the Not Module

Next, we compared the structure of the ubiquitylation module with that of the Not module. We superposed the structure of yeast Not1_C-Not4_C with those of yeast Not1_C-Not2-Not5_C (Bhaskar et al., 2013) and human CNOT1_C-CNOT2_C-CNOT3_C (Boland et al., 2013). While Not4_C binds the side surface of the first HEAT-repeat unit of Not1_C, yeast Not2-Not5_C and human CNOT2_C-CNOT3_C bind the top and the bottom surfaces (Figure 4A). Although there is a small overlap between the

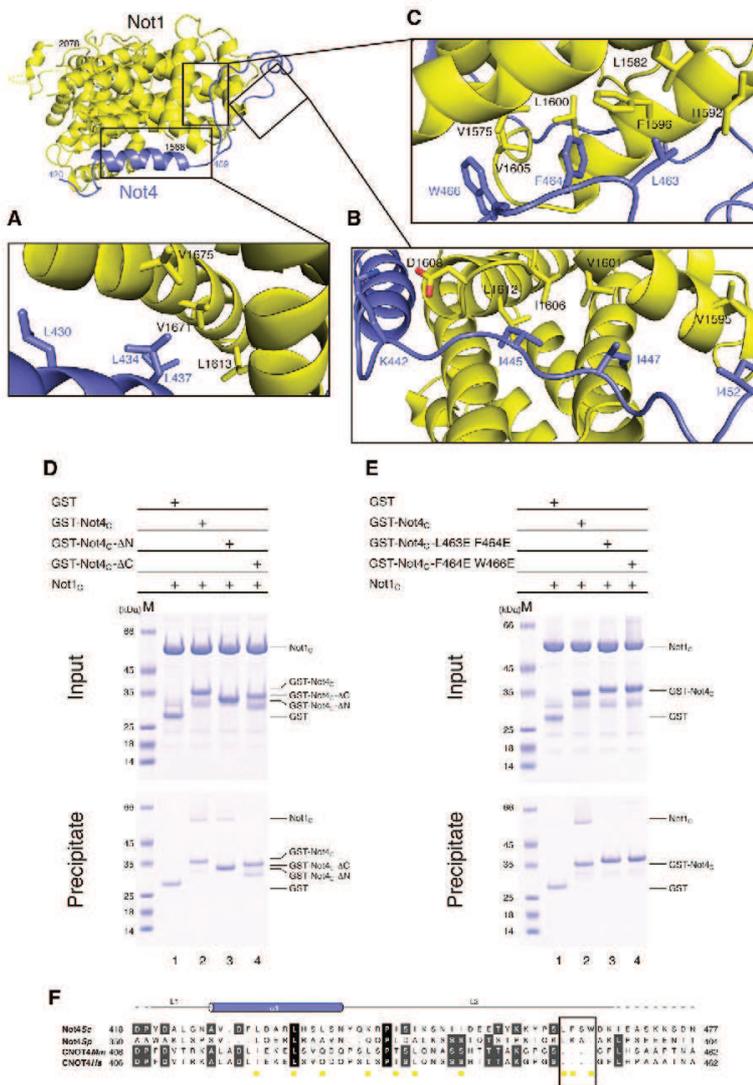


Figure 3. Not4_C Wraps around the N-terminal HEAT Repeats of Not1_C

(A–C) Close-up view of different segments of Not4_C that form the Not1_C interacting region. The position of each individual segment in the context of the complex is shown on the top left. The residues involved in interactions are shown as sticks and labeled.

(D and E) Pull-down experiments with GST-tagged versions of Not4 and untagged Not1_C, carried out as described in Figure 2B.

(F) Structure-based sequence alignment of Not4_C from different species, as mentioned in Figure 1D. The secondary structure elements are shown above the sequence. See also Figure S3.

model was refined against a 2.8-Å resolution native dataset (collected at 1 Å wavelength).

Not1_C-Not4_C was crystallized at 12 mg ml⁻¹ (see Supplemental Experimental Procedures). The crystals belong to space group *P*3₂21 with six copies in an asymmetric unit related by translational noncrystallographic symmetry (NCS). The structure was determined by molecular replacement using Not1_C from the Not1_C-Not2-Not5_C structure as the search model (Bhaskar et al., 2013). The model was refined for individual sites and individual B-factors along with torsion angle NCS restraints (in the initial rounds of refinement) that allow local conformational changes between the NCS-related copies.

Pull-down Assays

Pull-down assays of GST-tagged Not4 constructs with untagged Not1_C and/or Not2-Not5_C complex were performed as described in Bhaskar et al. (2013) (see Supplemental Experimental Procedures).

ACCESSION NUMBERS

The PDB accession numbers for the structure of Not4_N-Ubc4 and Not1_C-Not4_C reported in this paper are 5AIE and 5AJD, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedure and three figures and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2015.03.011>.

AUTHOR CONTRIBUTIONS

V.B. and J.B. performed the experiments; E.C. supervised the project; E.C. and V.B. wrote the manuscript.

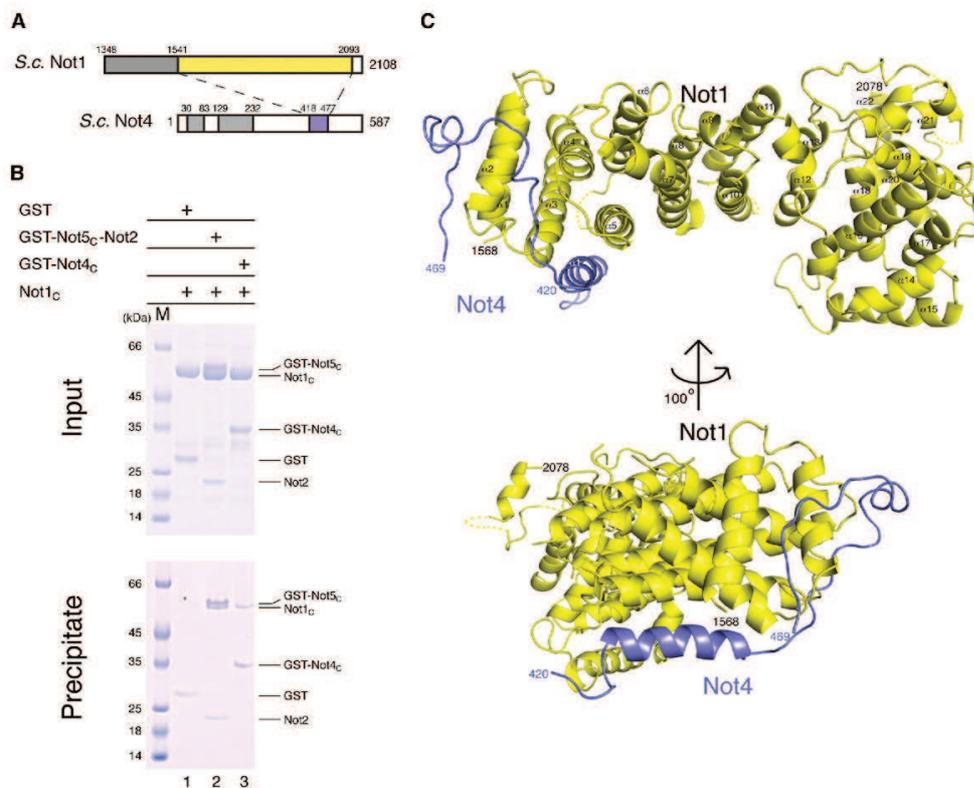


Figure 2. Structure of the Complex between Not4_C and Not1_C

(A) A schematic diagram of the domain architecture of *S. cerevisiae* Not1 C-terminal region and Not4. The colored rectangles indicate the regions present in the structure. The gray rectangles represent other folded domains, while the empty boxes represent low-complexity regions.
 (B) Protein co-precipitation by GST pull-down experiments. GST-Not4_C, GST-Not5_C-Not2 (positive control), or GST alone (negative control) were incubated with untagged Not1_C in a buffer containing 150 mM NaCl before co-precipitation with GSH-Sephadex beads, as indicated. Input (upper panel) and precipitates (lower panel) were analyzed on Coomassie stained 4%–12% Bis-Tris gradient gel (NuPage, Invitrogen). The proteins are labeled on the right.
 (C) Structure of the Not1_C-Not4_C complex shown in cartoon representation in two orientations. Not1_C is colored in yellow and Not4_C in blue. The N- and C-terminal residues of both proteins are marked. Disordered loops are indicated as dotted lines.
 See also [Figure S2](#).

N-terminal helix of Not4_C and the N-terminal region of Not5_C as observed in the yeast Not1_C-Not2-Not5_C complex, the structural analysis indicates that the interactions of Not4_C and Not2-Not5_C occur at largely separate surfaces of Not1_C. Indeed, pull-down assays showed that GST-tagged Not4_C could precipitate Not2-Not5_C in the presence of Not1_C ([Figure 4B](#)). Thus, the ubiquitylation module and the Not module can form simultaneously on the C-terminal domain of Not1. Finally, Not4_C binds at a completely different surface compared with the protein Nanos, which in metazoa is recognized by the C-terminal HEAT-repeat unit of CNOT1_C. Thus, the interactions of metazoan CNOT1 with CNOT2-CNOT3, CNOT4, and Nanos can in principle also occur simultaneously ([Figure 4C](#)). Whether and how bringing these proteins into close proximity by their concomitant interaction on the Not1_C platform affects the regulation or coordination of their functions are open questions for future studies.

EXPERIMENTAL PROCEDURES

Protein Purification

All proteins were cloned, expressed, and purified as previously described ([Bhaskar et al., 2013](#)) (see [Supplemental Experimental Procedures](#)).

Crystallization and Structure Determination

All crystals were obtained by vapor diffusion at room temperature. All data were collected at the PXII and PXIII beamlines of the Swiss Light Source, processed using XDS ([Kabsch, 2010](#)), and scaled and merged using Aimless ([Evans and Murshudov, 2013](#)). The structures were obtained after iterative rounds of model building using the program Coot ([Emsley et al., 2010](#)) and/or BUCCANEER ([Cowtan, 2006](#)) and refined using PHENIX.REFINE ([Adams et al., 2010](#)). The Not4_N-Ubc4 complex was crystallized at 48 mg ml⁻¹ (see [Supplemental Experimental Procedures](#)). Synchrotron data collected at the zinc edge (wavelength 1.28 Å) were used to solve the structure by molecular replacement-SAD in Phaser using the Ubc4 structure as a search model for molecular replacement and anomalous signal from the zinc atom ([Cook et al., 1993](#); [McCoy et al., 2007](#)). The final

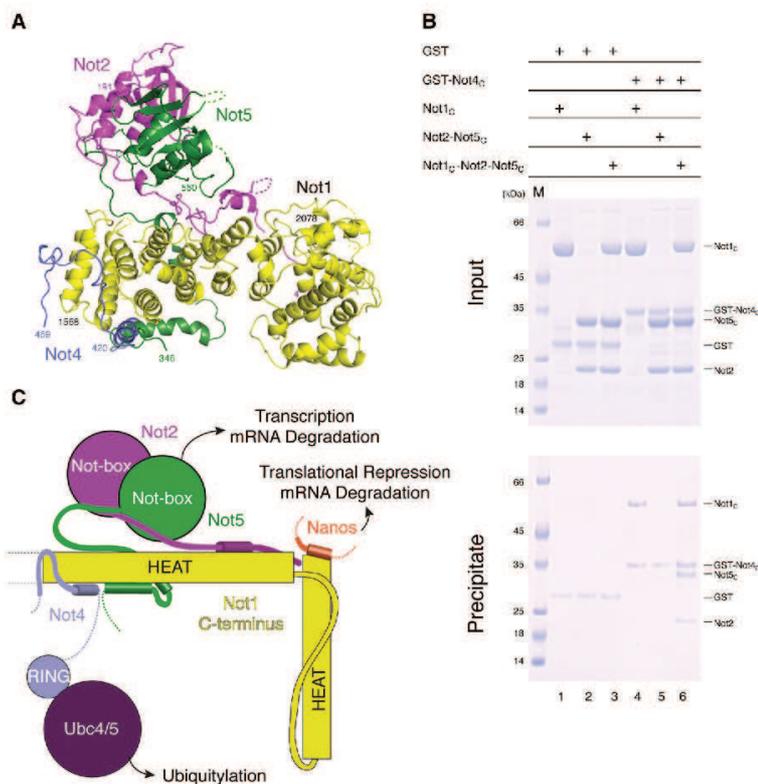


Figure 4. Not4_C Binds Not1_C Independently of Not2 and Not5

(A) Superposition of the yeast Not1_C-Not4_C and Not1_C-Not2-Not5_C structures. Not1_C is in yellow, Not2 in magenta, Not5_C in green, and Not4_C in blue.

(B) Pull-down experiments with GST-tagged Not4_C with untagged Not1_C and/or Not2-Not5_C, carried out as described in Figure 2B.

(C) Schematic diagram of the C-terminal domain of Not1 with the positions of the interacting proteins Not4, Not2-Not5 (or CNOT2-CNOT3 in humans), and Nanos.

ACKNOWLEDGMENTS

We would like to thank the Max-Planck-Institute of Biochemistry Core Facility and Crystallization Facility; the staff members at the beamlines PXII and PXIII of the Swiss Light Source, Airlie McCoy and Pavel Afonine for suggestions on NCS treatment; members of our lab for useful discussions and critical reading of the manuscript. This study was supported by the Max-Planck Gesellschaft, the European Commission (ERC Advanced Investigator Grant 294371, Marie Curie ITN RNPnet), and the Deutsche Forschungsgemeinschaft (DFG SFB646, SFB1035, GRK1721, FOR1680, CIPSM) to E.C.

Received: January 9, 2015

Revised: March 12, 2015

Accepted: March 13, 2015

Published: April 23, 2015

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221.

Albert, T.K., Hanzawa, H., Legtenberg, Y.J.A., de Ruwe, M.J., van den Heuvel, F.A.J., Collart, M.A., Boelens, R., and Timmers, H.T.M. (2002). Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO J.* **21**, 355–364.

Bai, Y., Salvatore, C., Chiang, Y.C., Collart, M.A., Liu, H.Y., and Denis, C.L. (1999). The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Mol. Cell. Biol.* **19**, 6642–6651.

Basquin, J., Roudko, V.V., Rode, M., Basquin, C., Séraphin, B., and Conti, E. (2012). Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Mol. Cell* **48**, 207–218.

Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S., and Izaurralde, E. (2013). NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol.* **10**, 228–244.

Bhandari, D., Raisch, T., Weichenrieder, O., Jonas, S., and Izaurralde, E. (2014). Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression. *Genes Dev.* **28**, 888–901.

Bhaskar, V., Roudko, V., Basquin, J., Sharma, K., Urlaub, H., Séraphin, B., and Conti, E. (2013). Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex. *Nat. Struct. Mol. Biol.* **20**, 1281–1288.

Boland, A., Chen, Y., Raisch, T., Jonas, S., Kuzuoğlu-Öztürk, D., Wohlbold, L., Weichenrieder, O., and Izaurralde, E. (2013). Structure and assembly of the NOT module of the human CCR4-NOT complex. *Nat. Struct. Mol. Biol.* **20**, 1289–1297.

Braun, J.E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44**, 120–133.

Chapat, C., and Corbo, L. (2014). Novel roles of the CCR4-NOT complex. *Wiley Interdiscip. Rev. RNA* **5**, 883–901.

Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226.

Chen, J., Rappsilber, J., Chiang, Y.C., Russell, P., Mann, M., and Denis, C.L. (2001). Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. *J. Mol. Biol.* **314**, 683–694.

- Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol. Cell* **54**, 737–750.
- Collart, M.A., and Struhl, K. (1994). NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. *Genes Dev.* **8**, 525–537.
- Cook, W.J., Jeffrey, L.C., Xu, Y., and Chau, V. (1993). Tertiary structures of class I ubiquitin-conjugating enzymes are highly conserved: crystal structure of yeast Ubc4. *Biochemistry* **32**, 13809–13817.
- Cooper, K.F., Scamati, M.S., Krasley, E., Mallory, M.J., Jin, C., Law, M.J., and Strich, R. (2012). Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction. *J. Cell Sci.* **125**, 1015–1026.
- Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011.
- Daugeron, M.C., Mauxion, F., and Séraphin, B. (2001). The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res.* **29**, 2448–2455.
- Dimitrova, L.N., Kuroha, K., Tatematsu, T., and Inada, T. (2009). Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. *J. Biol. Chem.* **284**, 10343–10352.
- Dominguez, C., Bonvin, A.M.J.J., Winkler, G.S., van Schaik, F.M.A., Timmers, H.T.M., and Boelens, R. (2004). Structural model of the UbcH5B/CNOT4 complex revealed by combining NMR, mutagenesis, and docking approaches. *Structure* **12**, 633–644.
- Draper, M.P., Liu, H.Y., Nelsbach, A.H., Mosley, S.P., and Denis, C.L. (1994). Ccr4 is a glucose-regulated transcription factor whose leucine-rich repeat binds several proteins important for placing Ccr4 in its proper promoter context. *Mol. Cell. Biol.* **14**, 4522–4531.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
- Erben, E., Chakraborty, C., and Clayton, C. (2014). The CAF1-NOT complex of trypanosomes. *Front. Genet.* **4**, 299.
- Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214.
- Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., et al. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217.
- Fabian, M.R., Frank, F., Rouya, C., Siddiqui, N., Lai, W.S., Karetnikov, A., Blackshear, P.J., Nagar, B., and Sonenberg, N. (2013). Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat. Struct. Mol. Biol.* **20**, 735–739.
- Gulshan, K., Thommandru, B., and Moye-Rowley, W.S. (2012). Proteolytic degradation of the Yap1 transcription factor is regulated by subcellular localization and the E3 ubiquitin ligase Not4. *J. Biol. Chem.* **287**, 26796–26805.
- Hanzawa, H., de Ruwe, M.J., Albert, T.K., van Der Vliet, P.C., Timmers, H.T., and Boelens, R. (2001). The structure of the C4C4 ring finger of human NOT4 reveals features distinct from those of C3HC4 RING fingers. *J. Biol. Chem.* **276**, 10185–10190.
- Hodson, C., Purkiss, A., Miles, J.A., and Walden, H. (2014). Structure of the human FANCL RING-Ube2T complex reveals determinants of cognate E3-E2 selection. *Structure* **22**, 337–344.
- Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797.
- Larabee, R.N., Shibata, Y., Mersman, D.P., Collins, S.R., Kemmeren, P., Roguev, A., Weissman, J.S., Briggs, S.D., Krogan, N.J., and Strahl, B.D. (2007). CCR4/NOT complex associates with the proteasome and regulates histone methylation. *Proc. Natl. Acad. Sci. USA* **104**, 5836–5841.
- Lau, N.-C., Kolkman, A., van Schaik, F.M.A., Mulder, K.W., Pijnappel, W.W.M.P., Heck, A.J.R., and Timmers, H.T.M. (2009). Human Ccr4-not complexes contain variable deadenylase subunits. *Biochem. J.* **422**, 443–453.
- Maillet, L., and Collart, M.A. (2002). Interaction between Not1p, a component of the Ccr4-not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. *J. Biol. Chem.* **277**, 2835–2842.
- Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. *Mol. Cell* **54**, 751–765.
- Matsuda, R., Ikeuchi, K., Nomura, S., and Inada, T. (2014). Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes Cells* **19**, 1–12.
- Mauxion, F., Prêve, B., and Séraphin, B. (2013). C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol.* **10**, 267–276.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
- Mersman, D.P., Du, H.-N., Fingerman, I.M., South, P.F., and Briggs, S.D. (2009). Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. *Genes Dev.* **23**, 951–962.
- Mulder, K.W., Inagaki, A., Camerini, E., Mousson, F., Winkler, G.S., De Virgilio, C., Collart, M.A., and Timmers, H.T.M. (2007a). Modulation of Ubc4p/Ubc5p-mediated stress responses by the RING-finger-dependent ubiquitin-protein ligase Not4p in *Saccharomyces cerevisiae*. *Genetics* **176**, 181–192.
- Mulder, K.W., Brenkman, A.B., Inagaki, A., van den Broek, N.J.F., and Timmers, H.T.M. (2007b). Regulation of histone H3K4 tri-methylation and PAF complex recruitment by the Ccr4-Not complex. *Nucleic Acids Res.* **35**, 2428–2439.
- Panasenko, O.O. (2014). The role of the E3 ligase Not4 in cotranslational quality control. *Front. Genet.* **5**, 141.
- Panasenko, O.O., and Collart, M.A. (2011). Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29. *Mol. Cell Biol.* **31**, 1610–1623.
- Panasenko, O.O., and Collart, M.A. (2012). Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase. *Mol. Microbiol.* **83**, 640–653.
- Panasenko, O., Landrieux, E., Feuermann, M., Finka, A., Paquet, N., and Collart, M.A. (2006). The yeast Ccr4-Not complex controls ubiquitination of the nascent-associated polypeptide (NAC-EGD) complex. *J. Biol. Chem.* **281**, 31389–31398.
- Petit, A.P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2012). The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res.* **40**, 11058–11072.
- Sandler, H., Kreth, J., Timmers, H.T.M., and Stoecklin, G. (2011). Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res.* **39**, 4373–4386.
- Suzuki, A., Igarashi, K., Aisaki, K.-I., Kanno, J., and Saga, Y. (2010). NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs. *Proc. Natl. Acad. Sci. USA* **107**, 3594–3599.
- Temme, C., Zhang, L., Kremmer, E., Ihling, C., Chartier, A., Sinz, A., Simonelig, M., and Wahle, E. (2010). Subunits of the drosophila CCR4-NOT complex and their roles in mRNA deadenylation. *RNA* **16**, 1356–1370.
- Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. *EMBO J.* **21**, 1427–1436.
- Wahle, E., and Winkler, S. (2013). RNA decay machines: deadenylation by the Ccr4-Not and Pan2–Pan3 complexes. *Biochim. Biophys. Acta* **1829**, 561–570.

Structure, Volume 23

Supplemental Information

**Architecture of the Ubiquitylation Module
of the Yeast Ccr4-Not Complex**

Varun Bhaskar, Jérôme Basquin, and Elena Conti

SUPPLEMENTAL FIGURES AND LEGENDS

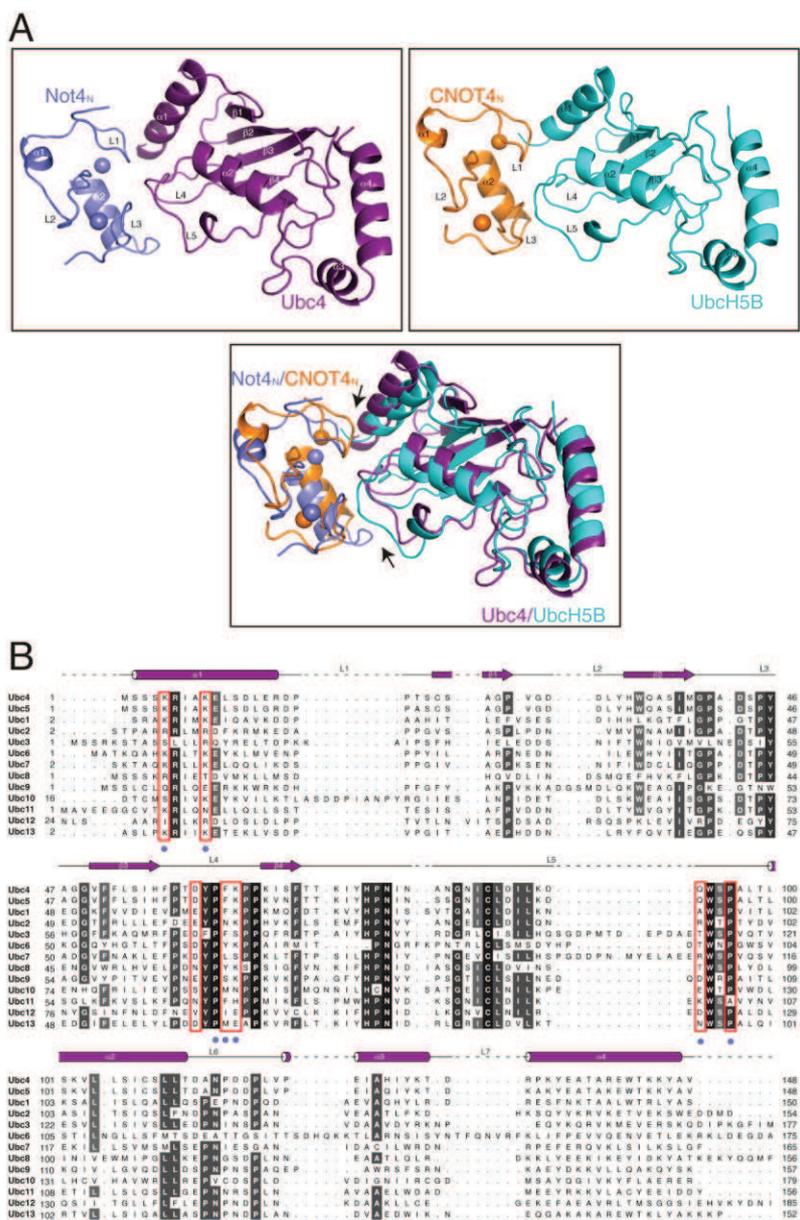


Figure S1

Detailed analysis of the Not4_N-Ubc4 crystal structure (Related to Figure 1)

(A) Not4_N-Ubc4 crystal structure and CNOT4-UbcH5B model are shown in similar orientation in the top panels. Superposition of the same is shown in the bottom panel. The difference in

the orientation of helix α_1 of E2 and the loop regions of E3 at the interface are highlighted by arrows.

(B) Structure-based sequence alignment of all the E2 enzymes in *S. cerevisiae*. Not4 interacting residues of Ubc4 are indicated with blue dots. Residues providing specificity for this E2-E3 interaction are highlighted. The secondary structure elements are shown above the sequence.

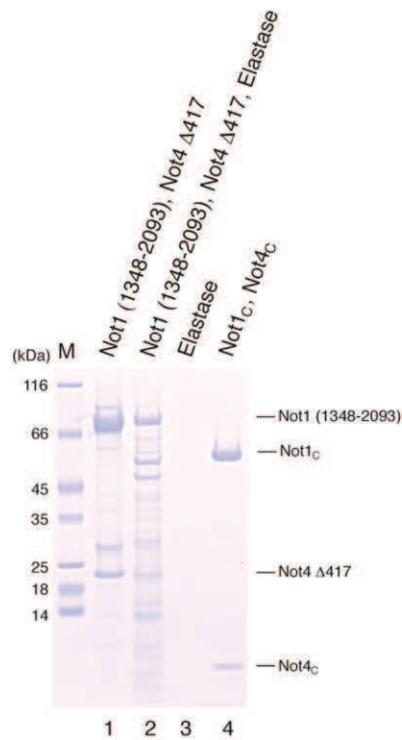


Figure S2

Identification of the Not1_c-Not4_c minimal complex (Related to Figure 2)

Not1 (1348-2093)-Not4 Δ 417 complex is shown in lane1. Limited proteolysis of Not1 (1348-2093)-Not4 Δ 417 was carried out by incubating the complex at 0.6 mg ml⁻¹ with elastase (Roche) for 60 minutes on ice at an enzyme to protein ratio of 1:10 and is shown in lane2. The mixture was then subjected to size-exclusion chromatography in a buffer containing 20 mM Tris-Cl pH 7.5, 250 mM NaCl and 2 mM DTT. The peaks were analyzed on 4-12% Bis-Tris NuPage gel with MES-SDS as the running buffer. The interacting fragments were identified by N-terminal sequencing and Liquid chromatography-Mass spectrometry (LC-MS) analysis. Not1_c-Not4_c complex that was used for structural studies is shown in lane4.

SUPPLEMENTAL EXPERIMENTAL PROCEDURE

Protein purification

All proteins were cloned and expressed in *E. coli* BL21 pLysS cells (Stratagene) in TB medium with 0.5 mM IPTG induction overnight at 18 °C. Not1 constructs were expressed as previously described in (Bhaskar et al., 2013). Not4_N and full-length Ubc4 were expressed as a fusion protein (connected by the linker TGSTGSTETG) with a N-terminal His-SUMO tag cleavable by Senp2 protease. The Not4_c, Not4_c-ΔN and Not4_c-ΔC (Not4 residues 418-477, 442-477 and 418-462, respectively) constructs were expressed as N-terminal His-GST fusion proteins followed by a 3C cleavage site. The proteins were purified using similar protocols as previously described (Bhaskar et al., 2013). Briefly, a first step of Nickel-based affinity chromatography was followed by tag cleavage and size-exclusion chromatography. For pull-down experiments, the GST-tagged proteins were purified with the same protocol but omitting the tag cleavage step.

Crystallization

The Not4_N-Ubc4 complex was crystallized at 48 mg ml⁻¹ by vapour diffusion using 10% (w/v) PEG 8000, 0.02 M L-Na-Glutamate, 0.02 M Alanine (racemic), 0.02 M Glycine, 0.02 M Lysine HCl (racemic), 0.02 M Serine (racemic), 0.1 M Bicine/Tris-Cl pH 8.5 and 20% (w/v) ethylene glycol as crystallization buffer at room temperature.

Not1_c-Not4_c complex was crystallized at 12 mg ml⁻¹ by vapour diffusion using 10% (w/v) PEG 4000, 0.02 M 1,6-Hexanediol, 0.02 M 1-Butanol, 0.02 M 1,2-Propanediol (racemic), 0.02 M 2-Propanol, 0.02 M 1,4-Butanediol, 0.02 M 1,3-Propanediol, 0.1 M MOPS/Hepes-Na 7.5 and 20% Glycerol as crystallization buffer at room temperature.

Pull-down assays

100 pmol of GST-tagged protein was incubated with 200 pmol of the untagged prey protein for 1 hr at 4 °C in the binding buffer (BB150 – 20 mM Tris-Cl pH 7.5, 150 mM NaCl, 2 mM DTT, 12.5% (v/v) glycerol and 0.1% (w/v) NP40). 400 μL of BB150 buffer and 20 μL of 50% GSH-Sepharose resin were added to the protein mix and incubated for 1 hr with gentle rocking at 4 °C. The resin was washed 3 times with BB150 and the proteins were eluted with 15 μL of BB150 containing 20 mM Glutathione. Input and precipitate were mixed with 3X SDS loading dye and resolved on 4-12% Bis-Tris NuPage gel (Invitrogen) using MES-SDS as running buffer, and visualized by Coomassie blue staining.

F. Conclusions et perspectives

Depuis quelques années, de nombreux travaux ont permis d'établir que le complexe CCR4-NOT n'était pas uniquement impliqué dans la dégradation des ARN, mais que bien au contraire il était un important acteur du contrôle de l'intégrité des programmes de régulation des gènes. Ces nouvelles connaissances ont permis de mieux comprendre la biologie du complexe CCR4-NOT. Toutes ces nouvelles données impliquant le complexe, aussi bien chez la levure que chez les eucaryotes supérieurs prouvent que le complexe CCR4-NOT est impliqué dans le contrôle des gènes à différents niveaux. Ce complexe intervient dans des processus cellulaires variés, tels que la réparation de l'ADN (Gaillard et al. 2009), le contrôle de la longueur de microtubules (DeBella et al. 2006) ou la méthylation des histones (Lau et al. 2009). Il interagit directement avec les facteurs de transcription et la polymérase II dans le noyau, mais aussi avec les ribosomes dans le cytoplasme. Sa présence dans différents compartiments cellulaires fait penser qu'il jouerait un rôle dans le contrôle de la qualité cellulaire. A ce titre, il peut contrôler la structure de la chromatine, et dans le cas d'un événement de transcription avorté, il agit sur l'élongation de la transcription en recrutant l'ubiquitine ligase NOT4 sur la chromatine (Reese 2013). Le complexe joue aussi un rôle sur les ribosomes dans le cas d'une traduction erratique et est capable de corriger les erreurs en jouant sur la répression de la traduction et la dégradation des ARN messagers suivis d'une terminaison anormale de la traduction en utilisant respectivement deux activités enzymatiques distinctes, l'ubiquitine ligase et la déadénylase (Reese 2013).

Des expériences chez des souris génétiquement modifiées ont mis en évidence les fonctions intégrées de ce complexe au sein de différents organes et tissus, intervenant par exemple dans la fonction cardiaque (Morita et al. 2007), la spermatogenèse (Nakamura et al. 2004) et le contrôle du renouvellement des cellules souches (Zheng et al. 2012). Il est vraisemblable que, puisque le complexe est impliqué dans ces grandes fonctions régulatrices, il soit un acteur principal de la régulation et de la coordination des programmes d'expression des gènes.

Les organismes vivants sont en permanence exposés à des agressions diverses, comme celles de pathogènes, des oncogènes ou encore des réponses inflammatoires. Dans ces conditions, il est crucial que la machinerie cellulaire soit capable de réagir très rapidement en réponse à ces signaux afin de garantir une bonne maintenance cellulaire pour assurer la viabilité des organismes. Chacun de ces signaux déclenche des événements en cascade impliquant de nombreuses interactions dans les différents chemins de transmission du signal. Une étude a montré que le complexe CCR4-NOT est impliqué dans la régulation du signal TNF et suggère qu'il peut agir comme plateforme chaperonne régulatrice (Collart and Timmers 2004). De plus le complexe CCR4-NOT peut adapter l'expression des gènes en réponse aux stimuli et aux changements cellulaires. La réponse du

système immunitaire innée est enclenchée par une phase d'induction pour répondre rapidement à une agression extérieure, suivie d'une répression de cette réponse. Comme le complexe CCR4-NOT agit sur les niveaux contrôlant l'expression des gènes, de la transcription à la dégradation des ARN, il permet une rapide adaptation des cellules aux agressions extérieures. Il serait très intéressant de comprendre les mécanismes moléculaires qui coordonnent le complexe CCR4-NOT, à la fois sur la synthèse des ARN messagers et sur leur dégradation.

Des progrès significatifs ont été accomplis ces quatre dernières années afin d'élucider la structure tridimensionnelle du complexe CCR4-NOT. Les interactions des différentes sous-unités ont pu être mises en évidence et les structures tridimensionnelles sont disponibles pour le module de déadénylation et le module NOT. Plus récemment, il a été possible de déterminer la structure d'autres facteurs interagissant directement avec le complexe CCR4-NOT (DDX6, TTP). En combinant toutes les structures déterminées par cristallographie aux rayons X et la structure du complexe déterminée par microscopie électronique en coloration négative, il est possible de proposer un modèle présentant le complexe dans sa totalité. Ce modèle devra être confirmé par une structure à plus haute résolution, probablement par cryo-microscopie électronique. De plus, d'autres études devront être menées afin de mieux comprendre les rôles fonctionnels du complexe, en particulier des modules déadénylase et NOT et leurs relations éventuelles. Pourquoi le complexe contient deux enzymes de déadénylation (CAF1 et CCR4) est une autre question intrigante qu'il reste à élucider. En effet, il n'a pas encore été déterminé si ces enzymes coopéraient ou avaient une fonction redondante (Winkler et al. 2013). Ces deux enzymes ont des préférences pour les substrats poly(A) mais il est à noter que chez la levure *S.cerevisiae*, le site actif est suboptimale et n'est probablement pas actif *in vivo*, servant uniquement à recruter CCR4 sur le complexe (Basquin et al. 2012). De plus chez le complexe humain, il n'est pas clair pourquoi il existe deux paralogues pour CCR4 (CNOT6, CNOT6L) et pour CAF1 (CNOT7, CNOT8).

Enfin, il n'a pas encore été identifié comment le complexe CCR4-NOT cible des ARN messagers spécifiques, ce qui doit impliquer des facteurs extérieurs pouvant se fixer sur le complexe, tels que la protéine PABP (poly(A)-binding protein), des facteurs de traduction ou des protéines pouvant se lier à l'ARN (Collart and Panasenko 2012). D'autres études structurales pourront se montrer très utiles afin de mieux définir ces interactions et à terme comprendre leurs rôles dans les grandes fonctions biologiques.

G. Références

- Albert TK, Hanzawa H, Legtenberg YIA, de Ruwe MJ, van den Heuvel FAJ, Collart MA, Boelens R, Timmers HTM. 2002. Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO J* **21**: 355–364.
- Albert TK, Lemaire M, van Berkum NL, Gentz R, Collart MA, Timmers HT. 2000. Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. *Nucleic Acids Res* **28**: 809–817.
- Andersen KR, Andersen KR, Jonstrup AT, Jonstrup AT, Van LB, Van LB, Brodersen DE, Brodersen DE. 2009. The activity and selectivity of fission yeast Pop2p are affected by a high affinity for Zn²⁺ and Mn²⁺ in the active site. *RNA* **15**: 850–861.
- Andrade MA, Petosa C, O'Donoghue SI, Müller CW, Bork P. 2001. Comparison of ARM and HEAT protein repeats. *J Mol Biol* **309**: 1–18.
- Arraiano CM, Mauxion F, Viegas SC, Matos RG, Séraphin B. 2013. Intracellular ribonucleases involved in transcript processing and decay: precision tools for RNA. *Biochim Biophys Acta* **1829**: 491–513.
- Azzouz N, Panasencko OO, Deluen C, Hsieh J, Theiler G, Collart MA. 2009. Specific roles for the Ccr4-Not complex subunits in expression of the genome. *RNA* **15**: 377–383.
- Badarinarayana V, Badarinarayana V, Chiang YC, Chiang YC, Denis CL, Denis CL. 2000. Functional interaction of CCR4-NOT proteins with TATAA-binding protein (TBP) and its associated factors in yeast. *Genetics* **155**: 1045–1054.
- Bardwell VJ, Kalkhoven E, Timmers H. 2006. Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription. *The EMBO*
- Bartlam M, Yamamoto T. 2010. The structural basis for deadenylation by the CCR4-NOT complex. *Protein Cell* **1**: 443–452.
- Basquin J, Roudko VV, Rode M, Basquin C, Séraphin B, Conti E. 2012. Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Mol Cell* **48**: 207–218.
- Bawankar P, Bawankar P, Loh B, Loh B, Wohlbold L, Wohlbold L, Schmidt S, Schmidt S, Izaurralde E, Izaurralde E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biol* **10**: 228–244.
- Bella J, Hindle KL, McEwan PA, Lovell SC. 2008. The leucine-rich repeat structure. *Cell Mol Life Sci* **65**: 2307–2333.
- Benson JD, Benson M, Howley PM, Struhl K. 1998. Association of distinct yeast Not2 functional domains with components of Gcn5 histone acetylase and Ccr4 transcriptional regulatory complexes. *EMBO J* **17**: 6714–6722.
- Berg JM, Stryer L, Tymoczko JL. 2015. *Stryer Biochimie*. Springer-Verlag.
- Bhaskar V, Roudko V, Basquin J, Sharma K, Urlaub H, Séraphin B, Conti E. 2013. Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex. *Nat Struct Mol Biol* **20**: 1281–1288.
- Bianchin C, Mauxion F, Sentis S, Séraphin B, Corbo L. 2005. Conservation of the deadenylase activity of proteins of the Caf1 family in human. *RNA* **11**: 487–494.
- Braun JE, Huntzinger E, Fauser M, Izaurralde E. 2011. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* **44**: 120–133.
- Bromberg J, Darnell JE. 2000. The role of STATs in transcriptional control and their impact on cellular function.

Oncogene **19**: 2468–2473.

- Buchwald G, Schüssler S, Basquin C, Le Hir H, Conti E. 2013. Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. *Proc Natl Acad Sci USA* **110**: E4611–8.
- Chapat C, Corbo L. 2014. Novel roles of the CCR4-NOT complex. *Wiley Interdiscip Rev RNA* **5**: 883–901.
- Chapat C, Kolytcheff C, Le Romancer M, Auboeuf D, La Grange De P, Chettab K, Sentis S, Corbo L. 2013. hCAF1/CNOT7 regulates interferon signalling by targeting STAT1. *EMBO J* **32**: 688–700.
- Chekulaeva M, Chekulaeva M, Mathys H, Mathys H, Zipprich JT, Zipprich JT, Attig J, Attig J, Colic M, Colic M, et al. 2011. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat Struct Mol Biol* **18**: 1218–1226.
- Chen J, Chiang Y-C, Denis CL. 2002. CCR4, a 3'–5' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. *EMBO J* **21**: 1414–1426.
- Cheng Z, Collier J, Parker R, Song H. 2005. Crystal structure and functional analysis of DEAD-box protein Dhh1p. *RNA* **11**: 1258–1270.
- Clark LB, Viswanathan P, Quigley G, Chiang Y-C, McMahon JS, Yao G, Chen J, Nelsbach A, Denis CL. 2004. Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. *J Biol Chem* **279**: 13616–13623.
- Collart MA. 2003. Global control of gene expression in yeast by the Ccr4-Not complex. *Gene* **313**: 1–16.
- Collart MA, Collart MA, Struhl K, Struhl K. 1993. CDC39, an essential nuclear protein that negatively regulates transcription and differentially affects the constitutive and inducible HIS3 promoters. *EMBO J* **12**: 177–186.
- Collart MA, Collart MA, Struhl K, Struhl K. 1994. NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. *Genes Dev* **8**: 525–537.
- Collart MA, Panasenko OO. 2012. The Ccr4--not complex. *Gene* **492**: 42–53.
- Collart MA, Timmers HTM. 2004. The eukaryotic Ccr4-not complex: a regulatory platform integrating mRNA metabolism with cellular signaling pathways? *Prog Nucleic Acid Res Mol Biol* **77**: 289–322.
- COOK WJ, Jeffrey LC, Xu Y, Chau V. 1993. Tertiary structures of class I ubiquitin-conjugating enzymes are highly conserved: crystal structure of yeast Ubc4. *Biochemistry* **32**: 13809–13817.
- COOK WJ, Tange Y, Jeffrey LC, Kurabayashi A, Xu Y, Goto B, Chau V, Hoe K-L, Kim D-U, Park H-O, et al. 2012. The CCR4-NOT complex is implicated in the viability of aneuploid yeasts. *PLoS Genet* **8**: e1002776.
- Cooper KF, Cooper KF, Scarnati MS, Scarnati MS, Krasley E, Krasley E, Mallory MJ, Mallory MJ, Jin C, Jin C, et al. 2012. Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction. *J Cell Sci* **125**: 1015–1026.
- Dahan N, Choder M. 2013. The eukaryotic transcriptional machinery regulates mRNA translation and decay in the cytoplasm. *Biochim Biophys Acta* **1829**: 169–173.
- DeBella LR, DeBella LR, Hayashi A, Hayashi A, Rose LS, Rose LS. 2006. LET-711, the *Caenorhabditis elegans* NOT1 ortholog, is required for spindle positioning and regulation of microtubule length in embryos. *Mol Biol Cell* **17**: 4911–4924.
- Deluen C, Deluen C, James N, James N, Maillet L, Maillet L, Molinete M, Molinete M, Theiler G, Theiler G, et al. 2002. The Ccr4-not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions. *Mol Cell Biol* **22**: 6735–6749.
- Denis CL, Denis CL, Chen J, Chen J. 2003. The CCR4-NOT complex plays diverse roles in mRNA metabolism. *Prog Nucleic Acid Res Mol Biol* **73**: 221–250.
- Dimitrova LN, Kuroha K, Tatematsu T, Inada T. 2009. Nascent peptide-dependent translation arrest leads to

Not4p-mediated protein degradation by the proteasome. *J Biol Chem* **284**: 10343–10352.

Dorf ME. 1981. *The Role of the Major Histocompatibility Complex in Immunobiology*.

Dori-Bachash M, Dori-Bachash M, Shema E, Shema E, Tirosh I, Tirosh I. 2011. Coupled evolution of transcription and mRNA degradation. *PLoS Biol* **9**: e1001106.

Draper MP, Draper MP, Liu HY, Liu HY, NELSBACH AH, NELSBACH AH, MOSLEY SP, MOSLEY SP, Denis CL, Denis CL. 1994. Ccr4 Is a Glucose-Regulated Transcription Factor Whose Leucine-Rich Repeat Binds Several Proteins Important for Placing Ccr4 in Its Proper Promoter Context. *Mol Cell Biol* **14**: 4522–4531.

Dupressoir A, Morel AP, Barbot W, Loireau MP, Corbo L, Heidmann T. 2001. Identification of four families of yCCR4- and Mg²⁺-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. *BMC Genomics* **2**: 9.

Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. 2007. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol Cell Biol* **27**: 3970–3981.

Ezzeddine N, Chang T-C, Zhu W, Yamashita A, Chen C-YA, Zhong Z, Yamashita Y, Zheng D, Shyu A-B. 2007. Human TOB, an antiproliferative transcription factor, is a poly(A)-binding protein-dependent positive regulator of cytoplasmic mRNA deadenylation. *Mol Cell Biol* **27**: 7791–7801.

Fabian MR, Fabian MR, Frank F, Frank F, Rouya C, Rouya C, Siddiqui N, Siddiqui N, Lai WS, Lai WS, et al. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol* **20**: 735–739.

Gaillard H, Gaillard H, Tous C, Tous C, Botet J, Botet J, González-Aguilera C, González-Aguilera C, Quintero MJ, Quintero MJ, et al. 2009. Genome-wide analysis of factors affecting transcription elongation and DNA repair: a new role for PAF and Ccr4-not in transcription-coupled repair. *PLoS Genet* **5**: e1000364.

Garapaty S, Garapaty S, Mahajan MA, Mahajan MA, Samuels HH, Samuels HH. 2008. Components of the CCR4-NOT complex function as nuclear hormone receptor coactivators via association with the NRC-interacting Factor NIF-1. *J Biol Chem* **283**: 6806–6816.

Garces RG, Garces RG, Gillon W, Gillon W, Pai EF, Pai EF. 2007. Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties. *Protein Sci* **16**: 176–188.

Goldstrohm AC, Hook BA, Seay DJ, Wickens M. 2006. PUF proteins bind Pop2p to regulate messenger RNAs. *Nat Struct Mol Biol* **13**: 533–539.

Goldstrohm AC, Seay DJ, Hook BA, Wickens M. 2007. PUF protein-mediated deadenylation is catalyzed by Ccr4p. *J Biol Chem* **282**: 109–114.

Goler-Baron V, Selitrennik M, Barkai O, Haimovich G, Lotan R, Choder M. 2008. Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. *Genes Dev* **22**: 2022–2027.

Grönholm J, Grönholm J, Kaustio M, Kaustio M, Myllymäki H, Myllymäki H, Kallio J, Kallio J, Saarikettu J, Saarikettu J, et al. 2012. Not4 enhances JAK/STAT pathway-dependent gene expression in *Drosophila* and in human cells. *FASEB J* **26**: 1239–1250.

Gulshan K, Thommandru B, Moye-Rowley WS. 2012. Proteolytic Degradation of the Yap1 Transcription Factor Is Regulated by Subcellular Localization and the E3 Ubiquitin Ligase Not4. *Journal of Biological Chemistry* **287**: 26796–26805.

Halter D, Collart MA, Panasenko OO. 2014. The Not4 E3 ligase and CCR4 deadenylase play distinct roles in protein quality control. *PLoS ONE* **9**: e86218.

Hanzawa H, Hanzawa H, de Ruwe MJ, de Ruwe MJ, Albert TK, Albert TK, van Der Vliet PC, van Der Vliet PC, Timmers HT, Timmers HT, et al. 2001. The structure of the C4C4 ring finger of human NOT4 reveals features distinct from those of C3HC4 RING fingers. *J Biol Chem* **276**: 10185–10190.

Harel-Sharvit L, Eldad N, Haimovich G, Barkai O, Duek L, Choder M. 2010. RNA polymerase II subunits link transcription and mRNA decay to translation. *Cell* **143**: 552–563.

- Henn A, Bradley MJ, La Cruz De EM. 2012. ATP utilization and RNA conformational rearrangement by DEAD-box proteins. *Annu Rev Biophys* **41**: 247–267.
- Hiraishi H, Shimada T, Ohtsu I, Sato T-A, Takagi H. 2009. The yeast ubiquitin ligase Rsp5 downregulates the alpha subunit of nascent polypeptide-associated complex Egd2 under stress conditions. *FEBS J* **276**: 5287–5297.
- Holm L, Rosenström P. 2010. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**: W545–9.
- Horiuchi M, Horiuchi M, Takeuchi K, Takeuchi K, Noda N, Noda N, Muroya N, Muroya N, Suzuki T, Suzuki T, et al. 2009. Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. *J Biol Chem* **284**: 13244–13255.
- Houseley J, Tollervey D. 2009. The many pathways of RNA degradation. *Cell* **136**: 763–776.
- Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* **12**: 99–110.
- Jonstrup AT, Jonstrup AT, Andersen KR, Andersen KR, Van LB, Van LB, Brodersen DE, Brodersen DE. 2007. The 1.4-Å crystal structure of the *S. pombe* Pop2p deadenylase subunit unveils the configuration of an active enzyme. *Nucleic Acids Res* **35**: 3153–3164.
- Kadlec J, Izaurralde E, Cusack S. 2004. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat Struct Mol Biol* **11**: 330–337.
- Kahvejian A, Roy G, Sonenberg N. 2001. The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harb Symp Quant Biol* **66**: 293–300.
- Kerr SC, Kerr SC, Azzouz N, Azzouz N, Fuchs SM, Fuchs SM, Collart MA, Collart MA, Strahl BD, Strahl BD, et al. 2011. The Ccr4-Not complex interacts with the mRNA export machinery. *PLoS ONE* **6**: e18302.
- Khusial P, Plaag R, Zieve GW. 2005. LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem Sci* **30**: 522–528.
- Kruk JA, Dutta A, Fu J, Gilmour DS, Reese JC. 2011. The multifunctional Ccr4-Not complex directly promotes transcription elongation. *Genes Dev* **25**: 581–593.
- Laribee RN, Shibata Y, Mersman DP, Collins SR, Kemmeren P, Roguev A, Weissman JS, Briggs SD, Krogan NJ, Strahl BD. 2007. CCR4/NOT complex associates with the proteasome and regulates histone methylation. *Proc Natl Acad Sci USA* **104**: 5836–5841.
- Lau N-C, Kolkman A, van Schaik FMA, Mulder KW, Pijnappel WWMP, Heck AJR, Timmers HTM. 2009. Human Ccr4-Not complexes contain variable deadenylase subunits. *Biochem J* **422**: 443–453.
- Lei J, Mesters JR, Brunn von A, Hilgenfeld R. 2011. Crystal structure of the middle domain of human poly(A)-binding protein-interacting protein 1. *Biochem Biophys Res Commun* **408**: 680–685.
- Lemaire M, Lemaire M, Collart MA, Collart MA. 2000. The TATA-binding protein-associated factor yTafII19p functionally interacts with components of the global transcriptional regulator Ccr4-Not complex and physically interacts with the Not5 subunit. *J Biol Chem* **275**: 26925–26934.
- Lenssen E, Oberholzer U, Labarre J, De Virgilio C, Collart MA. 2002. *Saccharomyces cerevisiae* Ccr4-not complex contributes to the control of Msn2p-dependent transcription by the Ras/cAMP pathway. *Mol Microbiol* **43**: 1023–1037.
- Leppek K, Schott J, Reitter S, Poetz F, Hammond MC, Stoecklin G. 2013. Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. *Cell* **153**: 869–881.
- Leung AKW, Nagai K, Li J. 2011. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* **473**: 536–539.
- Liu HY, Badarinarayana V, Audino DC, Rappsilber J, Mann M, Denis CL. 1998. The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *EMBO J* **17**: 1096–1106.

- Loh B, Loh B, Jonas S, Jonas S, Izaurralde E, Izaurralde E. 2013. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev* **27**: 2125–2138.
- Lotan R, Bar-On VG, Harel-Sharvit L, Duek L, Melamed D, Choder M. 2005. The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. *Genes Dev* **19**: 3004–3016.
- Mahadevan S, Struhl K. 1990. Tc, an unusual promoter element required for constitutive transcription of the yeast HIS3 gene. *Mol Cell Biol* **10**: 4447–4455.
- Maillet L, Maillet L, Collart MA, Collart MA. 2002. Interaction between Not1p, a component of the Ccr4-not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. *J Biol Chem* **277**: 2835–2842.
- Marcotrigiano J, Lomakin IB, Sonenberg N, Pestova TV, Hellen CU, Burley SK. 2001. A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol Cell* **7**: 193–203.
- Mathys H, Basquin J, Ozgur S, Czarnocki-Cieciura M, Bonneau F, Aartse A, Dziembowski A, Nowotny M, Conti E, Filipowicz W. 2014. Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. *Mol Cell* **54**: 751–765.
- Mauxion F, Mauxion F, Prève B, Prève B, Séraphin B, Séraphin B. 2013. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA Biol* **10**: 267–276.
- Mazza C, Ohno M, Segref A, Mattaj JW, Cusack S. 2001. Crystal structure of the human nuclear cap binding complex. *Mol Cell* **8**: 383–396.
- Mersman DP, Du HN, Fingerhahn IM, South PF, Briggs SD. 2009. Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. *Genes Dev* **23**: 951–962.
- Miller JE, Miller JE, Reese JC, Reese JC. 2012. Ccr4-Not complex: the control freak of eukaryotic cells. *Crit Rev Biochem Mol Biol* **47**: 315–333.
- Montpetit B, Thomsen ND, Helmke KJ, Seeliger MA, Berger JM, Weis K. 2011. A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* **472**: 238–242.
- Morel A-P, Morel A-P, Sentis S, Sentis S, Bianchin C, Bianchin C, Le Romancer M, Le Romancer M, Jonard L, Jonard L, et al. 2003. BTG2 antiproliferative protein interacts with the human CCR4 complex existing in vivo in three cell-cycle-regulated forms. *J Cell Sci* **116**: 2929–2936.
- Morita M, Morita M, Suzuki T, Suzuki T, Nakamura T, Nakamura T, Yokoyama K, Yokoyama K, Miyasaka T, Miyasaka T, et al. 2007. Depletion of mammalian CCR4b deadenylase triggers elevation of the p27Kip1 mRNA level and impairs cell growth. *Mol Cell Biol* **27**: 4980–4990.
- Mulder KW, Inagaki A, Cameron E, Mousson F, Winkler GS, De Virgilio C, Collart MA, Timmers HTM. 2007. Modulation of Ubc4p/Ubc5p-mediated stress responses by the RING-finger-dependent ubiquitin-protein ligase Not4p in *Saccharomyces cerevisiae*. *Genetics* **176**: 181–192.
- Mulder KW, Mersman DP, Winkler GS, Du H-N, Timmers HTM, Fingerhahn IM, South PF, Briggs SD. 2009. Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. *Genes Dev* **23**: 951–962.
- Nakamura T, Yao R, Ogawa T, Suzuki T, Ito C, Tsunekawa N, Inoue K, Ajima R, Miyasaka T, Yoshida Y, et al. 2004. Oligo-astheno-teratozoospermia in mice lacking Cnot7, a regulator of retinoid X receptor beta. *Nat Genet* **36**: 528–533.
- Nasertorabi F, Nasertorabi F, Batisse C, Batisse C, Diepholz M, Diepholz M, Suck D, Suck D, Böttcher B, Böttcher B. 2011. Insights into the structure of the CCR4-NOT complex by electron microscopy. *FEBS Lett* **585**: 2182–2186.
- Neely GG, Neely GG, Kuba K, Kuba K, Cammarato A, Cammarato A, Isobe K, Isobe K, Amann S, Amann S, et al. 2010. A global in vivo *Drosophila* RNAi screen identifies NOT3 as a conserved regulator of heart function. *Cell* **141**: 142–153.

- Oberholzer U, Oberholzer U, Collart MA, Collart MA. 1998. Characterization of NOT5 that encodes a new component of the Not protein complex. *Gene* **207**: 61–69.
- Ohn T, Chiang Y-C, Lee DJ, Yao G, Zhang C, Denis CL. 2007. CAF1 plays an important role in mRNA deadenylation separate from its contact to CCR4. *Nucleic Acids Res* **35**: 3002–3015.
- Okochi K, Suzuki T, Inoue J-I, Matsuda S, Yamamoto T. 2005. Interaction of anti-proliferative protein Tob with poly(A)-binding protein and inducible poly(A)-binding protein: implication of Tob in translational control. *Genes Cells* **10**: 151–163.
- Panasenko OO, Collart MA. 2011. Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29. *Mol Cell Biol* **31**: 1610–1623.
- Panasenko OO, David FPA, Collart MA. 2009. Ribosome association and stability of the nascent polypeptide-associated complex is dependent upon its own ubiquitination. *Genetics* **181**: 447–460.
- Panasenko OO, Panasenko OO, Collart MA, Collart MA. 2012. Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase. *Mol Microbiol* **83**: 640–653.
- Parker R, Sheth U. 2007. P bodies and the control of mRNA translation and degradation. *Mol Cell* **25**: 635–646.
- Petit A-P, Petit A-P, Wohlbold L, Wohlbold L, Bawankar P, Bawankar P, Huntzinger E, Huntzinger E, Schmidt S, Schmidt S, et al. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. *Nucleic Acids Res* **40**: 11058–11072.
- Prévôt D, Morel AP, Voeltzel T, Rostan MC, Rimokh R, Magaud JP, Corbo L. 2001. Relationships of the antiproliferative proteins BTG1 and BTG2 with CAF1, the human homolog of a component of the yeast CCR4 transcriptional complex: involvement in estrogen receptor alpha signaling pathway. *J Biol Chem* **276**: 9640–9648.
- Reese JC. 2013. The control of elongation by the yeast Ccr4-not complex. *Biochim Biophys Acta* **1829**: 127–133.
- Salgado-Garrido J, Bragado-Nilsson E, Kandels-Lewis S, Séraphin B. 1999. Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J* **18**: 3451–3462.
- Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA. 2002. Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol Cell Biol* **22**: 4723–4738.
- Sandler H, Sandler H, Kreth J, Kreth J, Timmers HTM, Timmers HTM, Stoecklin G, Stoecklin G. 2011. Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic Acids Res* **39**: 4373–4386.
- Schütz P, Bumann M, Oberholzer AE, Bieniossek C, Trachsel H, Altmann M, Baumann U. 2008. Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci USA* **105**: 9564–9569.
- Sun M, Sun M, Schwalb B, Schwalb B, Schulz D, Schulz D, Pirkl N, Pirkl N, Etzold S, Etzold S, et al. 2012. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res* **22**: 1350–1359.
- Temme C, Temme C, Zaessinger S, Zaessinger S, Meyer S, Meyer S, Simonelig M, Simonelig M, Wahle E, Wahle E. 2004. A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in *Drosophila*. *EMBO J* **23**: 2862–2871.
- Thore S, Thore S, Mauxion F, Mauxion F, Séraphin B, Séraphin B, Suck D, Suck D. 2003. X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. *EMBO Rep* **4**: 1150–1155.
- Törö I, Basquin J, Teo-Dreher H, Suck D. 2002. Archaeal Sm proteins form heptameric and hexameric complexes: crystal structures of the Sm1 and Sm2 proteins from the hyperthermophile *Archaeoglobus fulgidus*. *J Mol Biol* **320**: 129–142.
- Tucker M, Staples RR, Valencia-Sanchez MA, Muhlrud D, Parker R. 2002. Ccr4p is the catalytic subunit of a

Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. *EMBO J* **21**: 1427–1436.

Venturini G, Venturini G, Rose AM, Rose AM, Shah AZ, Shah AZ, Bhattacharya SS, Bhattacharya SS, Rivolta C, Rivolta C. 2012. CNOT3 is a modifier of PRPF31 mutations in retinitis pigmentosa with incomplete penetrance. *PLoS Genet* **8**: e1003040.

Viswanathan P, Viswanathan P, Ohn T, Ohn T, Chiang Y-C, Chiang Y-C, Chen J, Chen J, Denis CL, Denis CL. 2004. Mouse CAF1 can function as a processive deadenylase/3'-5'-exonuclease in vitro but in yeast the deadenylase function of CAF1 is not required for mRNA poly(A) removal. *J Biol Chem* **279**: 23988–23995.

Wahle E, Winkler GS. 2013. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829**: 561–570.

Wang H, Morita M, Yang X, Suzuki T, Yang W, Wang J, Ito K, Wang Q, Zhao C, Bartlam M, et al. 2010. Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *EMBO J* **29**: 2566–2576.

Weill L, Belloc E, Bava F-A, Méndez R. 2012. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat Struct Mol Biol* **19**: 577–585.

Winkler GS, Winkler GS, Albert TK, Albert TK, Dominguez C, Dominguez C, Legtenberg YIA, Legtenberg YIA, Boelens R, Boelens R, et al. 2004. An altered-specificity ubiquitin-conjugating enzyme/ubiquitin-protein ligase pair. *J Mol Biol* **337**: 157–165.

Winkler GS, Winkler GS, Balacco DL, Balacco DL. 2013. Heterogeneity and complexity within the nuclease module of the Ccr4-Not complex. *Front Genet* **4**: 296.

Yamashita A, Chang T-C, Yamashita Y, Zhu W, Zhong Z, Chen C-YA, Shyu A-B. 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol* **12**: 1054–1063.

Zekri L, Kuzuoğlu-Öztürk D, Izaurralde E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J* **32**: 1052–1065.

Zheng D, Ezzeddine N, Chen C-YA, Zhu W, He X, Shyu A-B. 2008. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *J Cell Biol* **182**: 89–101.

Zheng X, Zheng X, Dumitru R, Dumitru R, Lackford BL, Lackford BL, Freudenberg JM, Freudenberg JM, Singh AP, Singh AP, et al. 2012. Cnot1, Cnot2, and Cnot3 maintain mouse and human ESC identity and inhibit extraembryonic differentiation. *Stem Cells* **30**: 910–922.

Zielezinski A, Karlowski WM. 2015. Early origin and adaptive evolution of the GW182 protein family, the key component of RNA silencing in animals. *RNA Biol* **0**.

Etude structurale du complexe CCR4-NOT

Résumé

Le recyclage des ARN débute par un étape de déadénylation où la queue poly (A) est enzymatiquement clivée. La deadénylation est l'étape limitante dans le processus de dégradation des ARN. *In vivo* la deadénylation s'effectue successivement par les complexes multi-protéiques Pan2-Pan3 et Ccr4-Not. Le complexe Ccr4-not est conservé chez les eucaryotes et considéré comme le complexe prédominant responsable de l'activité de déadénylation dans la cellule. Le complexe est composé de neuf protéines organisées autour de la protéine d'échafaudage Not1. Le complexe comprend quatre modules distincts ; le module de déadénylation, le module Caf40, le module N-terminal et le module C-terminal. Mon mémoire de thèse regroupe les études structurales qui ont contribué à caractériser les structures des différents modules à la fois chez la levure et chez l'humain.

Complexe Ccr4-Not-Cristallographie aux rayons X -Caractérisation biochimique

Résumé en anglais

mRNA turnover begins with deadenylation wherein the poly(A) tail at the 3' end of the mRNA is removed. Deadenylation is the rate-limiting step of the decay pathway. *In vivo*, deadenylation is carried out by two major macromolecular complexes, the Pan2-Pan3 complex and the Ccr4-Not complex. The Ccr4-Not complex is a multi-protein complex that is evolutionarily conserved in all eukaryotes and is considered to be the major deadenylase complex in the cell. In *S. cerevisiae*, the Ccr4-Not complex is composed of nine subunits and is built around the scaffolding protein Not1. Structurally, the Ccr4-Not complex assembles into four separate modules with distinct domains of Not1 acting as a scaffold for individual modules. The four modules include the N-terminal module, the deadenylase module, the Caf40 module and the C-terminal module. With the exception of the C-terminal module, the architecture and biochemical role of all other modules of the yeast Ccr4-Not complex has been characterized. My doctoral thesis is focused on the elucidation of the architecture of the human and yeast Ccr4-Not complex.

Ccr4-Not complexe-X-ray crystallography-biochemical characterization