

UNIVERSITÉ DE SHERBROOKE

EXPRESSION, PURIFICATION ET CARACTÉRISATION STRUCTURALE ET
THERMODYNAMIQUE DES COMPLEXES FORMÉS DE LA RECONNAISSANCE
MOLÉCULAIRE DES DOMAINES B-HLH-LZ DES PROTÉINES HUMAINES
C-MYC ET MAX

par

FRANÇOIS-OLIVIER MC DUFF

Département de Pharmacologie

Mémoire présenté à la Faculté de médecine et des sciences de la santé
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.) en Pharmacologie

28 avril 2008

Ph.D. Pierre Lavigne, Département de Pharmacologie, Université de Sherbrooke
Ph.D. Michel Grandbois, Département de Pharmacologie, Université de Sherbrooke
Ph.D. Martin Bisailon, Département de Biochimie, Université de Sherbrooke

*Pour tous ceux et celles qui n'y auront pas crû,
Et a tous ceux et celles qui n'ont jamais cessé d'y croire.*

*Parce qu'à tout homme est donné,
Avec temps, travail et acharnement,
De prouver à un autre homme,
Qu'il a tort!*

*Merci maman, papa, Nico, Marie, Élie, Madeleine et François
Merci aux amis(e)... Etienne, Nic, André, Grease, Yanick et Anne-Marie
Merci aux membres du labo... JF, Simon, Mart, Fred, Gab, Luc... ah oui Rej!
Merci aux profs... Pierre, Michel, Martin, Marc et François
Et un merci spécial... à mononcle Guy!*

*Et surtout, un immense merci à ma chérie, Marie-Eve...
Pour tout ton amour, ton support et ta compréhension!*

*À la mémoire de mon grand chum Philippe Tartier (1975 - 2001)
Et à l'avenir de ma filleule Lydia, pour un monde toujours meilleur.*



TABLE DES MATIÈRES

TABLE DES MATIÈRES.....	iii
LISTE DES ILLUSTRATIONS.....	vii
LISTE DES SIGLES ET ABRÉVIATIONS.....	x
RÉSUMÉ.....	14
INTRODUCTION.....	15
1. HISTORIQUE.....	15
2. LES FACTEURS DE TRANSCRIPTION DU RÉSEAU MYC/MAX/MAD ET LE CANCER.....	17
2.1. Le motif b-HLH-LZ.....	20
2.1.1. Présentation.....	20
2.1.2. Nomenclature spécifique au texte.....	22
2.2. Biologie moléculaire du réseau Myc/Max/Mad.....	24
2.3. Problématique et présentation de l'article.....	29
3. SPECTROSCOPIE DE DICHROÏSME CIRCULAIRE.....	31
3.1. Introduction.....	31
3.2. Rayonnement électromagnétique, polarisation planaire et circulaire ..	32
3.2.1. Rayonnement électromagnétique.....	32
3.2.2. Réfringence et polarisation circulaire.....	33
3.2.3. Combinaison circulaire et polarisation planaire.....	35
3.3. Interactions matière : Biréfringence et Dichroïsme circulaire	36
3.3.1. Biréfringence circulaire.....	36
3.3.2. Dichroïsme circulaire.....	37
3.4. Chromophore amide, dichroïsme circulaire et structures secondaires.	40
3.4.1. Chromophore amide et dichroïsme circulaire.....	41

3.4.2.	Dichroïsme circulaire : hélice α	44
3.4.3.	Dichroïsme circulaire : feuillet β	46
3.4.4.	Dichroïsme circulaire : structures aléatoires r « <i>random coil</i> »	47
3.5.	CD vs analyse structurale et thermodynamique	48
4.	THERMODYNAMIQUE-COMPUTATIONNELLE.....	52
4.1.	L'analyse thermodynamique du système de c-Myc & Max	52
4.2.	Simulation via Gibbs-Helmholtz et Tableur CD_F.....	57
4.2.1.	Influence des intensités spectroscopiques extensives	63
4.2.2.	Détermination des populations via les équations de Gibbs-Helmholtz.....	65
4.2.3.	Influence de l'enthalpie standard de Van't Hoff : $\Delta H_{\text{ndu}}(T_{\text{ndu}})$	71
4.2.4.	Influence de la température de transition : T_{ndu}	73
4.2.5.	Influence de la variation de la capacité calorifique : $\Delta C_{p_{\text{ndu}}}$	74
4.2.6.	Tableur CD_F « <i>CD Fitting</i> »	75
4.2.7.	Irréductibilité des équations du système de c-Myc'SH/Max'SH.....	77
4.3.	Simulation de c-Myc'SH + Max'SH à l'aide du programme CD_SI	81
4.3.1.	Programme CD_SI « <i>CD Simulator Iterator</i> »	82
4.3.2.	Retro-calcul des concentrations de monomères et d'homodimères.....	86
4.3.3.	Résultats de la simulation à l'aide de CD_SI.....	89
4.3.4.	Calcul des $K_D(T)$ s « réels » et des $K_D^*(T)$ s « apparents ».....	91
4.4.	Simulation de l'équilibre global à l'aide du programme KD_I	94
4.4.1.	Programme KD_I « <i>KD Iterator</i> »	95
4.4.2.	Équilibre de K_D et populations.....	100
4.4.3.	Influences des concentrations et des K_D s homo et hétéro.....	102
5.	CONTRIBUTION DES AUTEURS À L'ARTICLE.....	110
	REVERSIBLE DIMERIZATION BETWEEN HUMAN C-MYC AND MAX B-HLH-LZ DOMAINS.....	111
	Résumé	112
	DISCUSSION ET CONCLUSION	113
6.	LES OBJECTIFS DE L'ÉTUDE	113
7.	CARACTÉRISATION STRUCTURALE ET RECONNAISSANCE MOLÉCULAIRE DE C-MYC ET MAX....	115

8.	THERMODYNAMIQUE DE L'ÉQUILIBRE GLOBAL INTRIQUÉ DE C-MYC ET MAX.....	118
9.	CONSIDÉRATIONS THÉRAPEUTIQUES.....	125
	MATÉRIEL ET MÉTHODES	130
10.	PRODUCTION DU B-HLH-LZ DE C-MYC	130
10.1.	Clonage et expression du b-HLH-LZ de c-Myc	130
10.1.1.	Clonage du b-HLH-LZ de c-Myc	130
10.1.2.	Expression du b-HLH-LZ de c-Myc	130
10.1.3.	Laemmli à bactéries	133
10.2.	Isolement du b-HLH-LZ de c-Myc.....	136
10.2.1.	Lyse et isolement des corps d'inclusion.....	138
10.2.2.	Solubilisation des corps d'inclusion et élimination des hauts poids moléculaires.	140
10.3.	Purification du surnageant par FPLC	141
10.3.1.	Chargement de la colonne.....	142
10.3.2.	Lavage à l'aide de Tampon U8	142
10.3.3.	Réalisation du gradient de force ionique.....	143
10.4.	Dessalage et lyophilisation.....	145
10.4.1.	Dessalage par Silice C8.....	145
10.4.2.	Dessalage par Exclusion de taille.....	146
10.4.3.	Lyophilisation	147
11.	PRODUCTION DU B-HLH-LZ DE MAX.....	148
11.1.	Clonage et Expression du b-HLH-LZ de Max	148
11.2.	Isolement du b-HLH-LZ de Max.....	150
11.3.	Purification du surnageant par FPLC	152
12.	PRODUCTION DES B-HLHS DE C-MYC ET MAX.....	153
12.1.	Clonage et expression des b-HLHs de c-Myc et Max.....	153
12.2.	Purification du b-HLH de Max.....	154
12.3.	Purification du b-HLH de c-Myc.....	154

BIBLIOGRAPHIE.....	157
ANNEXE I	I
Équations de Gibbs-Helmholtz pour un système monomérique à 2 états.....	I
ANNEXE II.....	II
Dérivation de P_U à partir de la fonction de partition d'un système homodimérique à 2 états.....	II
Dérivation de P_N à partir de la fonction de partition d'un système homodimérique à 2 états.....	III
Équations de Gibbs-Helmholtz pour un système homodimérique à 2 états.....	IV
ANNEXE III	VI
Dérivation de P_U à partir de la fonction de partition d'un système hétérodimérique à 2 états.	VI
Équations de Gibbs-Helmholtz pour un système hétérodimérique à 2 états.....	VII
ANNEXE IV	IX
Équations de Gibbs-Helmholtz pour un système monomérique à 3 états.....	IX
ANNEXE V	XI
Équations de Gibbs-Helmholtz pour un système homodimérique à 3 états.....	XI
ANNEXE VI	XIV
Équations de Gibbs-Helmholtz pour un système hétérodimérique à 3 états.....	XIV

LISTE DES ILLUSTRATIONS

Figure 1 : Schéma de la transcription par les membres du réseau Myc/Max/Mad.	19
Figure 2 : Représentation en ruban des domaines b-HLH-LZ.....	20
Figure 3 : Domaines fonctionnels des membres du réseau Myc/Max/Mad	24
Figure 4 : Activités « Histones Acetyl Transferase » et « Histones DeAcetylase »...	25
Figure 5 : Transactivation par c-Myc/Max	26
Figure 6 : Transrépression par Mad/Max	27
Figure 7 : Transrépression par c-Myc/Max.....	28
Figure 8 : Équilibres des espèces c-Myc et Max en présence de E-box.....	30
Figure 9 : Rayonnement électromagnétique polarisé linéairement à 45°	33
Figure 10 : Rayonnement électromagnétique polarisée circulairement à gauche ...	34
Figure 11 : Recomposition du rayonnement électromagnétique planaire.....	35
Figure 12 : Biréfringence circulaire de l'onde polarisée gauche (rotation horaire)	37
Figure 13 : Dichroïsme circulaire de l'onde polarisée gauche (résultante droite) ...	38
Figure 14 : Dichroïsme circulaire et biréfringence circulaire.....	39
Figure 15 : Spectres de dichroïsme circulaire des structures types α, β et r.....	40
Figure 16 : Chromophores amides et principales transitions : $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$....	43
Figure 17 : Hélice et spectre CD type de l'hélice α	45
Figure 18 : Brin β et spectre CD type du feuillet β antiparallèle.....	46
Figure 19 : Segment aléatoire et spectre CD type des formes aléatoires	47
Figure 20 : Différence de spectres types : $\alpha - r$, $\beta - r$ et $\alpha - \beta$.....	49
Figure 21 : Équilibre de Max'SH	54

Figure 22 : Équilibre de c-Myc'SH.....	54
Figure 23 : Équilibre global intriqué de c-Myc'SH et Max'SH.....	55
Figure 24 : Équilibres microscopiques de l'état natif d'un b-HLH-LZ.....	57
Figure 25 : Transition des populations d'un système à deux états	58
Figure 26 : Dénaturation thermique à l'équilibre thermodynamique	62
Figure 27 A) : Détermination des droites native et dénaturée (modèle à 2 états)	63
Figure 27 B) : Détermination de la droite intermédiaire (modèle à 3 états).....	64
Figure 28 : Influence de l'enthalpie standard de Van't Hoff sur la simulation	72
Figure 29 : Influence de la température de transition sur la simulation	73
Figure 30 : Influence de la variation de la capacité calorifique sur la simulation ...	74
Figure 31 : Exemple d'une simulation à l'aide des équations de Gibbs-Helmholtz	76
Figure 32 : Équilibre simplifié de c-Myc'SH et Max'SH	78
Figure 33 : Aperçu de la matrice d'exportation des résultats de CD_SI.....	83
Figure 34 : Reconnaissance moléculaire observée par CD.....	85
Figure 35 : Contributions individuelles des espèces au signal CD synthétique.....	90
Figure 36 : Concentrations individuelles des espèces en cours de dénaturation.....	91
Figure 37 : Distributions de population engendrées par des K_Ds homo et hétéro .	100
Figure 38 A) : Variations de population du complexe Max/Max en fonction des concentrations de c-Myc et Max	102
Figure 38 B) : Variations de population du complexe c-Myc/Max en fonction des concentrations de c-Myc et Max	103
Figure 39 A) : Variations de population du complexe c-Myc/Max en fonction de la concentration de protéine totale et du $K_{D, c-Myc/Max}$	105
Figure 39 B) : Variations de population du complexe c-Myc/Max en fonction de la concentration de protéine totale et du $K_{D, Max/Max}$	106

Figure 40 A) : Variations de population du complexe Max/Max en fonction de la concentration de protéine totale et du $K_{D, c-Myc/Max}$	107
Figure 40 B) : Variations de population du complexe Max/Max en fonction de la concentration de protéine totale et du $K_{D, Max/Max}$	108
Figure 41 : SDS-PAGE d'extrait crus de culture bactérienne.....	135
Figure 42 : SDS-PAGE de la purification du b-HLH-LZ de c-Myc.....	144

LISTE DES SIGLES ET ABRÉVIATIONS

^{13}C	Isotope carbone 13
^{15}N	Isotope azote 15
α	Hélice alpha; angle de rotation de la polarisation
AI	Souche bactérienne AI « Invitrogen »
AMP	Solution d'ampicilline
AUFS	« <i>Absorbance Unit Full Scale</i> »
b-HLH-LZ	« <i>basic-Helix-Loop-Helix-Leucine Zipper</i> »
β	Feuillet beta
β -Me	beta-mercaptoéthanol
C_α	Carbone alpha (carbone primaire des acides aminés)
C^*	Carbone asymétrique
C+	Souche bactérienne : Codons Plus « Stratagene »
CAM	Solution de chloramphénicol
CD	Dichroïsme Circulaire
CD_F	Tableur « <i>CD Fitting</i> »
CD_SI	Programme « <i>CD Simulator Iterator</i> »
CDK	Kinase Cyclines Dépendante
c-Myc'	Domaines b-HLH-LZ de c-Myc
c-Myc'SH	Domaines b-HLH-LZ de c-Myc réduit
c-Myc'SH/c-Myc'SH	Homodimère de b-HLH-LZ de c-Myc
c-Myc'SH/Max'SH	Hétérodimère de b-HLH-LZ de c-Myc et Max

c-Myc'SSc-Myc'	Homodimère covalent de b-HLH-LZ de c-Myc
c-Myc'SSMax'	Hétérodimère covalent de b-HLH-LZ de c-Myc et Max
ΔC_p	Variation de la capacité calorifique à pression constante
ΔG	Variation d'Énergie libre de Gibbs
ΔH	Variation d'Enthalpie
ΔS	Variation d'Entropie
D	État dissocié
DNaseI	DeoxyriboNucléase I, de pancréas bovin
DO	Densité optique
DTT	Dithiothréitol
ϵ	Coefficient d'extinction molaire; angle de l'ellipticité
e	Base du logarithme Népérien = 2,71828183
E-box	ADN « <i>Enhancer-Box</i> » (CANNTG)
ϕ	Angle de torsion du côté N-terminal du carbone α
GuHCl	Hydrochlorure de guanidine
HEPES	N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HPLC	Chromatographie liquide à haute pression
IPTG	Isopropyl- β -D-thiogalactopyranoside
K	Kelvin
K_D	Constante de dénaturation/dissociation
KD_I	Programme « <i>KD Iterator</i> »
λ	Longueur d'onde
LB	Milieu de culture Luria Bertani

M9	Milieu de culture minimal
Max'	Domaines b-HLH-LZ de Max
Max'SH	Domaines b-HLH-LZ de Max réduit
Max'SH/Max'SH	Homodimère de b-HLH-LZ de Max
Max'SSMax'	Homodimère covalent de b-HLH-LZ de Max
MCD	Signal CD molaire
<i>ndu</i>	indice « <i>native-dissociation-unfolding</i> »
<i>nidu</i>	indice « <i>native-intermediate-dissociation-unfolding</i> »
<i>nu</i>	indice « <i>native-unfolding</i> »
nm	nanomètre
N	État natif
NaAc	Acétate de sodium
ORD	Dispersion Rotatoire Optique (Biréfringence)
PAGE	Électrophorèse en gel de polyacrylamide
P	Population, probabilité
PCR	Réaction de polymérisation en chaîne
PEI	Polyéthylèneimine
pI	Point isoélectrique
θ	Signal spectroscopique, ellipticité
Q	Fonction de répartition d'un système
r	Structure aléatoire « <i>random coil</i> »
R	Constante des gaz parfaits = $8,314472 \text{ J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$
RMN	Résonance Magnétique Nucléaire

SDS	Sodium dodécyl sulfate (Lauryl Sulfate)
SID	« <i>mSin3 Interaction Domaine</i> »
TAD	« <i>TransActivation Domain</i> »
Tris	Tris(hydroxyméthyl)aminométhane
U	État dénaturé « <i>unfolded</i> »
UV	Lumière (Ultra-Violet)
ψ	Angle de torsion côté C-terminal du carbone α
[]	Concentration

RÉSUMÉ

EXPRESSION, PURIFICATION ET CARACTÉRISATION STRUCTURALE ET THERMODYNAMIQUE DES COMPLEXES FORMÉS DE LA RECONNAISSANCE MOLÉCULAIRE DES DOMAINES B-HLH-LZ DES PROTÉINES HUMAINES C-MYC ET MAX

FRANÇOIS-OLIVIER MC DUFF, Faculté de Médecine, Université de Sherbrooke, Département de pharmacologie.

Mémoire présenté à la Faculté de médecine en vue de l'obtention du grade de maître ès sciences (M.Sc.)

L'onco-protéine c-Myc est un facteur de transcription de type b-HLH-LZ "*basic-Helix-Loop-Helix-Leucine Zipper*" qui fait partie du réseau de facteur de transcription Myc/Max/Mad. Sous la forme d'hétérodimère avec Max (son partenaire obligatoire), c-Myc peut activer et réprimer la transcription. c-Myc est surexprimé dans les cellules cancéreuses et active la transcription de gènes essentiels à la croissance et la prolifération cellulaire et réprime les gènes menant à la différenciation et l'arrêt du cycle cellulaire. Le b-HLH-LZ de c-Myc fut démontré comme critique à la fois pour la transactivation en complexe avec l'ADN, de même que pour la transrépression suite à la formation d'un complexe avec la protéine Miz-1. On dénombre actuellement plus de 1500 gènes dont la transcription est régulée par c-Myc. La caractérisation de la dynamique de reconnaissance moléculaire des b-HLH-LZs de c-Myc et Max (étape initiale des fonctions cellulaires de c-Myc) est donc primordiale à la compréhension des mécanismes permettant l'édification et la suppression des complexes supérieurs d'activation et de répression de la transcription par c-Myc.

Dans le cadre de cette étude, nous avons cloné, surexprimé et purifié le domaine b-HLH-LZ de la protéine c-Myc humaine, uniformément marqué ^{13}C et ^{15}N , et avons formé son complexe hétérodimérique/hétéroisotopique avec le b-HLH-LZ de Max produit en abondance naturelle. Nous avons confirmé l'hétérodimérisation spécifique par dichroïsme circulaire (CD) et résonance magnétique nucléaire (RMN), et nos résultats montrent que le complexe c-Myc/Max est replié convenablement en absence d'ADN et lie de façon dimérique les cibles E-box. Nous montrons aussi que l'hétérodimérisation est optimale à pH 4.8 et inexistante à pH acide 2.8. Tel que discuté dans l'article joint, ces résultats suggèrent que des résidus glutamates sur c-Myc et une histidine sur Max sont impliqués dans l'hétérodimérisation spécifique. Pour réaliser une analyse thermodynamique adéquate du système dissociatif de c-Myc et Max, nous avons développé une nouvelle approche computationnelle qui permet d'estimer les populations simultanées d'hétérodimère et d'homodimère à partir de dénaturations thermiques suivies par CD. Tel qu'anticipé, nous montrons avec cette approche qu'une population d'homodimère Max/Max compétitionne avec l'hétérodimère en solution. Pris dans leur ensemble, nos résultats représentent la caractérisation la plus complète à ce jour de l'hétérodimérisation spécifique entre c-Myc et Max et mettent en relief la possibilité de l'existence d'une population non négligeable d'homodimère Max/Max *in vivo*.

n-myc « neuroblastome », (HIRVONEN *et al.*, 1993) et *s-myc* « suppressor » (SUGIYAMA *et al.*, 1989; KAGAYA *et al.*, 1997).

À l'époque, le pouvoir transformant de *v-myc* chez les volailles laissa présumer que les homologues humains pouvaient eux aussi être responsables de dérégulations de la prolifération cellulaire comme observées dans les cancers humains. Le séquençage de *v-myc* (ALITALO *et al.*, 1983) et de *c-myc*, (COLBY *et al.*, 1983; BERNARD *et al.*, 1983, WATT *et al.*, 1983) révéla les similitudes des deux transcrits de gènes et dès lors, *c-Myc* fut suspectée comme un proto-oncogène. Des études ultérieures permirent d'identifier des régions de *c-Myc* impliquées dans la transformation cellulaire et la localisation nucléaire de la protéine (STONE *et al.*, 1987).

2. LES FACTEURS DE TRANSCRIPTION DU RÉSEAU MYC/MAX/MAD ET LE CANCER

Les protéines c-Myc, Max et Mad sont des facteurs de transcription qui régulent l'expression de gènes impliqués, entre autres, dans la prolifération et l'arrêt du cycle cellulaire. On dénombre actuellement plus de 1500 gènes sous la gouverne de ce réseau de facteurs de transcription appelé : « Réseau Myc/Max/Mad » (LI *et al.*, 2003). En effet, des analyses de localisation « *location-analysis* » à l'échelle du génome sur des lignées cellulaires du lymphome de Burkitt, révèlent que les patrons d'expression de toutes protéines confondues corrélerent avec la liaison de c-Myc/Max aux promoteurs des gènes (LI *et al.*, 2003). Dans les lymphomes de Burkitt, la dérégulation de l'expression de c-Myc est causée par une translocation du gène *c-myc* en aval d'un promoteur fort d'immunoglobuline (KLEIN *et al.*, 1968). Parmi les 4839 promoteurs arbitrairement retenus pour l'étude on a retrouvé l'hétérodimère c-Myc/Max lié à près de 15% des cibles. La liaison de c-Myc corréle fortement avec ses activités de transcription dans l'ensemble du génome et celles-ci ressemblent d'avantage au profil de liaison génomique de la Polymérase II qu'à celui d'un facteur de transcription séquence spécifique dans les mêmes cellules. c-Myc et Max colocalisent fortement avec le facteur de transcription TFIID, ce qui laisse croire que c-Myc occuperait en réalité un rôle prépondérant dans la régulation de l'expression globale de tous les gènes dans les lymphomes de Burkitt (LI *et al.*, 2003).

Entre autres, c-Myc active la transcription de cyclines, de CDKs (Kinases Cyclines Dépendantes) et de la sous-unité catalytique de hTERT. Ces gènes sont respectivement impliqués dans la prolifération et l'immortalisation cellulaire (SAKAMURO et PRENDERGAST, 1999).

c-Myc réprime la transcription de gènes menant normalement à l'arrêt du cycle cellulaire (e.g. p15, p21). Les protéines Mad quant à elles, antagonisent les effets de c-Myc en réprimant une majorité des gènes activés par ce dernier (GRANDORI *et al.*, 2000). On ne connaît pas à ce jour d'effet de la famille Mad sur la répression de la transcription par c-Myc.

Pour lier l'ADN et contrôler l'expression de leurs gènes cibles, c-Myc et Mad doivent obligatoirement hétérodimériser avec la protéine Max. Des variations des niveaux d'expression des différents membres du réseau Myc/Max/Mad ont été observées dans plusieurs types cellulaires et démontrent une corrélation entre l'expression de c-Myc et la prolifération cellulaire comparable à la différenciation observée en présence de Mad (MARIADASON *et al.*, 2005).

Le mécanisme par lequel c-Myc active la transcription fut longtemps incompris, faute de l'identification de son partenaire obligatoire, la protéine Max (nommée ainsi pour « *Myc Associated protein X* »), découverte en 1991 par l'équipe de Blackwood et Eisenman (BLACKWOOD *et al.*, 1991). c-Myc étant incapable d'homodimériser pour lier seule l'ADN, Max agit pour elle comme partenaire d'hétérodimérisation et de liaison

à l'ADN. Dans les faits, les protéines Myc n'ont jamais pu être isolées seules *in-vivo* (ADHIKARY et EILERS, 2005). Dans les années qui suivirent, des groupes comme ceux d'Hurlin et Eisenman découvrirent que Max était aussi le partenaire obligatoire des protéines de la famille Mad (HURLIN *et al.*, 1995). Ainsi, à partir du milieu des années 90, un modèle pour l'implication du réseau Myc/Max/Mad dans la biologie cellulaire normale et tumorale devint apparent (Figure 1).

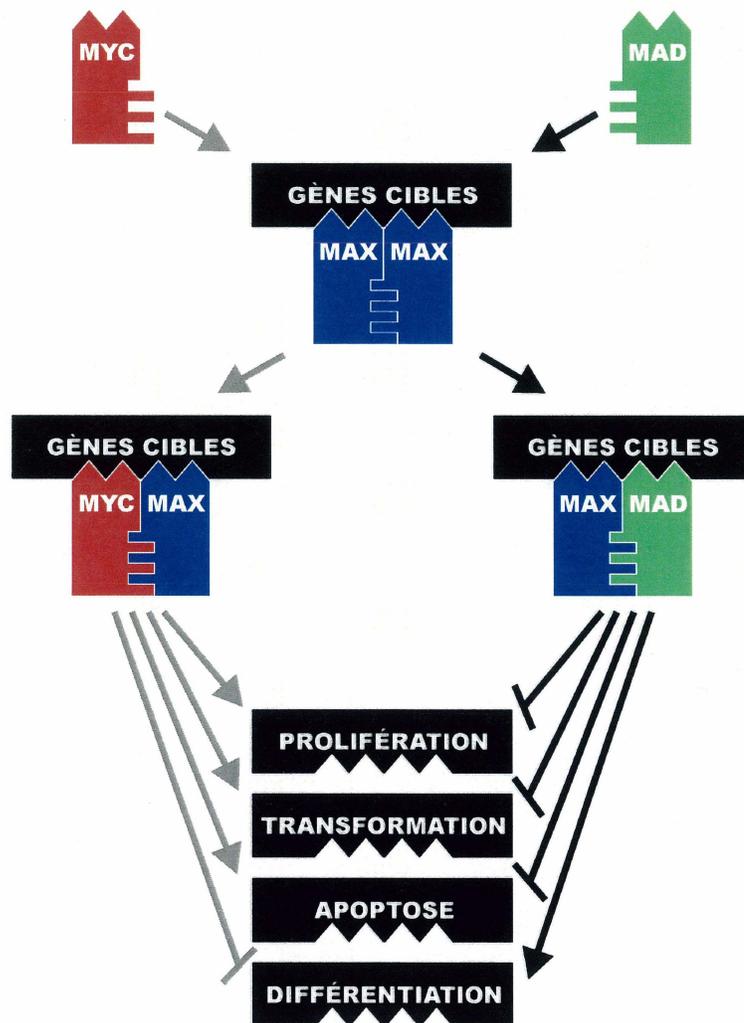


Figure 1 : Schéma de la transcription par les membres du réseau Myc/Max/Mad
(Adapté de ROTTMANN et LÜSCHER, 2006).

2.1. Le motif b-HLH-LZ

2.1.1. Présentation

Les protéines c-Myc, Max et Mad font partie de la famille des protéines du type b-HLH-LZ, pour « *basic-Helix-Loop-Helix-Leucine-Zipper* ». Le motif b-HLH-LZ est une conformation structurale qui permet à la fois la dimérisation des protéines entre elles par leurs domaines HLHs et LZs, et la liaison des partenaires à des cibles d'ADN double brins *via* leurs régions basiques.

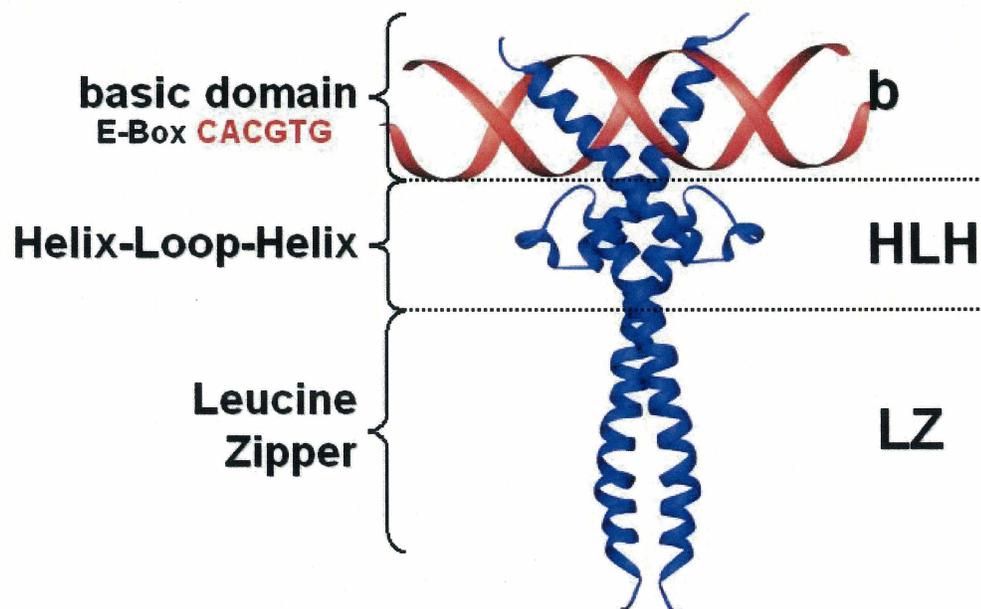


Figure 2 : Représentation en ruban des domaines b-HLH-LZ d'un homodimère Max (bleu) lié à l'ADN (rouge). Illustration réalisée à partir de la structure 1R05.pdb : Protein Data Bank, (<http://www.rcsb.org>) (SAUVÉ *et al.*, 2003).

Les régions basiques forment les principaux domaines de liaison au sillon majeur de l'ADN, bien que l'on observe également des contacts entre ce dernier et les HLHs. Les domaines HLHs des partenaires forment conjointement un ballot de quatre hélices α au coeur hydrophobe, tandis que les leucines zippers complètent l'interface de dimérisation protéine/protéine en formant une super hélice enchâssée de résidus leucines. Le b-HLH-LZ de c-Myc fut démontré comme critique à la fois pour la transactivation en complexe avec l'ADN, (AMATI *et al.*, 1993; MAO *et al.*, 2003) de même que pour la transrépression suite à la formation d'un complexe avec un autre facteur de transcription, la protéine Miz-1 (PEUKERT *et al.*, 1997).

Les b-HLH-LZ de c-Myc, Max et Mad reconnaissent spécifiquement des cibles d'ADN dont la séquence est 5'-CACA/GTG-3' et que l'on appelle E-box pour « *Enhancer box* » (NAIR et BURLEY, 2003; AMATI *et al.*, 2001). La séquence E-box reconnue par c-Myc, Max et Mad, représente une ponctuation de la séquence palindromique consensus 5'-CANNTG-3' reconnue invariablement par tous les facteurs de transcription de type b-HLH (ADHIKARY et EILERS, 2005). De récentes études estiment que le génome humain comporterait 25 000 sites E-box pouvant être reconnus par c-Myc, soit plus de sites qu'il n'y a de molécules c-Myc (≈ 3 à 5000) dans une cellule en prolifération. c-Myc régule donc potentiellement une partie non-négligeable des quelques 30 000 gènes connus chez l'homme (ADHIKARY et EILERS, 2005). Une évidence ressortant de l'estimation précédente est que pour accomplir l'ensemble de sa tâche, c-Myc doit agir par « *frappe et court* ». Ainsi, les mécanismes de régulation de c-Myc doivent avoir des effets de plus longue durée que son temps de présence en chacun de ses sites. La

conclusion directe de ce raisonnement est que c-Myc doit manifestement s'articuler dans une dynamique complexe d'échange de partenaire et de sites, et donc, dans une valse de reconnaissance moléculaire complexe.

2.1.2. Nomenclature spécifique au texte

Nous présentons ici une courte nomenclature précisant les différentes dénominations utilisées au long du texte et de l'article, pour préciser les protéines et la nature des complexes en présence.

Nous référerons donc plus spécifiquement : « aux domaines b-HLH-LZ uniquement » des protéines c-Myc, Max et Mad en ajoutant l'apostrophe « ' » pour signifier qu'il s'agit de versions raccourcies (altérées) des protéines complètes. Nous retrouverons ainsi, pour les protéines complètes (et/ou lorsque le texte ne requiert pas de précision particulière), les nomenclatures c-Myc, Max et Mad, et c-Myc', Max' et Mad' pour les domaines b-HLH-LZ uniquement.

De plus, pour les besoins de l'étude, les gènes encodant les domaines b-HLH-LZ des protéines c-Myc et Max furent clonées de manière à produire des protéines recombinantes jonchées d'une séquence (Gly-Ser-Gly-Cys) en C-terminal. Cet artéfact de clonage a eut pour but d'introduire une portion « *linker* » et permettre de contrôler à volonté les assemblages des complexes homodimériques et hétérodimériques, réalisés de

façon spécifique *via* une liaison covalente du pont disulfure entre les cystéines de deux partenaires. Pour préciser la nature des états d'oxydation des cystéines, nous recourons donc à la nomenclature « SH » lorsque la ou les cystéines sont réduites et les protéines non liées, et « SS » lorsqu'un pont aura volontairement été établie par l'oxydation spécifique de deux partenaires.

Pour spécifier la nature dissociative ou non dissociative des complexes, ceux-ci sont écrit avec l'ajout d'une diagonale « / » pour signifier ou rehausser le fait que le complexe est dissociable. Ainsi, les complexes pourront être écrits tel : « Max'/Max' » ou « Max'SH/Max'SH », pour un homodimère dissociatif des domaines b-HLH-LZ de Max réduit et « Max'SSMax' » pour le complexe covalent oxydé non-dissociatif. (*Pour plus de précision, veuillez référer à la : Liste des Sigles et Abréviations, p. x à xiii*)

2.2. Biologie moléculaire du réseau Myc/Max/Mad

Les principales fonctions des membres du réseau sont actuellement attribuées aux partenaires Myc et Mad. Cependant, de récentes évidences expérimentales (NAUD *et al.*, 2005) ainsi que les résultats présentés ici permettent de proposer des rôles prépondérants pour Max dans les mécanismes de régulation de la transcription. Dans un souci de concision, nous limiterons pour le moment la discussion aux cas des membres c-Myc et Mad.

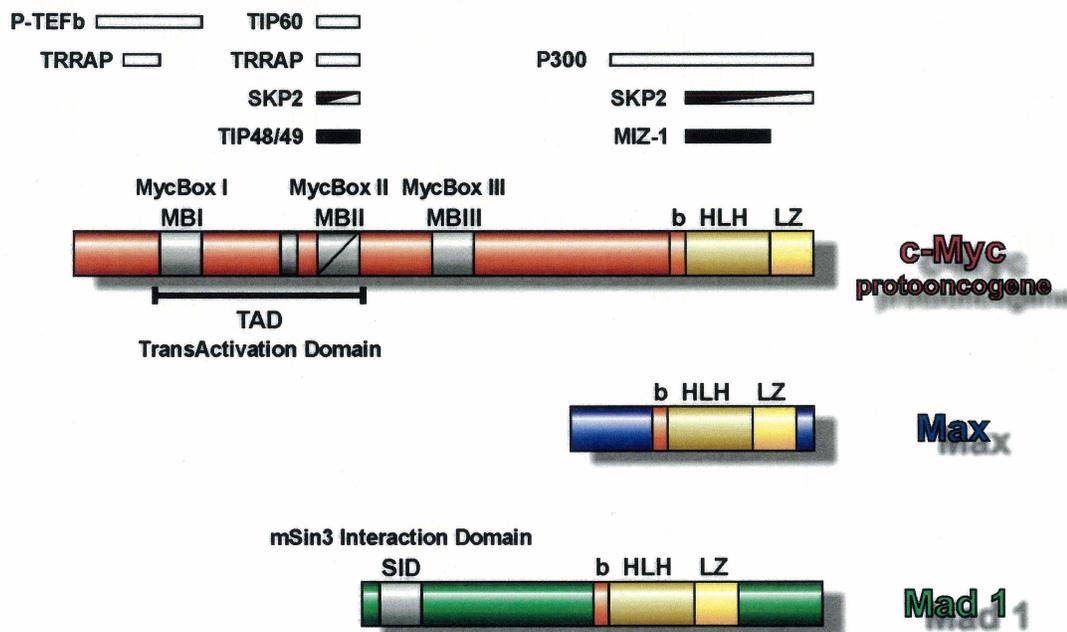


Figure 3 : Domaines fonctionnels des membres du réseau Myc/Max/Mad Les régions contacts de c-Myc avec d'autres protéines fonctionnelles hors réseau sont représentées en superposition. Les boîtes blanches représentent les contacts qui contribuent aux activités de transactivation alors que les boîtes noires sont des contributions à la transrépression (adaptée de (ADHIKARY et EILERS, 2005)).

En plus d'exhiber les domaines b-HLH-LZ qui leur permettent de dimériser avec la protéine Max et de lier l'ADN, c-Myc et Mad présentent tous deux des domaines spécifiques leur permettant d'exercer leurs fonctions cellulaires propres.

La région N-terminale de c-Myc présente trois éléments hautement conservés et connus sous les noms de « *Myc Boxes* » (boîtes Mycs, MBI-III). Conjointement, les éléments MBI et MB-II constituent le domaine appelé TAD pour « *TransActivation Domain* », par lequel c-Myc recrute le complexe de la protéine TRRAP (NIKIFOROV *et al.*, 2002) et les activités HAT d'acétylation des histones « *Histone Acetyl Transferase* ». L'acétylation des histones permet l'ouverture de la chromatine et ce faisant, donne plein accès à la machinerie de transcription des gènes.

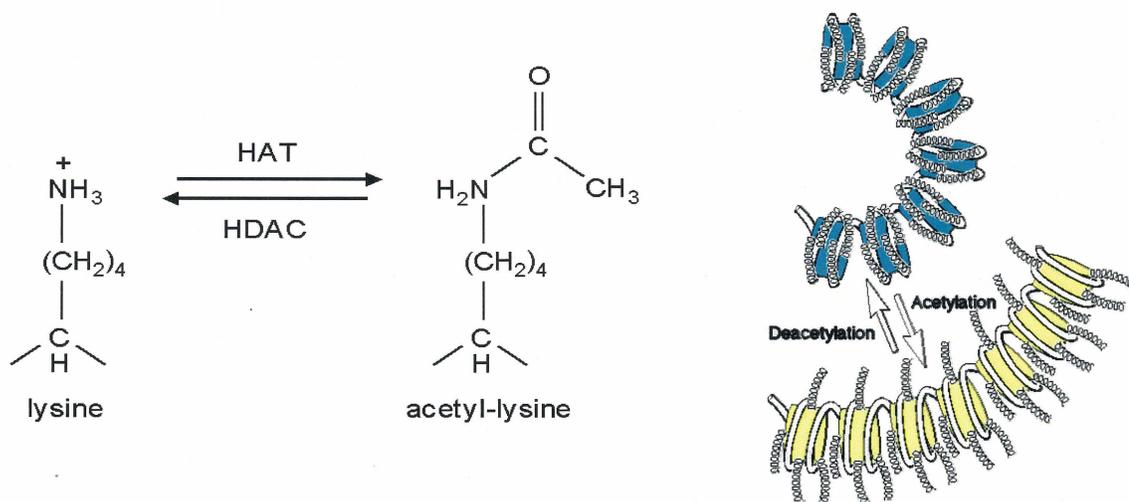


Figure 4 : Activités « Histones Acetyl Transferase » et « Histones DeAcetylase » permettant l'ouverture et la fermeture de la chromatine.

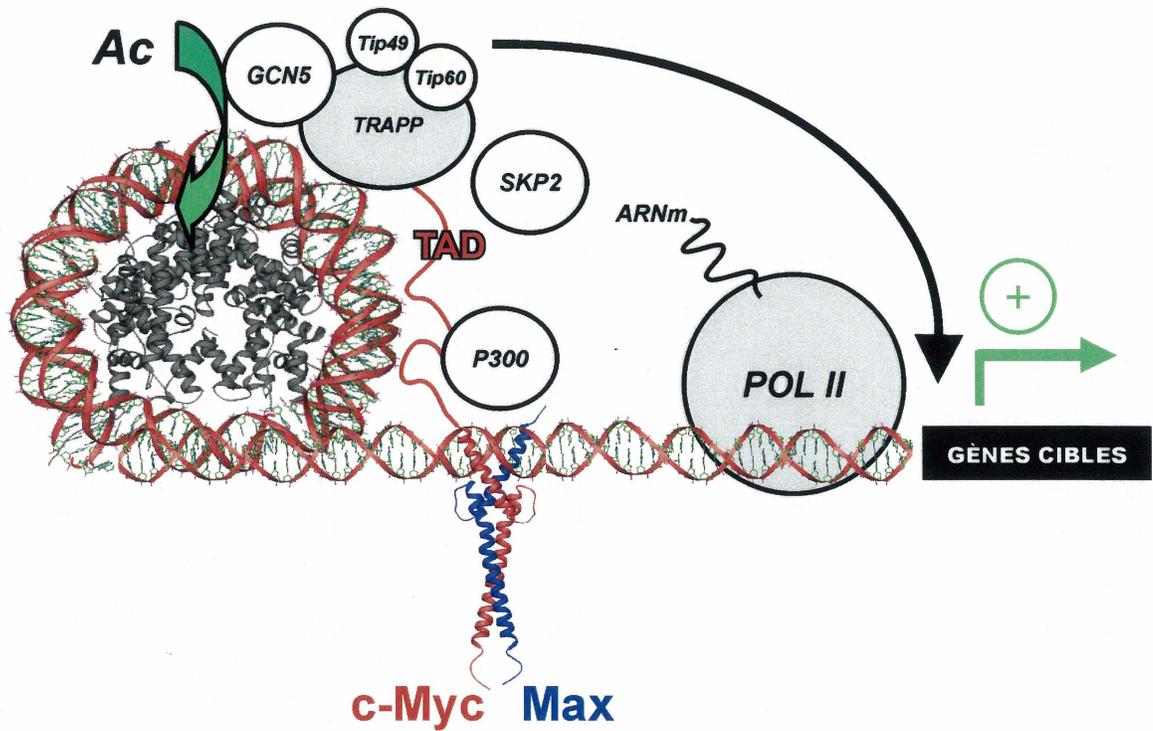


Figure 5 : Transactivation par c-Myc/Max via le recrutement du complexe de la protéine TRRAP et de GCN5 possédant l'activité HAT qui permet l'ouverture de la chromatine (NIKIFOROV *et al.*, 2002).

En réponse à c-Myc, Mad présente pour sa part le domaine SID « *mSin3 Interaction Domain* » qui tient lieu de domaine d'interaction avec la protéine mSin3 possédant l'activité HDAC « *Histone DeAcetylase* » permettant l'interruption de la transcription par le réemballage des promoteurs des gènes dans la chromatine.

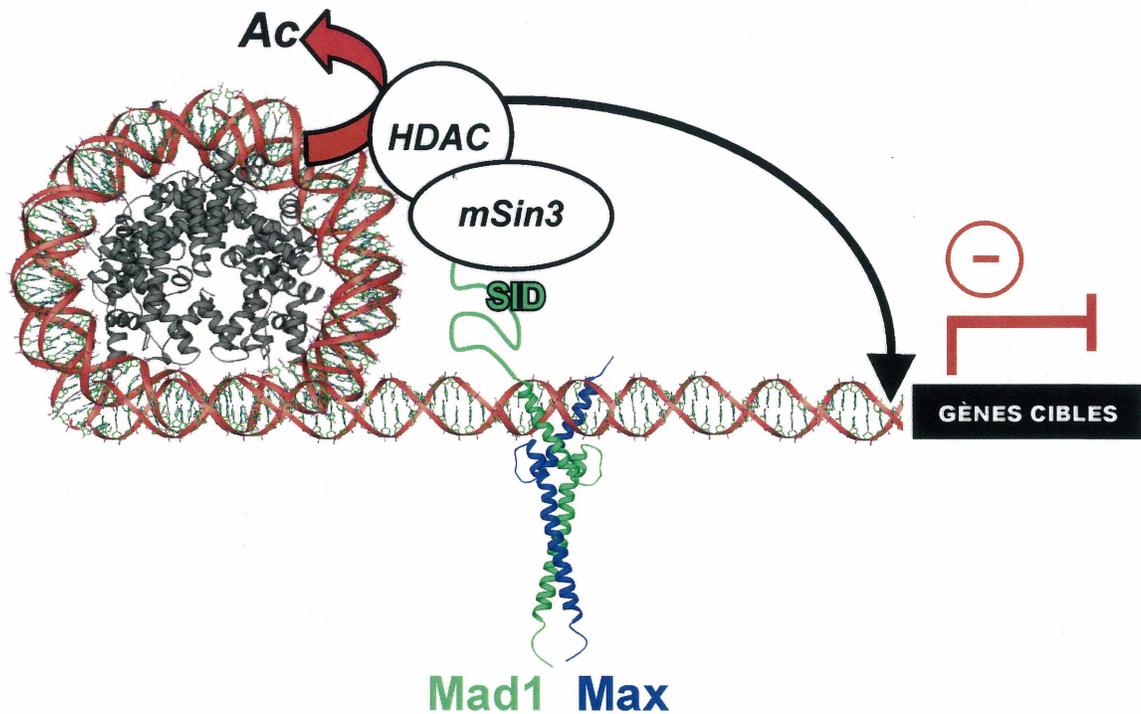


Figure 6 : Transrépression par Mad/Max via le recrutement de la protéine mSin3 et de son activité HDAC qui referme la chromatine des gènes (ROTTMANN et LÜSCHER, 2006).

Parmi les nombreuses autres fonctions connues de c-Myc (*pour une revue récente du sujet voir ADHIKARY et EILERS, 2005*), c-Myc peut également réprimer la transcription de gènes cytotostatiques (*e.g. p15, p21, Mad4*) en liant au niveau de son HLH la protéine Miz-1 et en déplaçant le cofacteur P300 essentiel à Miz pour l'activation de la transcription. Un second mécanisme non-exclusif du premier pourrait faire intervenir le recrutement de Dnmt3a par le TAD de c-Myc. La Dnmt3a possède des activités d'ADN méthyltransférase pouvant inhiber la transcription.

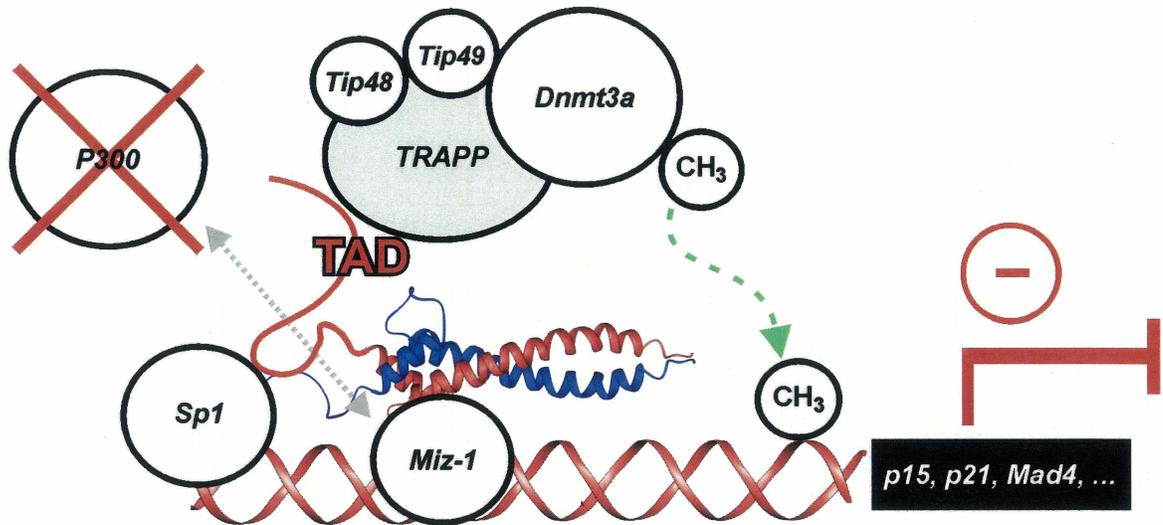


Figure 7 : Transrépression par c-Myc/Max de gènes cyostatiques. c-Myc agit sur Miz-1 en liant la protéine par son HLH et en déplaçant le cofacteur P300. Un second mécanisme représenté consisterait en le recrutement de la protéine Dnmt3a et de son activité d'ADN méthyltransférase pouvant inhiber la transcription (ADHIKARY et EILERS, 2005).

2.3.Problématique et présentation de l'article

Bien que la structure de l'hétérodimère de b-HLH-LZ de c-Myc/Max soit connue (NAIR et BURLEY, 2003), sa structure en absence d'ADN elle, ne l'est pas. Les déterminants structuraux responsables de l'hétérodimérisation spécifique en absence d'ADN sont encore aujourd'hui débattus et la caractérisation complète du processus d'hétérodimérisation est compliquée par le fait qu'*in vitro* et probablement davantage encore *in vivo*, l'homodimérisation de Max compétitionne contre l'hétérodimérisation de c-Myc/Max. Les mécanismes sous-jacents de cette reconnaissance moléculaire représentent une pierre d'assise pour la compréhension des activités de transactivation et de transrépression de c-Myc et de son rôle dans la tumorigénèse. La découverte des déterminants structuraux de la reconnaissance moléculaire entre c-Myc, Max et leurs partenaires est cruciale pour le développement d'éventuels inhibiteurs de la transactivation et de la transrépression par c-Myc.

Dans ce contexte nous avons procédé au clonage, à l'expression et la purification des domaines b-HLH-LZ de c-Myc sous forme de protéine recombinante dans *Escherichia coli*. Nous avons confirmé l'hétérodimérisation spécifique par dichroïsme circulaire (CD) et résonance magnétique nucléaire (RMN). Nous montrons que l'hétérodimérisation est optimale à pH 4,8 et inexistante à pH acide 2,8. Tel que discuté dans l'article joint au mémoire, ces résultats suggèrent que des résidus glutamates sur c-Myc et une histidine sur Max sont impliqués dans l'hétérodimérisation spécifique. Nous avons aussi développé une nouvelle approche computationnelle qui permet d'estimer les

populations simultanées d'hétérodimère et d'homodimère à partir de dénaturations thermiques suivies par CD. Tel qu'anticipé nous montrons avec cette approche computationnelle, qu'une population d'homodimère Max/Max compétitionne avec l'hétérodimère. Pris dans leur ensemble nos résultats représentent la caractérisation la plus complète à ce jour de l'hétérodimérisation spécifique entre c-Myc et Max et mettent en relief la possibilité de l'existence d'une population non-négligeable d'homodimère Max/Max *in vivo*.

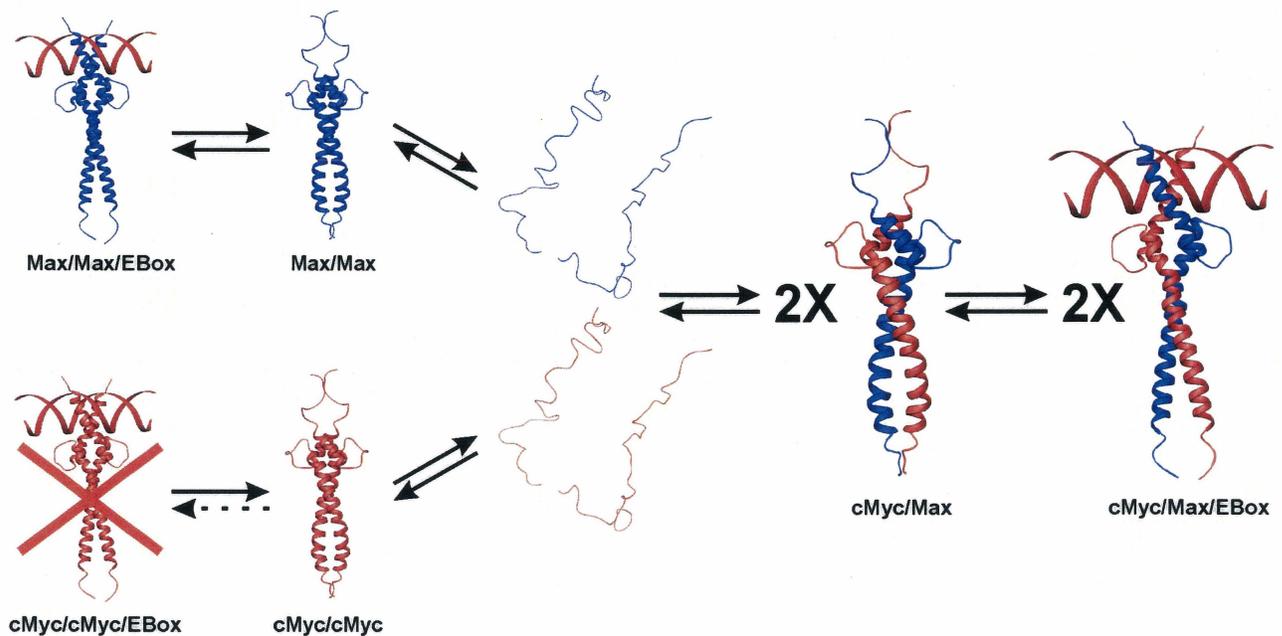


Figure 8 : Équilibres des espèces c-Myc et Max en présence de E-box

Avant de présenter cet article soumis à la revue « *Journal of Molecular Biology* » (JMB), nous procédons ici à l'introduction de la spectroscopie de dichroïsme circulaire (CD) utilisée pour cette étude, de même qu'à la présentation de la thermodynamique sous-jacente à notre approche computationnelle.

3. SPECTROSCOPIE DE DICHROÏSME CIRCULAIRE

3.1.Introduction

La spectroscopie de dichroïsme circulaire (CD; dichroïsme du grec « *dikhroos* » : de deux couleurs) est basée sur la spectroscopie conventionnelle d'absorption, mais mesure une différence d'absorption (absorption préférentielle) entre deux faisceaux de lumière polarisés circulairement, anti-sens l'un par rapport à l'autre. Les faisceaux, de même longueur d'onde (principalement UV pour les protéines), sont polarisés circulairement à droite et à gauche, et leur absorption préférentielle par certaines macromolécules (*e.g.* protéines, acides nucléiques, glucides), résulte d'une polarisation électrique et magnétique des transitions électroniques (BEYCHOK, 1966).

Comme nous le verrons plus loin, le dichroïsme circulaire implique d'abord une bande d'absorption lumineuse, mais rares sont les bandes d'absorption qui impliquent un phénomène de dichroïsme. En fonction de la longueur d'onde, les structures secondaires, tertiaires et quaternaires des protéines et des acides nucléiques renforcent ou atténuent cette propriété d'absorber préférentiellement une lumière plus que l'autre. L'appareil mesure ainsi, une différence d'absorption (de l'ordre de 0,03% (KELLY *et al.*, 2005)) entre la lumière transmise polarisée circulairement à gauche et celle polarisée circulairement à droite.

3.2. Rayonnement électromagnétique, polarisation plane et circulaire

Nombreuse sont les « diverses » interactions de la lumière avec la matière et les explications de ces différents phénomènes font appels aux équations de la mécanique quantique. Bien entendu, ceci dépasse de loin le cadre de cet ouvrage, mais il nous est tout de même possible d'appréhender conceptuellement les effets des certains processus d'interaction. (*Pour une introduction à la mécanique quantique du dichroïsme circulaire voir : (CANTOR et SCHIMMEL, 1980).*)

3.2.1. Rayonnement électromagnétique

Le rayonnement électromagnétique peut être vu comme la « résultante » de deux ondes sinusoïdales perpendiculaires : une onde électrique verticale et une onde magnétique horizontale, se propageant dans l'espace en phase l'une avec l'autre (Figure 9). Les ondes électrique et magnétique oscillent en phase et la résultante, l'onde EM obtenue de la somme des vecteurs, oscillent transversalement à un angle de 45° .

Les sources de lumière conventionnelles produisent de la lumière anisotrope, c'est-à-dire non polarisée, dont les photons de lumières voyagent dans l'espace avec des plans d'oscillations divers (VOLLHARDT et SCHORE, 1995). En filtrant la lumière anisotrope à l'aide d'un filtre ou d'un prisme polarisant, on obtient, par écrêtage ou

réfraction, un faisceau de lumière isotrope polarisé linéairement comme celui de la Figure 9.

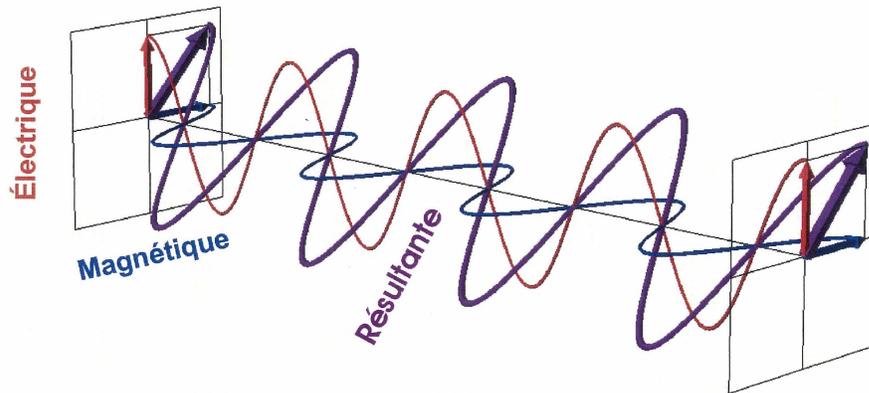


Figure 9 : Rayonnement électromagnétique polarisé linéairement à 45° Les ondes électrique et magnétique voyagent en phase et perpendiculairement l'une de l'autre, et la résultante électromagnétique oscille transversalement à 45°. Les représentations graphiques des diverses formes de rayonnements électromagnétiques et leurs interactions avec la matière, présentées dans ce chapitre, ont été réalisés grâce au gratuitiel EMANIM© Version 1.01, développé par Andras Szilagy, et disponible sur le site internet suivant : (<http://www.enzim.hu/~szia/emanim>)

3.2.2. Réfringence et polarisation circulaire

Le phénomène de « réfringence » (qui produit la réfraction) est observé lorsqu'un milieu exhibe des différences de permittivité à la propagation d'onde selon certains axes. La réfraction (ralentissement d'onde en milieu réfractaire) résulte d'une différence de

vitesse de propagation des ondes dans le milieu. La réfringence s'observe également lorsqu'une onde électromagnétique traverse un cristal présentant une polarisation électrique axiale. Dans ce contexte, la permittivité du milieu pour l'onde électrique diffèrera de celle pour l'onde magnétique et conséquemment, l'onde électrique sera retardée et déphasée de l'onde magnétique. En ajustant adéquatement les dimensions du prisme et/ou la tension électrique appliquée à ce dernier, on parvient à produire la polarisation circulaire, par le retardement sélectif approprié de l'onde électrique d'un rayonnement électromagnétique isotrope (Figure 10).

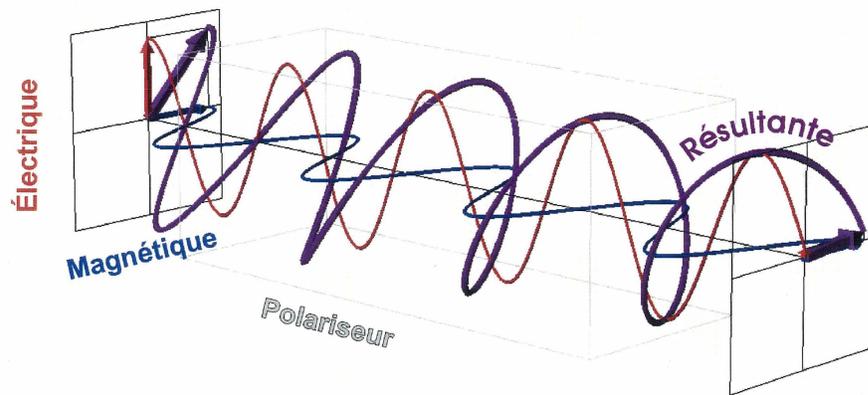


Figure 10 : Rayonnement électromagnétique polarisée circulairement à gauche

Le retardement est créé par la différence de vitesses de propagation des ondes électrique et magnétique lorsqu'elles traversent un cristal polariseur (*e.g.* ADP; Ammonium Phosphate ($\text{NH}_4\text{H}_2\text{PO}_4$) ou KDP; Potassium Phosphate (KH_2PO_4)) électriquement stressé (effet piezo-électrique). Comme seule l'onde électrique est ralentie, celle-ci subira un déphasage par rapport à la composante magnétique. Un déphasage de $\pi/2$ (90°) génèrera une rotation gauche de l'onde électro magnétique, et $3\pi/2$ (270°) résultera en une

rotation à droite. La résultante prend la forme d'une onde circulaire d'amplitude égale à la somme vectorielle (Figure 10).

3.2.3. Combinaison circulaire et polarisation planeaire

Lorsque deux ondes électromagnétiques de mêmes amplitudes mais polarisées circulairement anti-sens l'une par rapport à l'autre, voyagent en phase dans l'espace, la résultante (toujours la somme des vecteurs) prend la forme d'une onde planeaire polarisée linéairement, et dont l'amplitude égale, cette fois, au « double » des ondes individuelles (Figure 11).

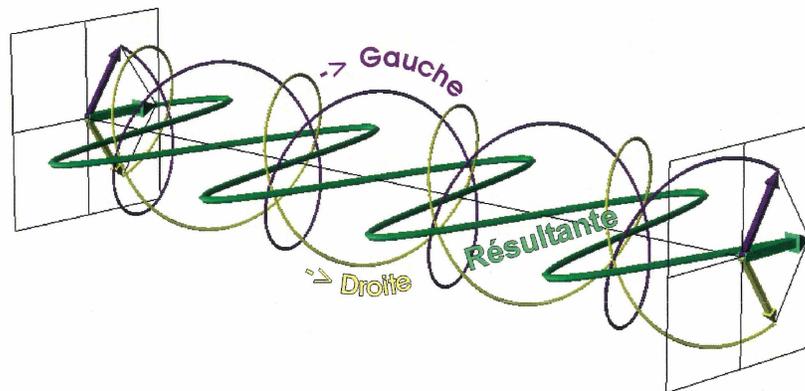


Figure 11 : Recomposition du rayonnement électromagnétique planeaire

Comme nous le verrons dans la prochaine section, bien qu'on puisse produire séparément deux ondes circulaires pour, par la suite, les recombinaer en une seul onde planeaire, l'inverse est tout aussi vrai : on peut extraire d'une onde polarisée planeaire ses deux composantes circulaires.

3.3. Interactions matière : Biréfringence et Dichroïsme circulaire

En traversant la matière, l'onde électromagnétique polarisée plane peut interagir avec celle-ci de différentes façons. Deux interactions retiendront plus particulièrement notre attention, soit : la « biréfringence circulaire » et le « dichroïsme circulaire ».

3.3.1. Biréfringence circulaire

Le processus de « bi » réfringence fait appel au concept de réfringence déjà mentionné précédemment, sauf qu'il implique cette fois-ci une différence entre les vitesses de propagation d'ondes polarisées circulairement à droite et à gauche. Ce processus permet d'expliquer l'activité optique observée par la rotation du plan de polarisation de la lumière, lorsqu'elle traverse une solution de molécule possédant un centre chiral ou asymétrique (*e.g.* C* sucres « D » et « L », C_α des acides aminés et des protéines, énantiomères optiquement actifs, etc.) (VOLLHARDT et SCHORE, 1995).

Lorsqu'un milieu ou une solution présente une différence de permittivité pour les composantes circulaires de l'onde polarisée plane, la différence de vitesse de propagation entraînera le retardement d'une composante par rapport à l'autre. Conséquemment, si l'une des deux composantes n'est pas préférentiellement absorbée, la résultante subira simplement une distorsion de son plan de polarisation (Figure 12). Si le milieu est plus réfringent à la composante gauche, l'angle de polarisation tournera

(pour l'observateur) dans le sens des aiguilles d'une montre, et inversement si la composante droite est plus retardée.

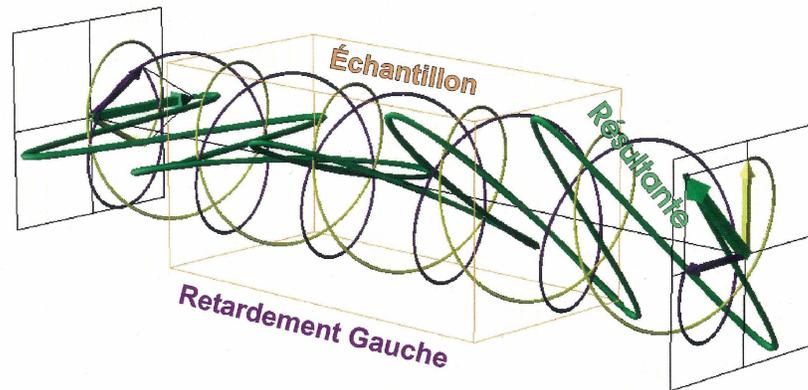


Figure 12 : Biréfringence circulaire de l'onde polarisée gauche (rotation horaire)

Notez la différence des plans de polarisation des lumières incidente et transmise.

3.3.2. Dichroïsme circulaire

Le dichroïsme circulaire s'observe lorsque, en traversant un échantillon, l'une des deux composantes de l'onde polarisée plane est « absorbée » préférentiellement à l'autre. La résultante transmise, somme des vecteurs, recouvre alors une polarisation circulaire contraire à l'onde absorbée, vestige des fondements circulaires de l'onde plane (Figure 13). C'est cette différence d'absorption : $\Delta A = (Abs_g - Abs_d)$ que l'appareil CD mesure et retourne sous la forme d'une ellipticité (θ) positive pour la résultante droite, ou négative pour la résultante gauche.

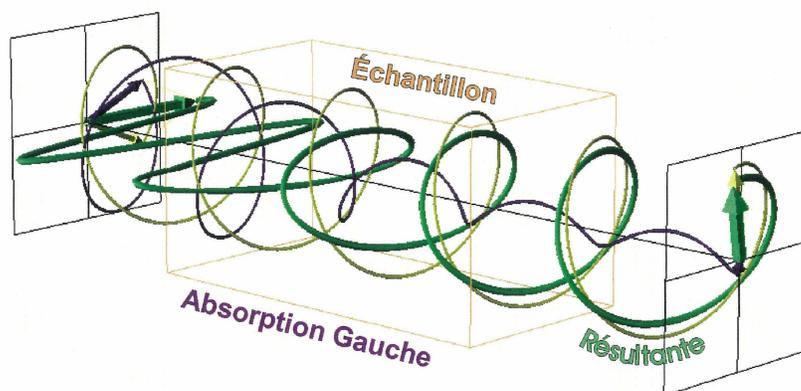


Figure 13 : Dichroïsme circulaire de l'onde polarisée gauche (résultante droite)

Physiquement, l'ellipticité est le résultat d'une différence de coefficient d'extinction moléculaire ($\Delta\varepsilon = (\varepsilon_g - \varepsilon_d)$). Comme la spectroscopie CD fait appel à l'absorption de lumière, la valeur de $\Delta\varepsilon$ est donc influencée par les mêmes paramètres optiques que ceux décrits par la loi de Beer-Lambert ($A = \varepsilon Cl$), soit : la concentration « C » et la longueur du trajet optique « l » (JOHNSON, 1988). La loi se traduit ici comme : $\Delta A = (\varepsilon_g - \varepsilon_d)Cl$, ou simplement $\Delta A = \Delta\varepsilon Cl$, d'où on isole $\Delta\varepsilon = \Delta A / Cl$ avec un résultat exprimé en millidegré d'ellipticité, positive ou négative ($\pm mdeg$).

Précisons qu'un changement de vitesse de propagation (ou d'indice de réfraction) de la lumière polarisée circulairement gauche ou droite, (ou biréfringence) conférée par la chiralité du carbone α et l'activité optique globale de la protéine, n'est pas un phénomène de dichroïsme circulaire. Ce phénomène affecte seulement l'angle α (Figure 14), et relève plutôt de la « dispersion rotatoire optique » (ORD) « *Optical Rotatory Dispersion* », l'ancêtre du dichroïsme circulaire. Bien que la lumière transmise au détecteur puisse être influencée d'une biréfringence circulaire, l'ellipticité résultante du

dichroïsme en est toutefois indépendante. Comme le montre la Figure 14 (un retardement de l'onde polarisée gauche superposé à une absorbance de l'onde polarisée droite), les deux phénomènes peuvent se chevaucher sans que l'un n'affecte l'autre, puisque si la biréfringence affecte l'angle « α » du plan de polarisation, l'ellipticité (θ) elle, est donnée par l'angle « ε » formé par la jonction des points OA'B'.

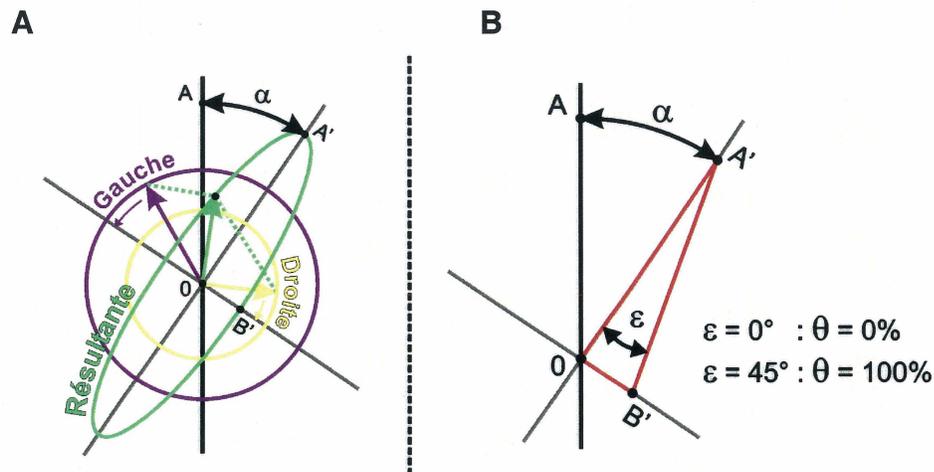


Figure 14 : Dichroïsme circulaire et biréfringence circulaire (Adapté de (BEYCHOK, 1966)) A) La biréfringence affecte l'angle « α » sans altérer l'effet du dichroïsme, B) le dichroïsme affecte l'angle « ε » indépendamment de l'angle « α ».

Ceci permet de comprendre que l'ellipticité reste toujours proportionnelle à la tangente de l'angle ε soit : $\theta \propto \tan(\varepsilon)$. Ainsi, puisque : $\tan(\varepsilon) = \sin(\varepsilon)/\cos(\varepsilon)$ l'ellipticité est nulle lorsque $\varepsilon = 0^\circ$, et maximum lorsque $\varepsilon = 45^\circ$ soit lorsque l'une des deux composantes est totalement absorbée. (Il existe plusieurs différentes façons par lesquelles les appareils modernes mesurent l'absorbance et calculent l'ellipticité (θ), nous référons donc le lecteur à (KELLY et al., 2005) pour plus de précisions.)

3.4. Chromophore amide, dichroïsme circulaire et structures secondaires

La convivialité de la spectroscopie CD pour l'étude des protéines vient de sa surprenante efficacité à discriminer, à même une solution de molécules anisotropes, la nature et les proportions en structures secondaires (Figure 15). D'une manière simple et rapide (comparé à la diffraction de rayon X ou la RMN) le CD permet d'estimer rapidement les quantités de structures secondaires tels l'hélice (α), le feuillet (β) et les structures aléatoires (r) « *random coil* » (JOHNSON, 1988).

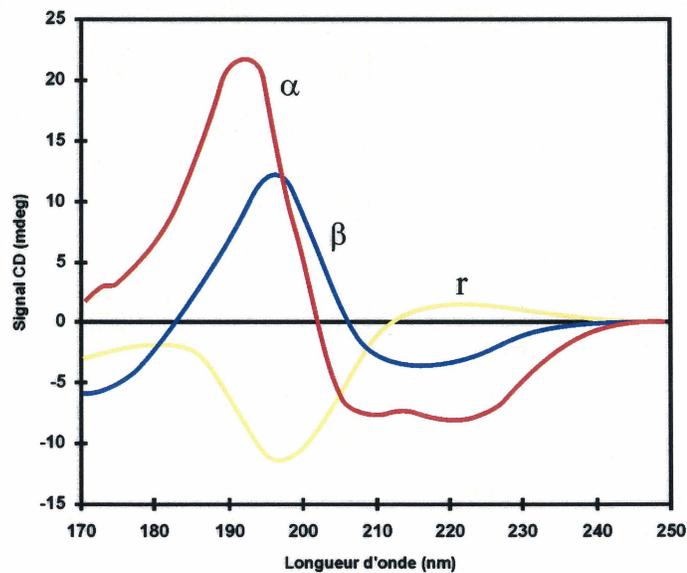


Figure 15 : Spectres de dichroïsme circulaire des structures types α , β et r (Adapté de (GRATZER et COWBURN, 1969).)

Les spectres caractéristiques présentés à la Figure 15 proviennent de mesures effectuées à partir de solutions de mêmes concentrations de l'homopolypeptide poly-L-lysine

{(Lys)_n} qui adopte les 3 différentes conformations en solution aqueuse selon : aléatoire à pH 7.0, hélice α à pH 10.8 et feuillet β à pH 11.1 après chauffage à 52°C et refroidissement.

Le dichroïsme circulaire permet ainsi de corréler directement certains maximums et minimums d'un spectre à des structures secondaires particulières. Comme nous le verrons plus en détails dans les prochains chapitres, il permet grâce aux chromophores amides des protéines, de suivre et analyser des processus thermodynamiques complexes comme : la reconnaissance moléculaire, la dénaturation et le repliement des macromolécules biologiques en solution.

3.4.1. Chromophore amide et dichroïsme circulaire

Les protéines possèdent toutes la propriété d'absorber la lumière aux longueurs d'onde UV avoisinantes de 200 nm. Cette aptitude leur est conférée par l'abondance de chromophores « amides » -CO-NH- engendrés par la liaison peptidique (Figure 16). *(Certaines protéines possèdent de plus, la propriété d'absorber la lumière aux longueurs d'onde avoisinantes de 300 nm, dû à la présence de résidus phénylalanine, tyrosine, tryptophane ou cystine. Dans un but de concision, nous ne traiterons ici que du dichroïsme du chromophore amide. Pour une revue exhaustive voir (KELLY et al., 2005).)*

Les états énergétiques des molécules sont définis par l'arrangement électronique de leurs orbitales atomiques et les modes vibratoires et rotatoires de leurs liaisons chimiques. La lumière peut être absorbée par interaction du vecteur électrique avec un électron. En absorbant un photon d'énergie équivalant à la différence énergétique existant entre deux orbitales ou deux modes vibratoires, les électrons d'une molécule peuvent être excités à un niveau énergétique supérieur et changer d'orbitale atomique ou moléculaire. L'état excité persiste alors quelque temps puis la molécule relaxe vers son état basal de moindre énergie en émettant à nouveau un photon. L'électron excité retourne à son orbital de moindre énergie, tandis que les modes vibratoires et rotatoires de la molécule décroissent. Ce faisant, l'énergie est émise ou transmise à une autre liaison, un autre atome, une autre molécule ou finalement, simplement perdu dans l'environnement (VOLLHARDT et SCHORE, 1995).

Contrairement à l'absorption UV conventionnelle, qui n'implique que le moment de transition électrique des électrons, le phénomène de dichroïsme circulaire, lui, fait intervenir également les moments de transition magnétiques, présents ou engendrés par le déplacement d'électrons lors d'une transition électronique (BEYCHOK, 1966). Comme tout déplacement de charge engendre un moment magnétique perpendiculaire au déplacement, et que tout champ magnétique perturbe invariablement le déplacement de charges électriques, le dichroïsme résulte de ces subtiles préférences dans l'achèvement de certaines transitions électroniques. Notons donc déjà que l'intensité d'une bande d'absorbance sur un spectre UV conventionnel ne dicte en rien l'intensité du dichroïsme qui peut lui être associé. Ainsi, même une très faible absorbance peut

s'avérée hautement préférentielle et, ce faisant, engendrer une forte ellipticité, tout comme à l'opposé, une très forte absorbance UV peut ne produire aucune ellipticité.

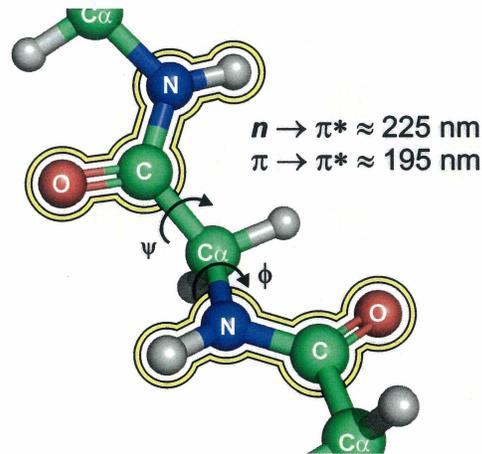


Figure 16 : Chromophores amides et principales transitions : $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$

L'environnement électromagnétique des liaisons peptidiques (CONH) d'une protéine dépendra : des valeurs des angles dièdres « ϕ » et « ψ » (Figure16), des structures secondaires dans lesquelles elles se trouvent (et de la taille de celles-ci), de même que de la présence de chaînes latérales (GREENFIELD et FASMAN, 1969). Il résultera, de cette variabilité des environnements électromagnétiques des amides, une polarisation électrique/magnétique des transitions électroniques et une absorbance préférentielle aux différentes longueurs d'onde polarisées circulairement droite ou gauche.

Les amides exhibent principalement deux transitions électroniques observables par dichroïsme circulaire dans l'UV lointain entre 180 et 250 nm. La première relève d'une transition $n \rightarrow \pi^*$, impliquant une paire d'électrons libre (n_1) promues à une orbitale

moléculaire anti-liante (π^-) et s'observe généralement autour de 210 à 230 nm (BEYCHOK, 1966). La bande d'absorption résultante possède un faible coefficient d'extinction puisque cette transition est considérée comme électriquement défendue, mais permise magnétiquement.

La deuxième transition observable pour les amides, se situe aux environs de 195 nm. Elle implique la promotion d'électrons « π » d'une orbitale moléculaire liante (π^0) vers une orbitale moléculaire anti-liante (π^-) soit une transition $\pi \rightarrow \pi^*$ (BEYCHOK, 1966). La polarisabilité de cette transition, favorisée par la résonance due à la proximité d'amides voisines, implique un « couplage d'excitons » et un dédoublement de la bande d'absorption. *(L'explication quantique sous-jacente de cette nature particulière de la transition $\pi \rightarrow \pi^*$ dépassant largement le cadre de ce travail, nous conseillons au lecteur de référer à (OAKLEY et al., 2006a et 2006b)) pour une discussion récente et plus approfondie du sujet.)*

3.4.2. Dichroïsme circulaire : hélice α

Le couplage d'excitons au sein de la transition $\pi \rightarrow \pi^*$ des hélices α doit être perçue comme le résultat d'une polarisation des amides, renforcée par la structure globale et la périodicité des liaisons peptidiques dans l'espace. Le couplage bénéficiant de la périodicité des chromophores amides, il naît de deux moments de transition induits : l'un parallèle à l'axe de l'hélice ($\pi \rightarrow \pi^* \parallel$) et l'autre perpendiculaire à l'hélice ($\pi \rightarrow \pi^* \perp$)

⊥). Pour les hélices (Figure 17), on observera donc principalement deux bandes pour la transition $\pi \rightarrow \pi^*$, soit un très fort maximum à 190 – 195 nm ($\pi \rightarrow \pi^*$ ⊥), et un fort minimum très diagnostique à 208 nm ($\pi \rightarrow \pi^*$ ∥). L'hélice compte de plus avec la transition $n \rightarrow \pi^*$, une troisième bande d'absorption centré à 222 nm, particulièrement susceptible à la fermeté de la structure et sensible à son organisation tertiaire sous forme de « *coil-coil* ». Un ratio $^{208\text{nm}}/_{222\text{nm}} < 1$ est souvent synonyme d'une structure tertiaire achevée tandis qu'un ratio > 1 implique habituellement une proportion d'hélice monomérique (GRATZER et COWBURN, 1969).

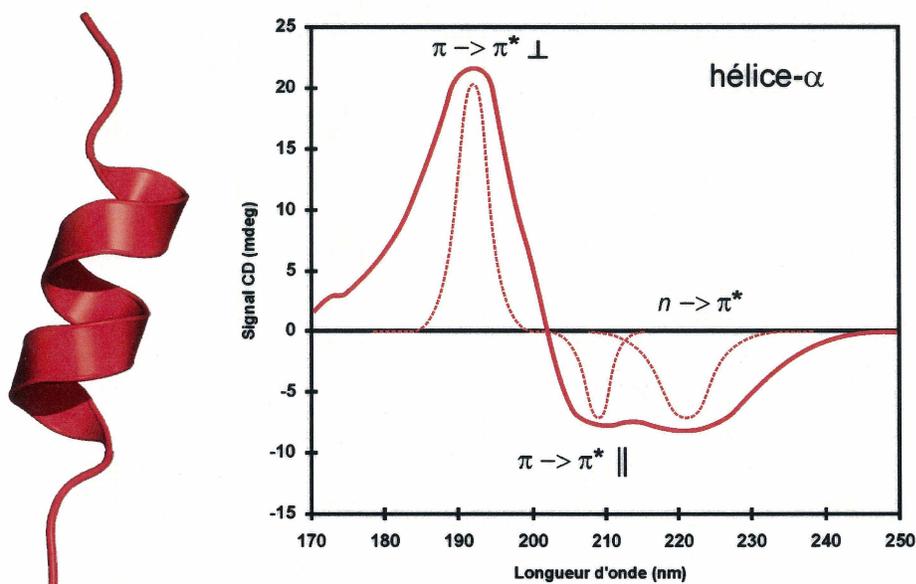


Figure 17 : Hélice et spectre CD type de l'hélice α

3.4.3. Dichroïsme circulaire : feuillet β

Le spectre CD du feuillet β antiparallèle (Figure 18) présente lui aussi des évidences d'un couplage d'excitons au sein de la transition $\pi \rightarrow \pi^*$ (largeur de la raie), quoi qu'il soit différent de celui de l'hélice α et moins bien résolu. Le spectre de feuillet β se dissèque en un unique maximum de la transition $\pi \rightarrow \pi^*$ à 195 nm, mais auquel s'ajoute un minimum de magnitude moyenne et de profil très caractéristique, venant de la transition $n \rightarrow \pi^*$, centrée autour de 215 – 218 nm (GRATZER et COWBURN, 1969).

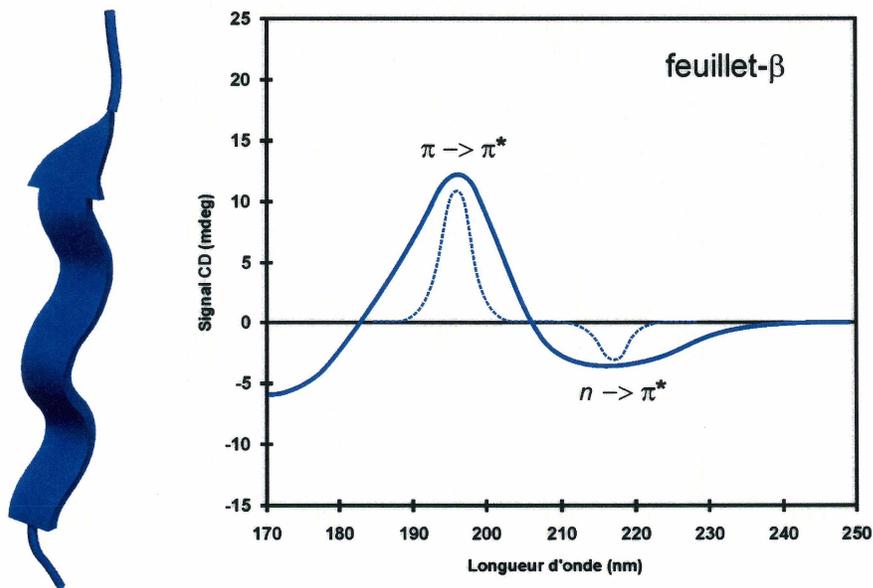


Figure 18 : Brin β et spectre CD type du feuillet β antiparallèle

3.4.4. Dichroïsme circulaire : structures aléatoires r « *random coil* »

Finalement, un polypeptide ou une protéine dans son état dénaturé présente un spectre similaire à celui d'un amide simple, à l'exception qu'il possède une magnitude supérieure (Figure 19). Le spectre décrit pour la transition $n \rightarrow \pi^*$ un léger maximum centré autour de 225 nm, et un profond minimum caractéristique à 195 nm pour l'unique transition $\pi \rightarrow \pi^*$ (GRATZER et COWBURN, 1969).

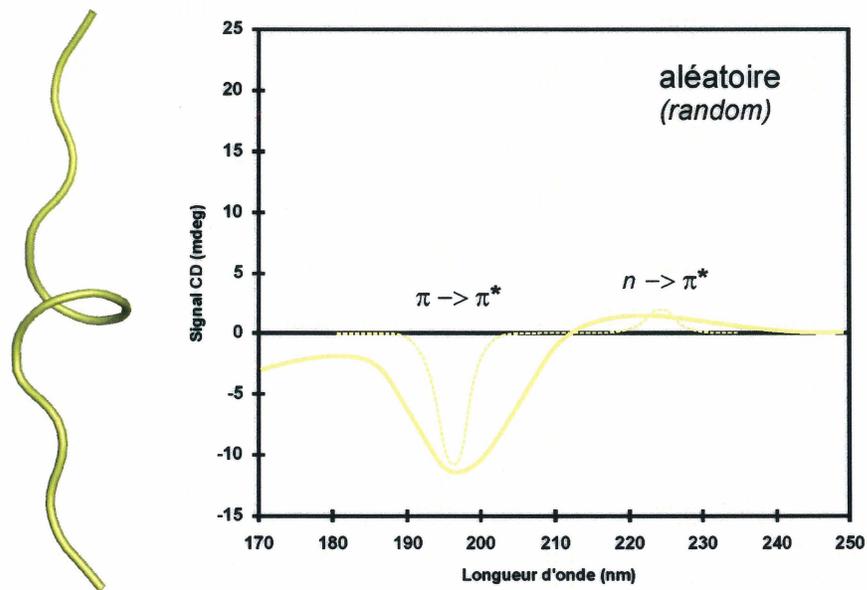


Figure 19 : Segment aléatoire et spectre CD type des formes aléatoires

3.5. CD vs analyse structurale et thermodynamique

Comme mentionné précédemment, le dichroïsme circulaire représente un outil de choix pour investiguer les changements de conformation des protéines et complexes protéiques ou pour suivre l'évolution des changements d'état thermodynamique en solution. Dans le premier cas, le recours à des balayages de la plage de fréquence 190 à 250 nm (voir 190 à 350 nm pour les résidus cys, phe, trp et tyr) est primordiale pour bien suivre les déplacements des minimums et maximums caractéristiques en cours d'expérience. L'enregistrement de spectres sur de larges plages de longueur d'onde présente par contre les défauts d'être lente et de ne fournir qu'un nombre restreint de mesures par longueur d'onde échantillonnées (faible ratio signal/bruit). Ainsi, lors d'analyse de changement d'état (*e.g.* dénaturation thermique, renaturation, reconnaissance moléculaire ou liaison en directe), on recourt plus spontanément au suivi d'une unique longueur d'onde, échantillonnée plusieurs fois à la minute et de façon continue tout au long de l'expérience. Il incombe donc de choisir avec discernement la longueur d'onde qui sera échantillonnée.

Théoriquement, la meilleure plage d'observation des transitions de structure secondaire, ce situe aux environs de 190 - 200 nm étant donné les importantes différences de signaux escomptables pour les transitions entre l'hélice α et les structure aléatoire, le feuillet β et les structures aléatoire ou finalement l'hélice α et le feuillet β (Figure 20). La pratique incombe par contre deux zones grisées, soit : entre 170 et 205 nm, et entre 240 et 250 nm. La première zone (170 à 205 nm) est malheureusement à proscrire

puisque'il s'agit d'une région particulièrement sensible aux bruits résultant de l'absorbance par les sels et impuretés des solutions tampons, de même qu'à l'absorbance par l'oxygène ambiant ou dissout. Il s'agit également d'une plage particulièrement difficile à couvrir pour les lampes conventionnelles au Xénon. La deuxième zone (240 – 250 nm) pour sa part ne présente pas d'intensité diagnostique (hormis d'éventuelles contributions de résidus cystines), et présente donc un défaut de ratio signal/bruit par simple manque de signal. Notons que des progrès pour les plages inférieures à 200 nm ont été récemment réalisés par le recours à l'utilisation de lumière synchrotronique dans des expériences de spectroscopie STCD (SynchroTron Circular Dichroism) (COWIESON *et al.*, 2008).

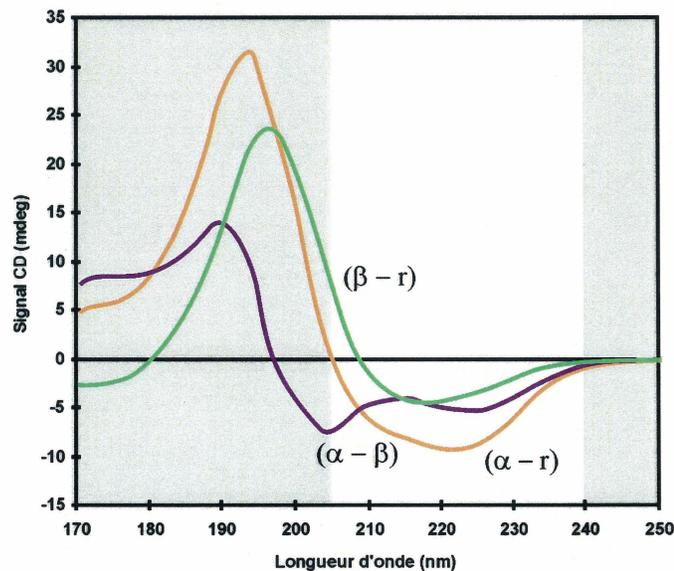


Figure 20 : Différence de spectres types : $\alpha - r$, $\beta - r$ et $\alpha - \beta$

Conséquemment il ne reste, pour la plupart des spectroscopistes, qu'une mince plage de 35 nm se situant entre 205 et 240 nm (Figure 20). Fort heureusement on y distingue,

pour chacune des altérations types de structure secondaire, un minimum caractéristique provenant de la transition $n \rightarrow \pi^*$. On distingue ainsi : pour la transition ($\alpha \leftrightarrow r$) un minimum à 222 nm, pour la transition ($\beta \leftrightarrow r$) un minimum à 218 nm et finalement pour la transition ($\alpha \leftrightarrow \beta$) deux minimums, soit un premier à 205 nm et un second à 227 nm. Bien entendu, les protéines et les complexes présentent souvent un assemblage varié de différentes structures secondaires, il incombe donc d'abord de caractériser le sujet par une série de balayage afin d'identifier les meilleures opportunités pour suivre le plus précisément possible les types de transition particulières qui s'appliquent.

Pour résumer ce chapitre, on peut dire que le dichroïsme circulaire des protéines aux longueurs d'onde comprises entre 180 et 250 nm provient d'une préférence de certaines transitions électroniques du chromophore amide, à l'absorption de rayonnement électromagnétique polarisée gauche et/ou droit. Cette préférence est due aux moments électriques et magnétiques des transitions électroniques $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$ des groupements amides, et donc à l'organisation électromagnétique globale des liens peptidiques au cœur de la protéine. L'agencement des chromophores étant elle-même conférée par les valeurs des angles ϕ et ψ , qui sont assujettis aux résidus substituant les carbones α et aux structures secondaires, tertiaires et quaternaires, l'ensemble de la structure protéique participe se faisant à l'émergence du dichroïsme circulaire. Notons en terminant que les spectres CD obtenus avec des protéines naturelles aux acides aminés « L » sont exactement inversés (\pm) pour leurs homonymes synthétisés tout « D » (image miroir) (PRITSKER *et al.*, 1998), ce qui démontre bien l'influence de

l'asymétrie du carbone α sur l'organisation structurale, et ce faisant, sur la polarisation du chromophore amide et ses répercussions sur les signaux CD.

4. THERMODYNAMIQUE-COMPUTATIONNELLE

4.1.L'analyse thermodynamique du système de c-Myc & Max

On peut, à l'aide de simulations de résultats expérimentaux, attribuer les paramètres thermodynamiques relatifs à la dénaturation d'un système de macromolécules en solution (VELAZQUEZ-CAMPOY *et al.*, 2004). On doit pour ce faire enregistrer l'évolution d'un signal diagnostique des états macroscopiques du système (natif ou dénaturé), au cours d'une expérience de dénaturation thermique. Le signal, habituellement de nature spectroscopique (*e.g.* absorbance et diffusion de lumière, dichroïsme circulaire, fluorescence, infra-rouge), doit essentiellement permettre l'observation des changements d'états macroscopiques en fonction de la température. On peut par exemple, à l'aide du dichroïsme circulaire, corrélérer les variations de teneur en structures secondaires d'une protéine avec ses états macroscopiques « natif » et « dénaturé ». En définissant d'abord l'intensité des signaux spectroscopiques propres à chacun des états natif et dénaturé, on peut par la suite simuler le signal correspondant à la transition entre ces deux états grâce aux populations générées *via* l'équation de Gibbs-Helmholtz (éq. 1.0)(PRIVALOV, 1996). La simulation par régression (minimisation de la différence entre une courbe simulée et des valeurs expérimentales) permet de trouver par ajustement progressif *via* l'équation 1.0, les paramètres thermodynamiques comme : la température de transition ($T_{dén}$), l'enthalpie de Van't Hoff associée à la dénaturation ($\Delta H_{dén}(T_{dén})$), la variation de la capacité calorifique ($\Delta C_{p_{dén}}$), la variation d'énergie libre ($\Delta G_{dén}$) et ultimement les constantes de dénaturations ($K_{dén}$).

éq. 1.0
$$\Delta G_{dén}^{\circ}(T) = \Delta H_{dén}^{\circ}(T_{dén}) - T \Delta S_{dén}^{\circ}(T_{dén}) + \Delta C p_{dén}^{\circ} \left(T - T_{dén} - T \ln \left(\frac{T}{T_{dén}} \right) \right)$$

Cette méthode s'est avérée très efficace pour caractériser de nombreux équilibres comme ceux de protéines globulaires (GAGNON *et al.*, 2000), de multimères (MILVEV *et al.*, 2003a et 2003b), et même de systèmes présentant des états intermédiaires de dénaturation (NAUD *et al.*, 2003). Elle présente par contre des limitations intrinsèques lorsque vient le moment de simuler des systèmes plus complexes, comportant des équilibres enchevêtrés. Les limitations proviennent de l'incapacité de la technique à déterminer une solution unique pour l'ensemble des variables à l'aide d'une seule et même simulation (*trop de variables pour trop peu d'équations*). Il devient alors impossible de discriminer objectivement les signaux propres aux nombreuses espèces possibles (*e.g. A, B, AA, AB, BB*). (*Notons que l'on pourrait théoriquement envisager ici le recours à des technologies de « transfert non-radiatif d'énergie » comme le BRET et le FRET (« Bioluminescence » et « Förster Resonance Energy Transfer ») pour parvenir à discriminer toutes les espèces d'une solution en une seule simulation. Cela dit, ces techniques présentent elles aussi leurs lots d'imperfections et de complications.*)

Pour résoudre de tels systèmes d'équations à multiples variables, on doit nécessairement déterminer au préalable certaines inconnues, puis résoudre par la suite les variables restantes. Ainsi, les analyses par simulation des systèmes isolés ($AA \rightleftharpoons A + A$) de Max'SH homodimérique (Figure 21) :

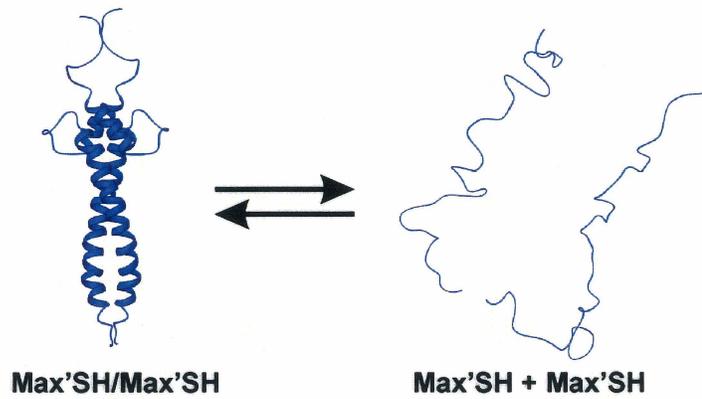


Figure 21 : Équilibre de Max'SH

ou de c-Myc'SH (Figure 22) ($BB \Leftrightarrow B + B$), ne présentent aucun problème d'inconsistance :

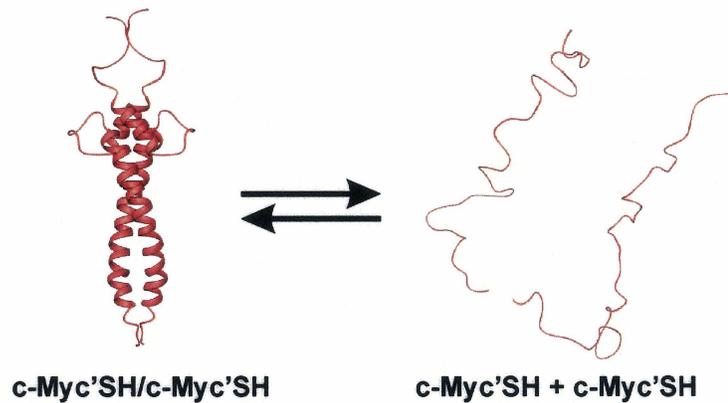


Figure 22 : Équilibre de c-Myc'SH

Mais il en va autrement dans le cas du système global où les homodimères ne sont plus isolés mais plutôt mis en présence l'un de l'autre de manière à pouvoir générer un nouveau complexe hétérodimère c-Myc'SH/Max'SH (Figure 23). Les équations deviennent alors « indéterminées » puisque le modèle ne peut soutenir le fait que plusieurs complexes (*en l'occurrence ici trois : AA, BB, AB et non seulement un seul :*

AB), constituent conjointement l'état natif global. La dynamique de dissociation des complexes homodimériques n'est plus *a priori* connue comme dans le cas des systèmes isolés, puisque les proportions de monomères et d'homodimères sont maintenant soumises à une nouvelle contrainte indéterminée, soit : « la possibilité de former des hétérodimères ». La dissociation d'un homodimère Max'SH/Max'SH n'engendrera plus seulement deux monomères de Max'SH libres, mais également une population d'hétérodimère c-Myc'SH/Max'SH asservie par la concentration de c-Myc'SH libre.

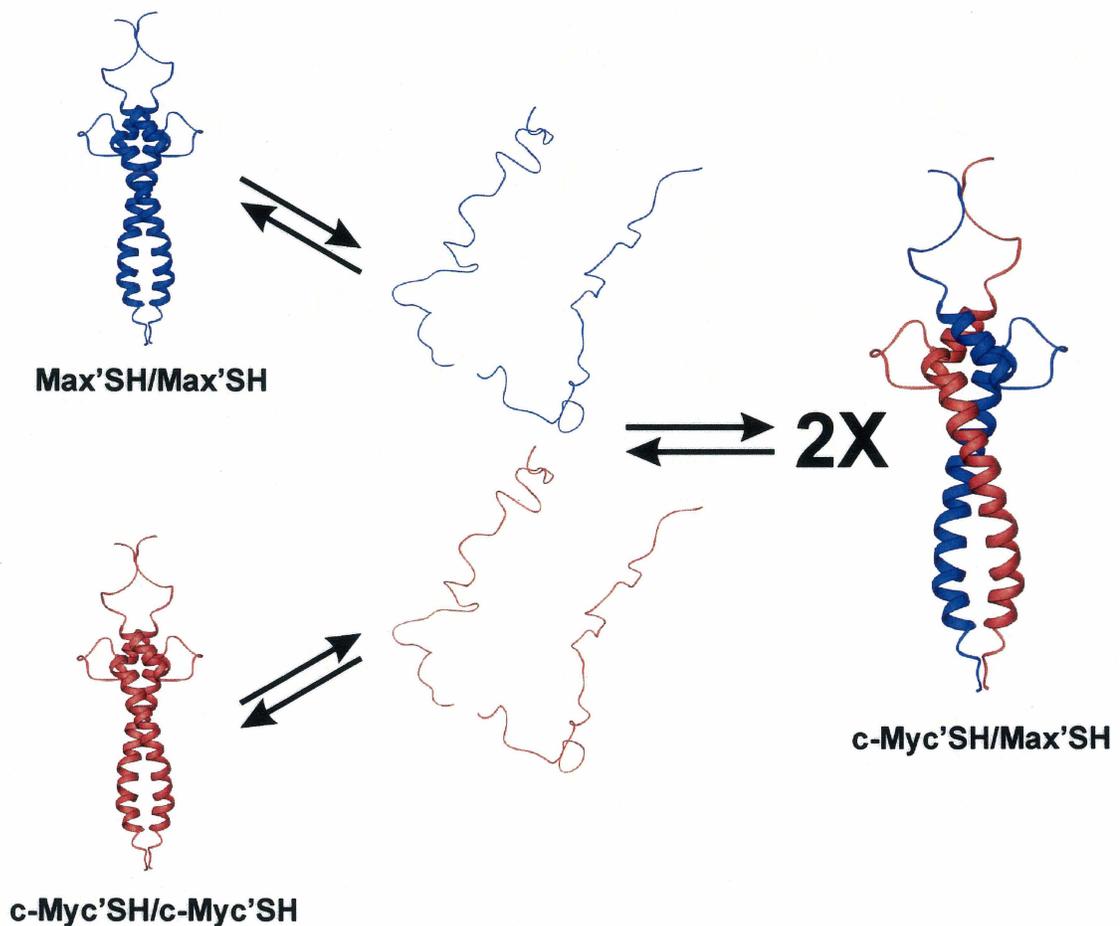


Figure 23 : Équilibre global intriqué de c-Myc'SH et Max'SH

De la même manière, la dissociation de c-Myc'SH/c-Myc'SH est, elle aussi, encouragée par la présence de Max'SH libre en solution. Les constantes de dissociations (K_{DS}) mesurées pour les systèmes isolés ne permettent donc plus de calculer les populations exactes des différentes espèces. Les K_{DS} « réels » des homodimères ($K_{D,Max'SH/Max'SH}$ et $K_{D,c-Myc'SH/c-Myc'SH}$) sont affectés d'une dévaluation (*c-à-d*: une augmentation de la valeur numérique du K_D) proportionnelle à la stabilité du nouvel hétérodimère et de son K_D . On doit conséquemment définir de nouveaux K_{DS} « apparents » (K_{D}^*) qui tiendront compte des compétitions engendrées dans l'équilibre global du mélange.

4.2. Simulation *via* Gibbs-Helmholtz et Tableau CD_F

Une solution de protéine est un système thermodynamique en équilibre, où la structure des protéines est sensible aux variations d'énergie. Comme aux conditions normales l'énergie d'une solution n'est jamais parfaitement répartie, les protéines doivent être vues comme un mélange plus ou moins hétérogène de structures « différentes » en équilibre dynamique (Figure 24).



Figure 24 : Équilibres microscopiques de l'état natif d'un b-HLH-LZ superposition de 6 structures RMN du b-HLH-LZ de MaxVL (SAUVÉ *et al.*, 2004).

En définissant des ensembles statistiques « natif » et « dénaturé » pour différencier les conformations structurales limites qu'une protéine peut adopter en solution, on peut alors poser une fonction de répartition (Q) pour le système et définir des populations propres à chacun des états. Il devient donc possible de caractériser la stabilité thermodynamique de la protéine en simulant l'énergie nécessaire pour engendrer sa transition de l'état natif vers l'état dénaturé (Figure 25).

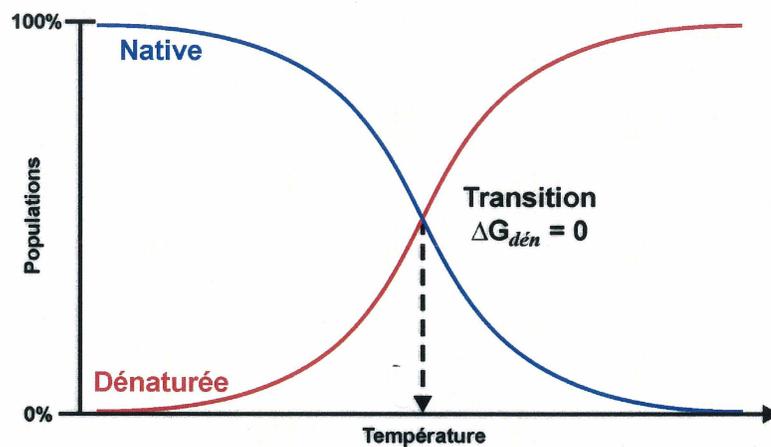


Figure 25 : Transition des populations d'un système à deux états

Pour une protéine « monomérique » (voir *Annexe II-III, V-VI pour les systèmes multimériques dissociatifs*), on pose d'abord la fonction de répartition (Q) (éq. 2.0) d'un système à « deux états » (Natif (N) et Dénaturé (U)) :

éq. 2.0
$$Q = N + U$$

On peut alors poser que la population (ou probabilité) de protéine (P) dans l'état dénaturé (U) soit (P_U) est égale à l'état dénaturé (U) divisé par la fonction de répartition du système, soit la somme de tous ses états possibles (éq. 2.1 (a)) :

éq. 2.1 (a)
$$P_U = \frac{U}{Q} = \frac{U}{(N + U)}$$

On peut alors exprimer P_U comme une fonction de concentration de protéine :

éq. 2.1 (b)
$$P_U = \frac{[U]}{([N] + [U])}$$

Connaissant l'expression de la constante de dénaturation (K_U), on peut alors en isoler la valeur de la concentration de protéine dénaturée ($[U]$) comme suit :

éq. 2.2 (a & b)
$$K_U = \frac{[U]}{[N]} \quad \text{donc:} \quad [U] = K_U [N]$$

Et ce faisant, remplacer la valeur de $[U]$ dans l'expression de P_U (éq. 2.1 (b)) :

éq. 2.3 (a)
$$P_U = \frac{K_U [N]}{([N] + K_U [N])}$$

Puis finalement, on peut réduire l'équation à une simple fonction de K_U :

éq. 2.3 (b & c)
$$P_U = \frac{K_U [N]}{[N](1 + K_U)} \quad \text{donc:} \quad P_U = \frac{K_U}{(1 + K_U)}$$

Comme on sait que K peut être exprimé comme une fonction de l'énergie libre du système (ΔG°), selon la loi de Stephan Boltzmann (éq. 2.4) :

$$\text{éq. 2.4} \quad K = e^{\left(\frac{-\Delta G^\circ}{RT}\right)}$$

On obtient que la population (probabilité) de protéine dénaturée P_U (éq. 2.3 (c)) est fonction de l'énergie libre du système (éq. 2.5) :

$$\text{éq. 2.5} \quad P_U = \frac{e^{\left(\frac{-\Delta G_U^\circ}{RT}\right)}}{\left(1 + e^{\left(\frac{-\Delta G_U^\circ}{RT}\right)}\right)}$$

Et puisque :

$$\text{éq. 2.6 (a \& b)} \quad P_N + P_U = 1 \quad \text{ou encore :} \quad P_N = 1 - P_U$$

Donc finalement :

$$\text{éq. 2.7} \quad P_N = 1 - \frac{e^{\left(\frac{-\Delta G_U^\circ}{RT}\right)}}{\left(1 + e^{\left(\frac{-\Delta G_U^\circ}{RT}\right)}\right)}$$

Ainsi, l'ensemble des états macroscopiques que peut explorer une protéine en solution à une température donnée peut être défini comme une distribution statistique (P_N vs P_U) déterminée par « G ». Les protéines d'une solution ont donc, à tout instant, la possibilité d'explorer des variations de leur structure, dans la mesure où celles-ci ne requièrent pas

davantage d'énergie qu'en possède localement le système. Celles-ci n'exploreront donc toujours qu'une gamme de structures plus ou moins restreinte en fonction de la valeur de G.

Comme l'énergie libre d'une solution est dépendante de sa température (T) selon la loi de Gibbs (éq. 2.8), on comprend qu'il est possible, en variant la température d'une solution, d'influer sur la distribution statistique des conformations qu'exploreront les protéines individuellement.

éq. 2.8
$$G^\circ(T) = H^\circ(T) - TS^\circ(T)$$

En réalisant une augmentation lente et progressive de la température d'une solution (*e.g.* 1°C/minute), on peut graduellement élever l'énergie libre totale du système (ΔG°) tout en assumant une variation d'énergie libre ($\Delta\Delta G^\circ$) à peu près nulle (condition de réversibilité thermodynamique). Si le système est réversible, il devient alors possible d'effectuer plusieurs mesures en série tout en considérant celui-ci à l'équilibre thermodynamique. Une expérience de la sorte est appelée : « dénaturation thermique à l'équilibre thermodynamique ». (*Pour alléger la suite du texte, l'exposant « ° » signifiant l'état « standard » normalisé (e.g. ΔG°) sera volontairement omis.*)

Au cours d'une dénaturation thermique, les protéines subissent généralement des altérations importantes de leurs structures secondaires, tertiaires et quaternaires. On parle ici d'un changement d'état thermodynamique.

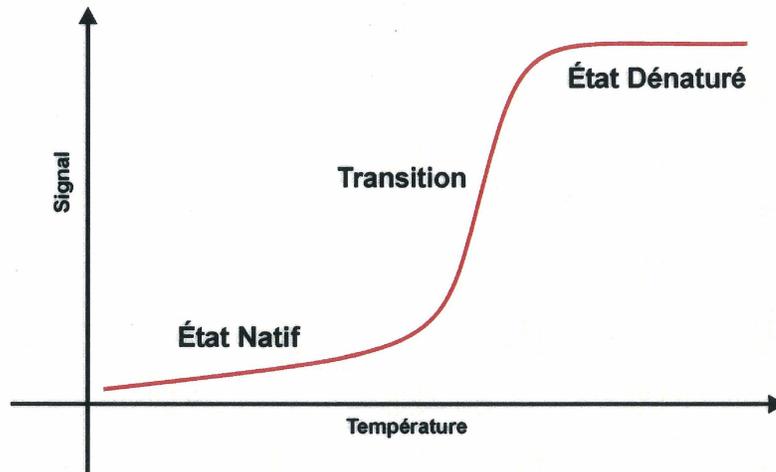


Figure 26 : Dénaturation thermique à l'équilibre thermodynamique

Il est possible, en simulant la progression d'un signal diagnostique de ce changement d'état (tel le dichroïsme circulaire) de corréler, grâce à l'équation de Gibbs-Helmholtz (éq. 1.0), la variation des populations de protéine natives et dénaturées en fonction de la température. Les paramètres thermodynamiques régissant les équations, et donc simulant l'énergie du système, sont alors raffinés progressivement de manière à ce que les populations s'ajustent le plus adéquatement possible aux valeurs empiriques obtenues aux différentes températures. Les populations sont obtenues à partir des constantes de dénaturation (K_U et/ou K_D pour les systèmes dissociatifs), elles-mêmes calculées à l'aide de l'équation de Gibbs-Helmholtz. On peut ainsi déterminer, pour n'importe quelle température, la constante de dénaturation propre au système.

4.2.1. Influence des intensités spectroscopiques extensives

Pour simuler les signaux empiriques en utilisant l'équation de Gibbs-Helmholtz, on doit préalablement attribuer des valeurs d'intensité spectroscopique extensives aux états : natif et dénaturé. Ces valeurs peuvent prendre des formes diverses en fonction du type de mesure que l'on désire simuler. Elles peuvent être exprimées sous la forme d'intensité de signal absolue par Kelvin ($\theta \cdot K^{-1}$) dans le cas où les équations produisent des valeurs de population de 0 à 100%, ou d'intensité par mole et par Kelvin ($\theta \cdot mole^{-1} \cdot K^{-1}$) si les équations sont élaborées de manière à produire des concentrations d'espèces.

La détermination des valeurs spectroscopiques s'effectue en ajustant le plus fidèlement possible une droite à chacune des régions limites de la dénaturation thermique. Il est donc primordial de pouvoir distinguer clairement, sur la dénaturation thermique, les régions propres aux deux états « natifs » et « dénaturés ». (Figure 27 A)

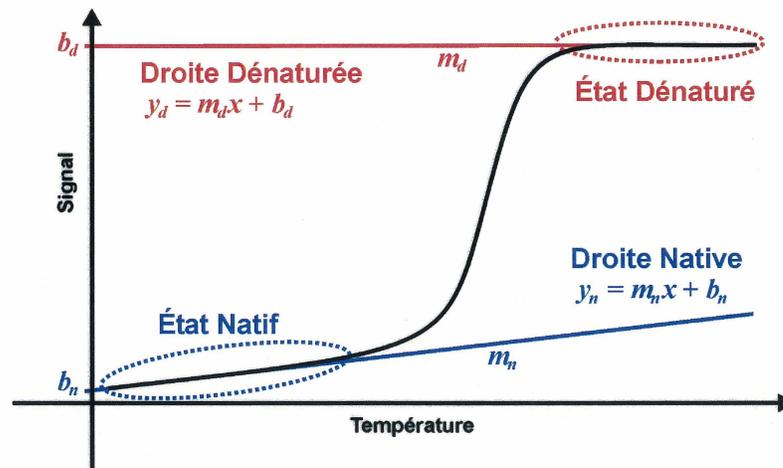


Figure 27 A) : Détermination des droites native et dénaturée (modèle à 2 états)

Dans la mesure où l'un ou l'autre des états est ambigu, une approximation devra être faite en tenant compte du maximum de connaissances disponibles sur le système à l'étude. Les choix des pentes et des ordonnées des droites représentant les états limites d'un système restent des décisions subjectives mais qui doivent être défendables. Une fois l'ajustement des droites complété, on extrait de celles-ci les valeurs des pentes « m » et des ordonnées à l'origine « b ». On résout alors les équations « $y = mx + b$ » qui serviront à générer les signaux synthétiques propres aux populations « natives » et « dénaturées », qui seront finalement additionnés pour générer le « signal simulé ».

Dans certains cas plus complexes, il peut s'avérer nécessaire de choisir un modèle à « trois états », impliquant le passage par un état « intermédiaire » de repliement. Cependant, on doit distinguer suffisamment l'état intermédiaire pour lui attribuer une droite d'intensité spectroscopique propre (Figure 27 B).

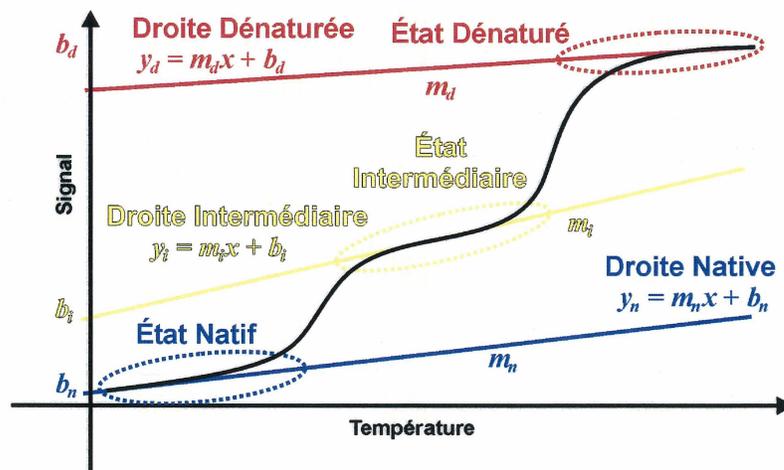
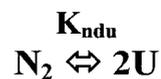


Figure 27 B) : Détermination de la droite intermédiaire (modèle à 3 états)

Dans le cas de l'homodimère Max/Max, le système se simule sans nécessiter un tel état supplémentaire. Les altérations mineures progressives de l'état dimérique (à haute température) sont, tout comme les contributions d'éventuelles structures résiduelles de l'état dissocié (à basse température), incluses dans les pentes des droites respectives des deux principaux états natif et dénaturé. L'homodimère MaxVL/MaxVL (NAUD *et al.*, 2003) lui par contre, nécessite le recours à un modèle en trois états (Figure 27 B), puisque les domaines Leucine Zipper sont artificiellement stabilisés par deux mutations et persistent en dimère malgré la dénaturation des HLH.

4.2.2. Détermination des populations via les équations de Gibbs-Helmholtz

Pour parvenir à exprimer les populations des différentes espèces d'un système, on doit définir les équations propres à celui-ci. Par exemple, en posant K_{ndu} (ndu : « *native-dissociated-unfolding* ») pour un système dissociatif comme celui de Max'SH, et où le complexe d'homodimère Max'SH/Max'SH (N_2) se dissocie en deux monomères Max'SH dénaturés (2U) :



On pose d'abord que :

éq. 3.0
$$K_{ndu} = \frac{[U][U]}{[N_2]} = \frac{[U]^2}{[N_2]}$$

Sachant que la concentration de protéine totale ($[T]$) est donnée par l'équation 3.1 (a) :

éq. 3.1 (a) $[T] = 2[N_2] + [U]$

On peut poser que la concentration de complexe natif $[N_2]$ est égale à :

éq. 3.1 (b) $[N_2] = \frac{[T] - [U]}{2}$

La constante de dénaturation/dissociation fonction de la concentration de protéine totale (éq. 3.0) peut donc être réécrite comme suit :

éq. 3.2 (a & b) $K_{ndu} = \frac{[U]^2}{\left(\frac{[T] - [U]}{2}\right)}$ ou $0 = 2[U]^2 + K_{ndu}[U] - K_{ndu}[T]$

D'où l'on peut extraire la concentration de protéine dénaturée $[U]$ via la racine quadratique (+) :

éq. 3.3 $[U] = \frac{-K_{ndu} + \sqrt{K_{ndu}^2 - 4 \cdot 2 \cdot -K_{ndu}[T]}}{4}$

En posant que la population de protéine dénaturée (P_U) est égale à la concentration de protéine dénaturée ($[U]$) divisé par la concentration de protéine totale ($[T]$) (éq. 3.4) :

éq. 3.4 $P_U = \frac{[U]}{[T]}$

On peut alors exposer P_U comme étant égale à :

$$\text{éq. 3.5 (a)} \quad P_U = \frac{-K_{\text{ndu}} + \sqrt{K_{\text{ndu}}^2 + 8K_{\text{ndu}}[T]}}{4[T]}$$

Et son corollaire donné en fonction de la température (T) de la solution :

$$\text{éq. 3.5 (b)} \quad P_U(T) = \frac{-K_{\text{ndu}}(T) + \sqrt{K_{\text{ndu}}(T)^2 + 8K_{\text{ndu}}(T)[T]}}{4[T]}$$

On obtient alors instantanément par soustraction, la population de protéine native (P_N) en fonction de la température :

$$\text{éq. 3.6} \quad P_N(T) = 1 - P_U(T)$$

On doit à présent simplement calculer $K_{\text{ndu}}(T)$ pour obtenir automatiquement les populations natives et dénaturées à n'importe qu'elles températures. C'est ici que l'on a recours à l'équations de Gibbs-Helmholtz (éq. 1.0) pour calculer, en fonction de la température, la variation d'énergie libre associée à la dénaturation ($\Delta G_{\text{ndu}}(T)$) et avoir accès à $K_{\text{ndu}}(T)$ *via* la loi de Boltzmann :

$$\text{éq. 3.7} \quad K_{\text{ndu}}(T) = e^{\left(\frac{-\Delta G_{\text{ndu}}(T)}{RT}\right)}$$

On pose donc d'abord l'équation de Gibbs où ΔG égale la variation d'énergie libre du système, ΔH est la variation d'enthalpie, T est la température et ΔS est la variation d'entropie :

$$\text{éq. 4.0} \quad \Delta G = \Delta H - T\Delta S$$

On exprime ensuite l'équation en tant que Gibbs-Helmholtz dépendante de la température et où l'indice « ndu » (« *native-dissociated-unfolding* ») symbolise les paramètres propres au changement d'état tel que : ΔG_{ndu} est la valeur de l'énergie libre, T est la température pour laquelle on souhaite mesurer ΔG_{ndu} , T_{ndu} est la température de transition, $\Delta H_{\text{ndu}}(T_{\text{ndu}})$ est l'enthalpie standard de dénaturation de Van't Hoff à la température de transition T_{ndu} , $\Delta S_{\text{ndu}}(T_{\text{ndu}})$ est la variation d'entropie, R est la constante des gaz parfaits et $\Delta C_{p_{\text{ndu}}}$ est la variation de la capacité calorifique à pression constante :

$$\text{éq. 4.1} \quad \Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - T\Delta S_{\text{ndu}}(T_{\text{ndu}}) + \Delta C_{p_{\text{ndu}}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

Sachant qu'à l'équilibre thermodynamique :

$$\text{éq. 4.2} \quad T\Delta S_{\text{ndu}}(T_{\text{ndu}}) = \frac{T\Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}}$$

On peut alors remplacer le terme associé à l'entropie :

$$\text{éq. 4.3} \quad \Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - \frac{T\Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}} + \Delta C_{p_{\text{ndu}}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

Il ne reste plus qu'à développer l'équation pour prendre en compte la dissociabilité du système en ajoutant $-RT \ln K_{\text{ndu}}(T)$:

$$\Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - \frac{T \Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}} - RT \ln K_{\text{ndu}}(T) + \Delta C p_{\text{ndu}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

éq. 4.4 (a)

$$\Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - T \left(\frac{\Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}} + R \ln K_{\text{ndu}}(T) \right) + \Delta C p_{\text{ndu}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

éq. 4.4 (b)

Dans le but de simplifier le logarithme de $K_{\text{ndu}}(T)$, on fait appel à l'approximation suivante (éq. 4.5 (a)), où $[N]$ est la concentration maximum possible de complexes natifs (*associés*), n est le nombre total de mères dans le complexe et m_i est le nombre de copies m de chaque différente mère i :

$$\ln K_{\text{ndu}}(T) \approx \frac{\ln [N]^{n-1} \cdot n^{n/2} \cdot \prod_{i=1}^k (m_i)^{m_i}}{(\sqrt{n} + 1)^{n-1}}$$

éq. 4.5 (a)

D'où l'on obtient que :

$$\text{Max}'\text{SH}/\text{Max}'\text{SH} \Leftrightarrow 2 \text{Max}'\text{SH}$$

$$A_n \Leftrightarrow m_i A$$

$$n = 2, m_i = 2$$

éq. 4.5 (b)
$$\ln K_{\text{ndu}}(T_{\text{ndu}})_{\text{HOMO}} \approx \frac{\ln[N]^1 \cdot 2^1 \cdot 2^2}{(\sqrt{2} + 1)^1} = \ln[N] \cdot \frac{8}{2,4142}$$

éq. 4.5 (c)
$$\ln K_{\text{ndu}}(T_{\text{ndu}})_{\text{HOMO}} \approx \ln[N] \cdot 3,3137$$

Que l'on remplace alors dans l'expression de $\Delta G_{\text{ndu}}(T)$ (éq. 4.4 (b)) :

$$\Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - T \left(\frac{\Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}} + R \ln([N] \cdot 3,3137) \right) + \Delta C p_{\text{ndu}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

éq. 4.6 (a)

On peut ce faisant exprimer la variation d'énergie libre comme une fonction dépendante de la concentration totale de protéine $[T]$ en substituant $[N]$ par $\frac{[T]}{2}$:

$$\Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - T \left(\frac{\Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}} + R \ln \left(\frac{3,3137[T]}{2} \right) \right) + \Delta C p_{\text{ndu}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

éq. 4.6 (b)

$$\Delta G_{\text{ndu}}(T) = \Delta H_{\text{ndu}}(T_{\text{ndu}}) - T \left(\frac{\Delta H_{\text{ndu}}(T_{\text{ndu}})}{T_{\text{ndu}}} + R \ln(1,6569[T]) \right) + \Delta C p_{\text{ndu}} \left(T - T_{\text{ndu}} - T \ln \left(\frac{T}{T_{\text{ndu}}} \right) \right)$$

éq. 4.6 (c)

On obtient alors l'expression 4.6 (c), permettant de calculer $\Delta G_{\text{ndu}}(T)$ et de simuler les variations de populations natives et dénaturées d'un système dissociatif homodimérique à l'aide de seulement trois différents variables soit : l'enthalpie standard de Van't Hoff ($\Delta H_{\text{ndu}}(T_{\text{ndu}})$), une température de transition (T_{ndu}) et une variation de capacité calorifique ($\Delta C p_{\text{ndu}}$). La concentration de protéine totale ($[T]$) étant connue préalablement à l'expérience, le seul autre paramètre restant est la température (T) à

laquelle on désire simuler les populations du système. Le nombre de point de la simulation doit être suffisant pour recouvrir adéquatement les valeurs expérimentales recueillies lors de l'enregistrement de la dénaturation thermique.

En variant les valeurs des différents paramètres de l'équation de Gibbs-Helmholtz on influe sur les valeurs de $\Delta G_{\text{ndu}}(T)$ et ce faisant sur les valeurs de $K_{\text{ndu}}(T)$, qui à leur tour modifient les valeurs des populations P_U et P_N . En multipliant les populations par les intensités spectroscopiques extensives des états natif et dénaturé on peut alors simplement additionner les signaux synthétiques et retrouver le signal simulé, qu'on superpose à la série de valeurs expérimentales et qu'on ajuste progressivement jusqu'à juger la simulation acceptable.

4.2.3. Influence de l'enthalpie standard de Van't Hoff : $\Delta H_{\text{ndu}}(T_{\text{ndu}})$

Des trois paramètres dont on dispose pour ajuster la simulation aux valeurs expérimentales, l'enthalpie de transition est de loin la plus significative. L'enthalpie influence l'allure de la transition en déterminant la rudesse de celle-ci. L'enthalpie correspond à la sensibilité thermique de la structure, ou plus simplement la tendance de celle-ci à résister aux altérations malgré de nouveaux apports de chaleur. Elle permet de simuler la coopérativité du système, soit l'interdépendance des interactions microscopiques les unes envers les autres et dans l'intégrité de la structure macroscopique native. Les protéines qui exhibent une forte coopérativité dans leur

transition sont moins assujetties à explorer de nombreuses conformations de structure puisque toutes les interactions contribuent au maintien de l'ensemble global. Les protéines de ce type présentent donc une transition franche et une température de transition claire. À l'opposé, les protéines lâches (moins rigides entropiquement) présenteront une faible solidarité et auront tendance à se répartir parmi de nombreuses structures homologues alternatives, puisque les réseaux d'interactions qui composent l'état natif sont moins bien organisés. L'influence de la valeur de l'enthalpie est propre aux systèmes étudiés, puisqu'elle est partie intégrante dans la nature associative/dissociative, homo ou hétéro et dépend des paramètres d'ordre des complexes à l'étude. Le graphique suivant présente l'influence de la variation d'enthalpie sur la coopérativité d'une transition simulée à l'aide des équations précédentes. Du bleu vers le rouge, l'enthalpie augmente et la pente de la transition se redresse. Les valeurs typiques d'enthalpie pour des systèmes coopératifs comme celui de l'homodimère de Max'SH sont de l'ordre de 75 ± 25 kilocalories par mole.

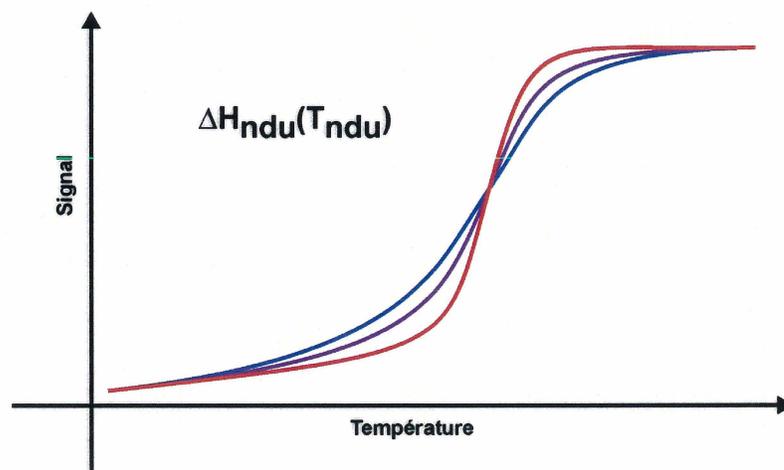


Figure 28 : Influence de l'enthalpie standard de Van't Hoff sur la simulation

4.2.4. Influence de la température de transition : T_{ndu}

La température de transition représente le deuxième paramètre en importance dans la simulation et détermine la température à laquelle l'intégrité de la structure subit l'essentiel de ses altérations. L'influence de la température de transition est présentée sur le graphique suivant, allant du bleu au rouge pour des températures de transition de plus en plus élevées. Les valeurs des températures de transition sont le reflet de la proportion d'interactions favorables/défavorables à l'intérieur de la structure native. Plus les interactions favorables seront nombreuses et/ou rigides au sein d'une structure, plus celle-ci tolérera de fortes températures avant d'exhiber une transition. Il n'existe pas de valeurs typiques de température de transition, puisque dans le cas de complexes multimériques, celles-ci peuvent être influencées par de nombreux facteurs dont la concentration de protéine, la force ionique (NAUD *et al.*, 2005) et la présence de système compétiteurs (résultat de l'article).

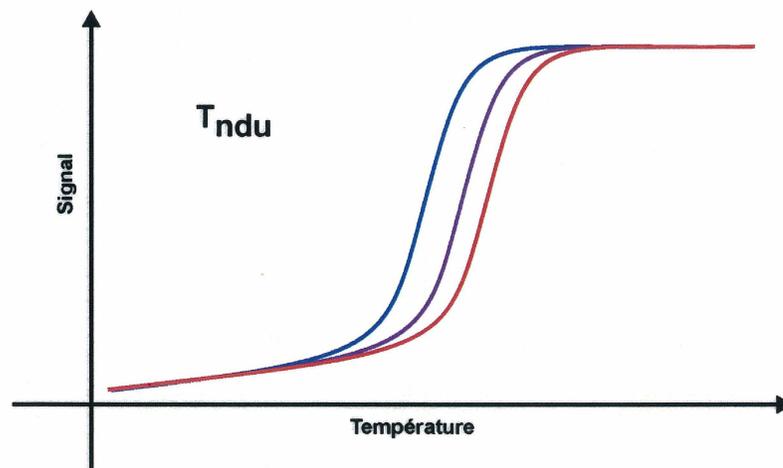


Figure 29 : Influence de la température de transition sur la simulation

4.2.5. Influence de la variation de la capacité calorifique : $\Delta C_{p_{ndu}}$

L'influence de la variation de la capacité calorifique associée à la transition est présentée sur le graphique ci-dessous avec encore une fois, de bleu à rouge, des valeurs de plus en plus élevées.

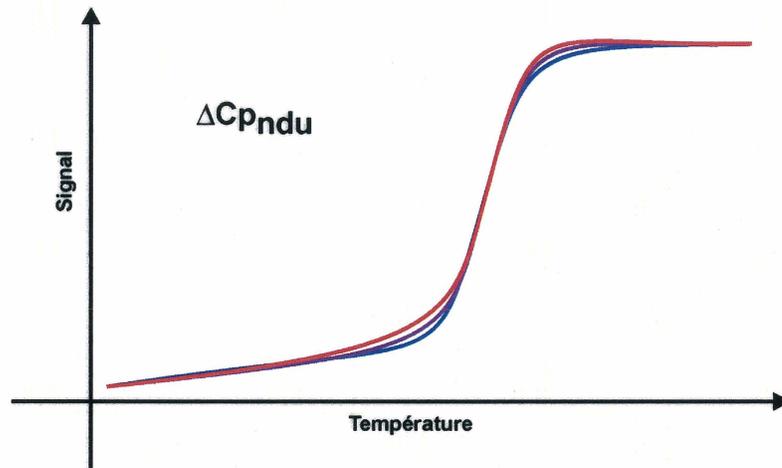


Figure 30 : Influence de la variation de la capacité calorifique sur la simulation

Bien que la variation de la capacité calorifique doive normalement être considérée comme dépendante de la température et que nous ne tenions pas compte de cette dépendance dans les équations précédentes, on remarquera qu'il s'agit du paramètre ayant le moins d'effet sur l'allure de la simulation mais aussi sur les valeurs de K_{ndu} calculées (MYERS *et al.*, 1995). En réalité, les effets les plus pervers liés à l'influence de la variation de la capacité calorifique sont plus fortement observés aux limites extrêmes de la simulation. Ils exposent d'une part, les effets de la dénaturation par le froid dans les systèmes plus sensibles, ou d'autre part provoquent des aberrations de calcul aux températures supérieures à 85°C. Pour les régions avoisinantes de la

transition, l'influence de la variation de la capacité calorifique permet d'ajuster l'ardeur parfois précipitée de la transition, un phénomène évoquant la susceptibilité de la structure à sa propre accumulation de chaleur. Plus une structure est sujette à accumuler une grande quantité de chaleur (grand ΔC_p), plus elle présentera une amorce et une fin de transition précoce. La structure surchauffe la dénaturation et la devance. Inversement, moins la structure peut accumuler de chaleur (petit ΔC_p) plus l'amorce de la transition sera franche et la fin élancée. La structure tempère la dénaturation et la retarde. L'ordre de grandeur de la variation de la capacité calorifique associée à la transition de structures de la taille des b-HLH-LZ se situe à environ $0,75 \pm 0,25$ kcal par mole et par Kelvin, et des déviations de 30% n'affectent pas de façon significative les valeurs de K_{ndu} déterminées.

4.2.6. Tableur CD_F « CD Fitting »

En dérivant les équations de Gibbs-Helmholtz comme précédemment, nous avons développé (annexes I-VI) l'ensemble des systèmes d'équation nécessaires à la simulation des dénaturations thermiques des équilibres monomériques et dimériques, de type homo et hétéro, à deux et à trois états (*c'est-à-dire incluant un état de transition intermédiaire*). Ces systèmes d'équation ont été regroupés dans un seul et même tableur « Microsoft Excel™ » de manière à rendre leur manipulations plus efficaces. Le tableur a été assorti d'outils de commande et d'une interface présentant les résultats en simultané afin d'accélérer le raffinement des simulations. Le tableur, intitulé « CD_F »

pour « *CD fitting* », permet de traiter jusqu'à 5 dénaturations thermiques en parallèle et de choisir pour chacune les systèmes d'équations de Gibbs-Helmholtz désirés. Les valeurs des paramètres spectroscopiques et thermodynamiques peuvent ainsi être altérées et les résultats de leurs implications sur la simulation peuvent être accédés sur une seule et même feuille de calcul.

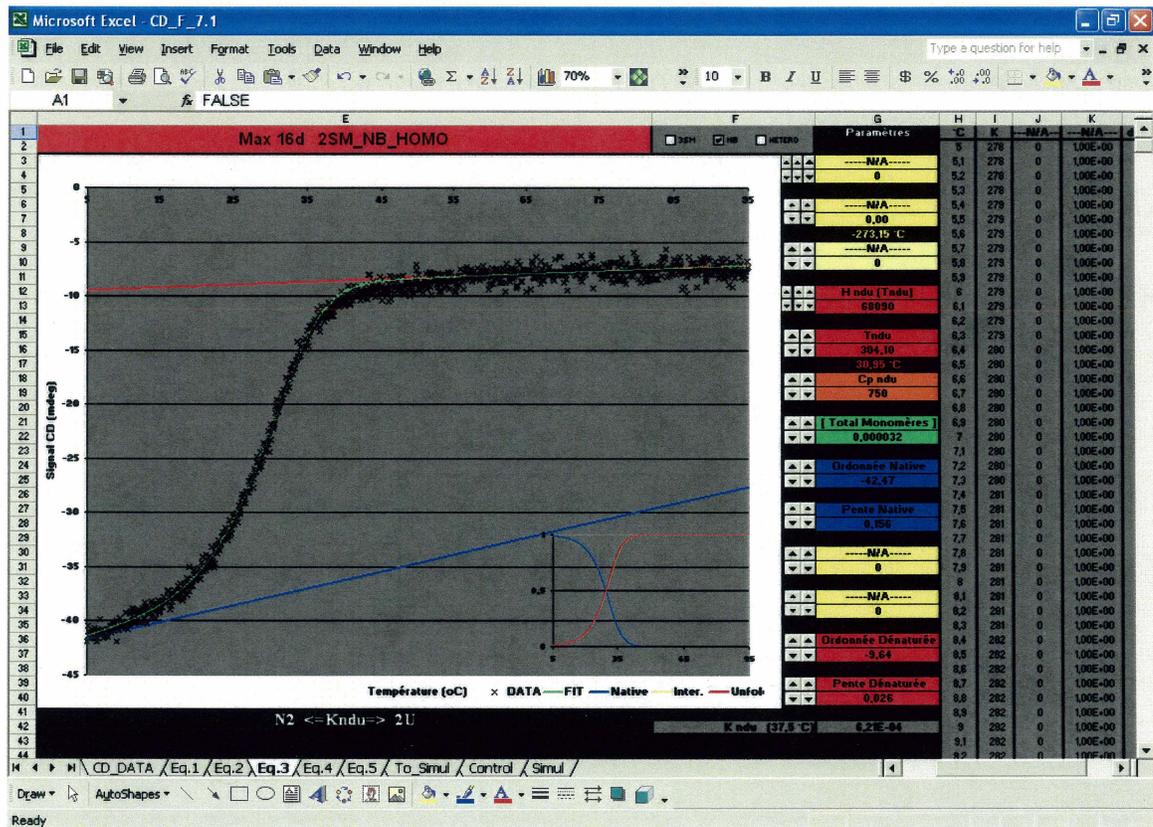


Figure 31 : Exemple d'une simulation à l'aide des équations de Gibbs-Helmholtz
Saisie d'écran de l'interface du tableur CD_F.

Le tableur CD_F a également été implémenté de macros commandes permettant l'export, sous forme de matrices, des résultats obtenus des simulations. Il permet ainsi de produire rapidement les listes des valeurs de ΔG , de ΔH , des K_{DS} de même que celles

des intensités spectroscopiques extensives des différents états, et ce pour l'ensemble des températures. Cette particularité est rendue nécessaire pour permettre la communication du tableur CD_F, avec les programmes complémentaires « CD_SI » (*CD Simulator-Iterator*) et « KD_I » (*KD Iterator*) qui ont été développés pour résoudre le système plus complexe de c-Myc'SH/Max'SH.

4.2.7. Irréductibilité des équations du système de c-Myc'SH/Max'SH

Bien que les systèmes d'équations permettant de simuler un système dissociatif hétérodimérique aient été dérivés et inclus dans le tableur CD_F, les résultats escomptables de telles simulations du système de c-Myc'SH/Max'SH ne sont en réalité d'aucune valeur significative. Ce n'est qu'après avoir dérivé l'ensemble des équations hétérodimériques que l'on s'aperçut que les valeurs de K_D déterminées de la sorte seraient toujours erronées. Car, bien que les équations soient exactes, le système d'équations développé pour les équilibres hétérodimériques ne tient pas compte de l'existence possible d'équilibres compétiteurs homodimériques. Il se résume à la simulation d'une simple dissociation sans tenir compte de la possibilité de former des complexes alternatifs. Les systèmes d'équation de Gibbs-Helmholtz ainsi développés ne permettent que d'approximer l'équilibre global de c-Myc'SH/Max'SH en le résumant bêtement à l'équilibre suivant :

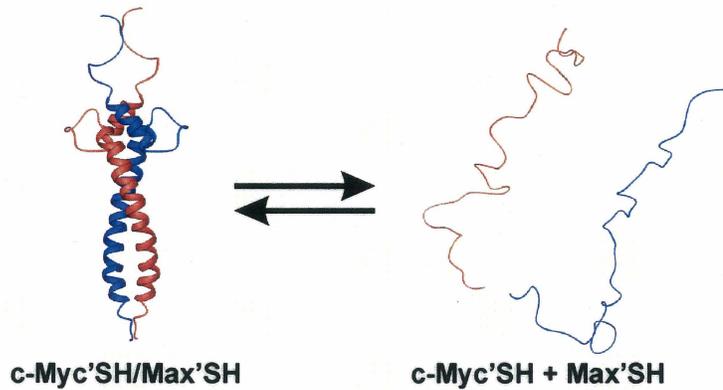


Figure 32 : Équilibre simplifié de c-Myc'SH et Max'SH

Autrement, le nombre d'inconnus devient incompatible avec la méthode. On ne peut ainsi, en une seule et même simulation, parvenir à déterminer trois K_D différents et cinq intensités spectroscopiques extensives, sans risquer de produire des aberrations grotesques. De plus, alors que les K_D s des homodimères sont seulement dépendants de leurs propres concentrations mais selon la puissance 2 (éq. 5.0 & 5.1). Le K_D d'un hétérodimère, lui, est dépendant des concentrations des deux partenaires, et selon la puissance 1 (éq. 5.2) :

éq. 5.0
$$K_{D, \text{Max'SH} / \text{Max'SH}} = \frac{[\text{Max'SH}][\text{Max'SH}]}{[\text{Max'SH} / \text{Max'SH}]} = \frac{[\text{Max'SH}]^2}{[\text{Max'SH} / \text{Max'SH}]}$$

éq. 5.1
$$K_{D, \text{c-Myc'SH} / \text{c-Myc'SH}} = \frac{[\text{c-Myc'SH}][\text{c-Myc'SH}]}{[\text{c-Myc'SH} / \text{c-Myc'SH}]} = \frac{[\text{c-Myc'SH}]^2}{[\text{c-Myc'SH} / \text{c-Myc'SH}]}$$

éq. 5.2
$$K_{D, \text{c-Myc'SH} / \text{Max'SH}} = \frac{[\text{c-Myc'SH}][\text{Max'SH}]}{[\text{c-Myc'SH} / \text{Max'SH}]}$$

Ce faisant, le système est non linéaire et ne peut être simplement réduit par l'isolement et la substitution d'une des variables [c-Myc'SH] ou [Max'SH], puisqu'on doit nécessairement conserver les deux membres pour rendre justice à l'équilibre global. On a donc ici affaire à un système d'équation non linéaire à multiples variables intriquées. Ce genre de système d'équation ne trouve de solution que par des méthodes de résolution numérique comme la méthode de Newton-Raphson ou par itérations récurrentes, et nécessite obligatoirement la détermination préalable de certaines variables.

Ainsi, toute expérience de titration des partenaires l'un à l'autre et/ou méthode d'approximation du système *via* la simplification précédente, ne peut conduire qu'à des valeurs de K_D inexactes, puisque la compétition des homodimères n'est prise en compte d'aucune manière. Comme la littérature n'a fait état d'aucune considération à ce sujet dans la détermination des K_{DS} de c-Myc/Max (BANERJEE *et al.*, 2006; FIEBER *et al.*, 2001; HU *et al.*, 2005; KOHLER *et al.*, 1999; NAIR et BURLEY, 2003), et puisque des évidences nous montrent que les homodimères entrent directement en jeu au cours des activités biologiques de c-Myc (NAUD *et al.*, 2005), il nous est apparu indispensable d'élaborer une méthode efficace et robuste pour résoudre de tels systèmes d'équation et parvenir à en simuler les équilibres et les populations sur de vastes plages de température et de concentration.

Comme la solution est de type non-linéaire mais continue, la méthode retenue fut celle d'une régression par itérations récurrentes. La méthode consiste essentiellement à

converger vers la solution exacte en circonscrivant d'abord celle-ci par ses limites physiques réelles, puis au raffinement de la solution de manière récurrente jusqu'à la précision escomptée. En plus de permettre un traitement rapide d'équations simples, cette méthode a l'avantage d'être flexible et aisément programmable.

4.3. Simulation de c-Myc'SH + Max'SH à l'aide du programme CD_SI

Parmi les nombreuses avenues disponibles pour résoudre les équations de systèmes non-linéaires présentant des variables intriquées, nous avons préconisé la régression par itérations récurrentes, soit une méthode d'optimisation par essais et erreurs. Ce choix est basé sur la simplicité de programmation d'une telle routine, comparé à l'implémentation de solutions numériques fournies par des méthodes comme celle de Newton-Raphson, de l'interpolation linéaire ou encore la méthode de Graeffe (REINHARDT et SOEDER, 1997). Concrètement, la différence se situe dans la façon de poser le problème à l'ordinateur. L'itération multiple à l'avantage d'utiliser des équations mathématiques simples et concrètes, avec des calculs rapides, effectués sous la forme d'optimisation, tandis que les autres méthodes utilisent des dérivations abstraites et des calculs fastidieux souvent difficiles à implémenter sans erreurs. La programmation d'un système d'itération présente également l'énorme avantage d'être versatile, ce que les autres méthodes ne sont pas. On peut ainsi conserver intact le cœur d'un programme d'itération et adapter rapidement les équations d'un nouveau problème sans avoir à repenser celui-ci en entier.

La méthode par itérations récurrentes a par contre un fâcheux défaut, celui d'engendrer un nombre considérable de calculs, comparativement aux autres méthodes qui trouvent les solutions du premier jet. Ainsi, le problème est déplacé à l'ordinateur et aux algorithmes qui traitent les solutions, ces derniers devant être optimisés pour utiliser efficacement les ressources computationnelles. Le jeu en vaut tout de même la

chandelle, puisque le temps dépensé en calcul est largement compensé par celui épargné en dérivation et en programmation.

4.3.1. Programme CD_SI « CD Simulator Iterator »

Le programme CD_SI a été élaboré sous forme d'un programme en macro-commandes VBA « Visual-Basic for Applications » à même le tableur « Microsoft Excel™ ». La méthode a été élaborée en partant des prémisses qui sont que : « *Le signal CD total d'un mélange ($CD_{c-Myc'SH+Max'SH}$) mesuré par l'appareil est égal à la somme des signaux CD individuels de tous les constituants du mélange* » (éq. 6.0) :

$$CD_{c-Myc'SH+Max'SH} = CD_{c-Myc'SH} + CD_{c-Myc'SH/c-Myc'SH} + CD_{Max'SH} + CD_{Max'SH/Max'SH} + CD_{c-Myc'SH/Max'SH}$$

éq. 6.0

Et que ce faisant : « *Il est juste de poser que chaque signal CD spécifique est égal au produit du signal CD molaire (extensif) de l'espèce (MCD_A) multiplié par la concentration de cette espèce $[A]$* » (éq. 6.1) :

$$\begin{aligned}
 CD_{c-Myc'SH+Max'SH} &= [c-Myc'SH] \cdot MCD_{c-Myc'SH} \\
 &+ [c-Myc'SH/c-Myc'SH] \cdot MCD_{c-Myc'SH/c-Myc'SH} \\
 &+ [Max'SH] \cdot MCD_{Max'SH} \\
 &+ [Max'SH/Max'SH] \cdot MCD_{Max'SH/Max'SH} \\
 &+ [c-Myc'SH/Max'SH] \cdot MCD_{c-Myc'SH/Max'SH}
 \end{aligned}$$

éq. 6.1

La méthode consiste donc à simuler les valeurs expérimentales $CD(T)$ de la dénaturation thermique du mélange de c-Myc'SH et de Max'SH ($CD_{c-Myc'SH+Max'SH}(T)$), en reconstituant celle-ci par itération à partir de ses constituants. Les calculs sont basés sur les paramètres extraits des simulations des dénaturations thermiques des protéines pures c-Myc'SH et Max'SH *via* Gibbs-Helmoltz. On procède donc d'abord aux dénaturations thermiques indépendantes de 32 μ M de c-Myc'SH pure et de 32 μ M de Max'SH pure. On procède ensuite à la simulation des données *via* Gibbs-Helmoltz et les équations du tableur CD_F, puis on exporte les résultats vers le programme CD_SI. Les données exportées sont celles des $K_D(T)$ s de c-Myc'SH/c-Myc'SH et de Max'SH/Max'SH, de même que les signaux spectroscopiques extensifs utilisés pour simuler les différentes espèces de c-Myc'SH et c-Myc'SH/c-Myc'SH d'une part, et de Max'SH et Max'SH/Max'SH d'autre part. Dans le cas présent, les valeurs des signaux spectroscopiques sont exprimées sous la forme de CD molaire « MCD » (*pour Molar CD*) soit l'intensité du signal de dichroïsme circulaire relativisé pour une mole d'espèce. Les valeurs de MCD étant dépendantes de la température « MCD(T) », elles sont, tout comme les valeurs de $K_D(T)$, regroupées en tableau et indexées pour chaque incrément de 0,1°C allant de 5 à 95°C. Les résultats exportés représentent donc une matrice de 8 colonnes par 902 lignes, ayant la forme du tableau suivant :

T (°C)	$K_D(T)$, c-Myc'SH/ c-Myc'SH	MCD(T) c-Myc'SH	MCD(T) c-Myc'SH/ c-Myc'SH	$K_D(T)$, Max'SH/ Max'SH	MCD(T) Max'SH	MCD(T) Max'SH/ Max'SH	MCD(T) c-Myc'SH/ Max'SH
5,0	5,73E-02	-176406	-349375	1,25E-05	-256563	-2605625	-2470000
5,1	5,80E-02	-176397	-349293	1,26E-05	-256544	-2604650	-2468962
5,2

Figure 33 : Aperçu de la matrice d'exportation des résultats de CD_SI

La huitième et dernière colonne de la matrice contient la liste des valeurs de $MCD(T)$ pour le complexe hypothétique c-Myc'SH/Max'SH. Bien que ces dernières valeurs ne nous soient accessibles directement, il est possible d'en déterminer une bonne approximation en réalisant une dénaturation thermique du complexe d'hétérodimère « spécifique » c-Myc'SSMax'. Le complexe c-Myc'SSMax', lié de façon covalente à l'aide d'un pont disulfure, est indépendant de la concentration et stable quant à d'éventuels réarrangements sous formes d'homodimères. Ce faisant, la construction nous permet d'approximer des valeurs de $MCD(T)$ pour une solution jugée 100% pure en hétérodimère c-Myc'SH/Max'SH. La détermination des valeurs de $MCD_{c-Myc'SH/Max'SH}(T)$ est donc réalisée de la même manière que pour les autres complexes (en ajustant une droite à l'état natif des valeurs expérimentales), mais cette fois-ci, d'une solution de 32 μM de c-Myc'SSMax' (voir article Figure 10 a).

Une fois l'ensemble des données exportées vers le programme CD_SI, on fournit à ce dernier les valeurs de la dénaturation thermique que l'on souhaite simuler, soit celle du vrai mélange de 32 μM de c-Myc'SH supplémenté de 32 μM de Max'SH. La simulation s'appuie sur l'existence d'une différence d'intensité de signal détectable entre : « la somme théorique (Σ) des dénaturations de c-Myc'SH et Max'SH pures » et d'autre part, « le résultat expérimental de leur vrai mélange ». Si aucune différence n'existe, le programme en conclut qu'il n'y a aucune hétérodimérisation des partenaires, mais si la courbe expérimentale possède un signal plus intense que la somme théorique, le programme tentera alors de combler cette différence en formant des complexes c-Myc'SH/Max'SH au détriment des homodimères.

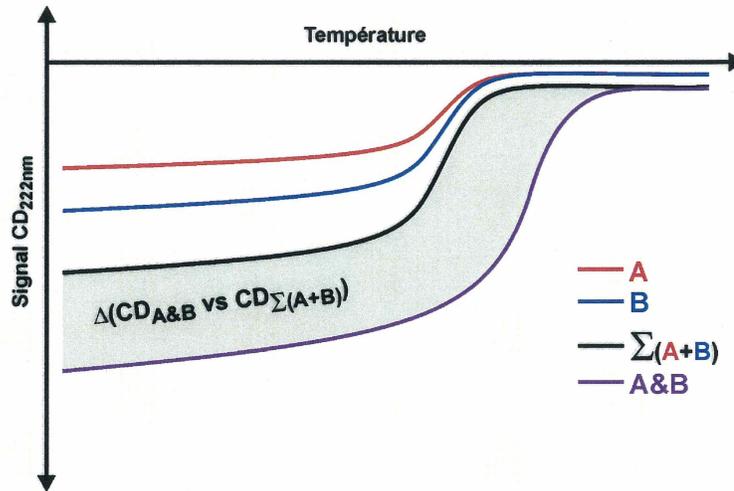


Figure 34 : Reconnaissance moléculaire observée par CD

Le programme procèdera alors à des « rondes de calcul » durant lesquelles il tentera de déterminer les concentrations optimales de complexe c-Myc'SH/Max'SH satisfaisant le mieux chaque point de la dénaturation thermique du mélange. Pour chaque ronde de calcul, le programme définit d'abord cinq différentes concentrations de complexe [c-Myc'SH/Max'SH]_i. Les concentrations sont choisies à l'intérieur d'une plage limitante ([*minimum*] et [*maximum*]) qui commence d'abord par les extrémités physiques 0 et 32 μM , et se rétrécit par la suite, de ronde en ronde, jusqu'à atteindre une précision de l'ordre du femto molaire (10^{-15}M). Pour chaque itération « *i* », le programme rétro-calcule les concentrations des espèces monomériques et homodimériques de [c-Myc'SH]_i, [c-Myc'SH/c-Myc'SH]_i, [Max'SH]_i et [Max'SH/Max'SH]_i à l'aide des valeurs de $K_D(T)$ importées. Finalement le programme multiplie chaque espèce par sa valeur propre de $MCD(T)$ pour obtenir la valeur de signal $CD(T)$ simulé (éq. 6.1). Les deux concentrations de [c-Myc'SH/Max'SH]_i dont les simulations circonscrivent le mieux la valeur empirique, sont retenues comme nouvelles limites minimum et

maximum de la plage de concentration et le programme reprend une nouvelle ronde avec cinq nouvelles concentrations un peu plus précises que les précédentes. Une fois que l'écart entre les limites se rétrécit à une précision de l'ordre de $10^{-15}M$ (*soit après environ quinze rondes de calcul* $\left(\frac{32E-6M}{5^{15}} = 1E-15M\right)$), le programme effectue la moyenne des deux concentrations limites, calcule et enregistre les valeurs des concentrations de toutes les espèces, puis recommence le processus avec un nouveau point de la dénaturation thermique et ce jusqu'à atteindre 95°C.

4.3.2. Retro-calcul des concentrations de monomères et d'homodimères

Pour chaque concentration de complexe $[c\text{-Myc}'SH/Max'SH]_i$ (itérée), le programme libère une quantité transitoire (*) de c-Myc et de Max ($[c\text{-Myc}'SH]_{i}^*$ et $[Max'SH]_{i}^*$) qu'il redistribue en fonction des $K_D(T)$ s importés de CD_F. Pour ce faire, il définit d'abord les concentrations de protéine non utilisées, par simple soustraction à partir des concentrations de protéine totale initiales ($[c\text{-Myc}'SH]_o$ et $[Max'SH]_o$) selon :

$$\text{éq. 7.0 (a)} \quad [c\text{-Myc}'SH]_{i}^* = [c\text{-Myc}'SH]_o - [c\text{-Myc}'SH/Max'SH]_i$$

$$\text{éq. 7.0 (b)} \quad [Max'SH]_{i}^* = [Max'SH]_o - [c\text{-Myc}'SH/Max'SH]_i$$

Le programme peut alors, connaissant les valeurs de $K_{D,c-Myc'SH/c-Myc'SH}(T)$ et $K_{D,Max'SH/Max'SH}(T)$, recalculer les concentrations proportionnelles des homodimères $[c-Myc'SH/c-Myc'SH]_i$ et $[Max'SH/Max'SH]_i$ à l'aide des deux équations suivantes:

$$[c-Myc'SH/c-Myc'SH]_i = \frac{4[c-Myc'SH]_i^* + K_{D,c-Myc'SH/c-Myc'SH} - \sqrt{(4[c-Myc'SH]_i^* + K_{D,c-Myc'SH/c-Myc'SH})^2 - 16[c-Myc'SH]_i^{*2}}}{8}$$

éq. 7.1 (a)

$$[Max'SH/Max'SH]_i = \frac{4[Max'SH]_i^* + K_{D,Max'SH/Max'SH} - \sqrt{(4[Max'SH]_i^* + K_{D,Max'SH/Max'SH})^2 - 16[Max'SH]_i^{*2}}}{8}$$

éq. 7.1 (b)

Ayant défini les concentrations des homodimères, le programme obtient les concentrations de chaque monomère par les différences :

$$\text{éq. 7.2 (a)} \quad [c-Myc'SH]_i = [c-Myc'SH]_i^* - 2[c-Myc'SH/c-Myc'SH]_i$$

$$\text{éq. 7.2 (b)} \quad [Max'SH]_i = [Max'SH]_i^* - 2[Max'SH/Max'SH]_i$$

Les équations quadratiques précédentes (éq. 7.1 (a & b)) sont obtenues par développement à partir des expressions des K_{DS} et des équations de conservation de la masse, selon la démarche suivante (*le développement est présenté seulement pour $[Max'SH/Max'SH]_i$ puisqu'il est identique pour $[c-Myc'SH/c-Myc'SH]_i$*).

Sachant que :

$$\text{éq. 7.3 (a)} \quad K_{D, \text{Max}'\text{SH} / \text{Max}'\text{SH}} = \frac{[\text{Max}'\text{SH}]^2}{[\text{Max}'\text{SH} / \text{Max}'\text{SH}]}$$

$$\text{éq. 7.3 (b)} \quad K_{D, \text{Max}'\text{SH} / \text{Max}'\text{SH}} [\text{Max}'\text{SH} / \text{Max}'\text{SH}] = [\text{Max}'\text{SH}]^2$$

On peut, selon la conservation de la masse, poser que les concentrations de monomère et de dimère pour une itération « i » sont données par :

$$\text{éq. 7.4 (a)} \quad [\text{Max}'\text{SH}]_i^* = [\text{Max}'\text{SH}]_i + 2[\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i$$

$$\text{éq. 7.4 (b)} \quad [\text{Max}'\text{SH}]_i = [\text{Max}'\text{SH}]_i^* - 2[\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i$$

En substituant la valeur de $[\text{Max}'\text{SH}]_i$ (éq. 7.4 (b)) dans l'équation du K_D (éq. 7.3 (b)), on obtient une expression qui se réarrange sous la forme d'un polynôme du type : $ax^2 + bx + c$:

$$\text{éq. 7.5 (a)} \quad K_{D, \text{Max}'\text{SH} / \text{Max}'\text{SH}} [\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i = ([\text{Max}'\text{SH}]_i^* - 2[\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i)^2$$

$$0 = 4[\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i^2 - (4[\text{Max}'\text{SH}]_i^* + K_{D, \text{Max}'\text{SH} / \text{Max}'\text{SH}}) [\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i + [\text{Max}'\text{SH}]_i^*{}^2$$

éq. 7.5 (b)

Qu'on peut donc résoudre pour $[\text{Max}'\text{SH} / \text{Max}'\text{SH}]_i$ via les racines quadratiques :

$$\text{éq. 7.6 (a & b)} \quad r_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad r_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Puisque seule la deuxième racine est valable, soit (r_2) (éq. 7.6 (b)), on retrouve

l'équation 7.1 (b) :

$$[\text{Max}'\text{SH}/\text{Max}'\text{SH}]_i = \frac{4[\text{Max}'\text{SH}]_i^* + K_{D,\text{Max}'\text{SH}/\text{Max}'\text{SH}} - \sqrt{\left(4[\text{Max}'\text{SH}]_i^* + K_{D,\text{Max}'\text{SH}/\text{Max}'\text{SH}}\right)^2 - 16[\text{Max}'\text{SH}]_i^{*2}}}{8}$$

éq. 7.1 (b)

De la même façon que pour Max'SH, le programme reproduit aussi fidèlement les proportions de monomère [c-Myc'SH]_i libre et d'homodimère [c-Myc'SH/c-Myc'SH]_i qui se formeraient normalement en fonction des $K_D(T)$ s déterminés *via* Gibbs-Helmoltz et des concentrations de [c-Myc'SH]_i* laissées pour compte dans la titration virtuelle.

4.3.3. Résultats de la simulation à l'aide de CD_SI

Les résultats obtenus de la simulation à l'aide de CD_SI se présentent sous la forme de deux graphiques et d'une matrice. On peut y lire, pour chaque incrément de 0,1°C, la concentration de chacune des différentes espèces c-Myc'SH, c-Myc'SH/c-Myc'SH, Max'SH, Max'SH/Max'SH et c-Myc'SH/Max'SH de même que les contributions respectives de chaque espèce au signal CD simulé.

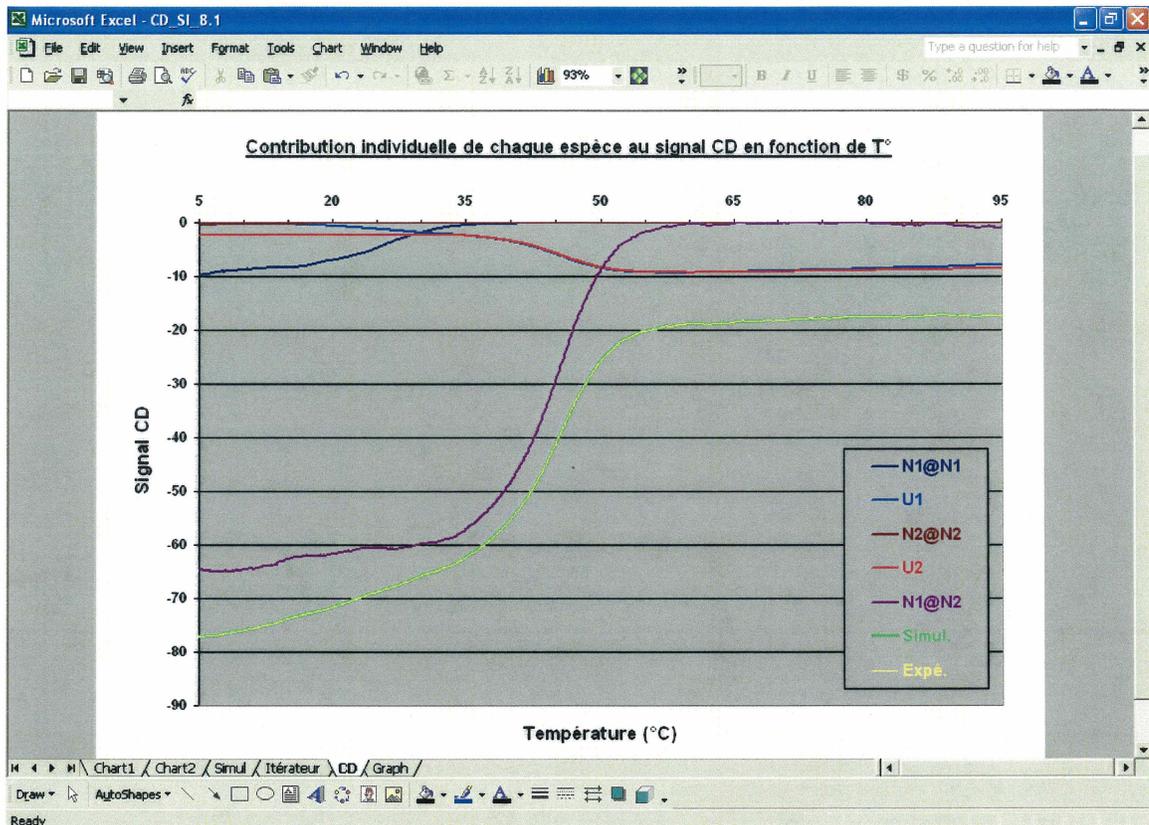


Figure 35 : Contributions individuelles des espèces au signal CD synthétique

Saisie d'écran de l'interface du programme CD_SI.

Les signaux CD individuels étant seulement le produit des concentrations des espèces individuelles multipliées par leur intensité spectroscopique extensive propre.

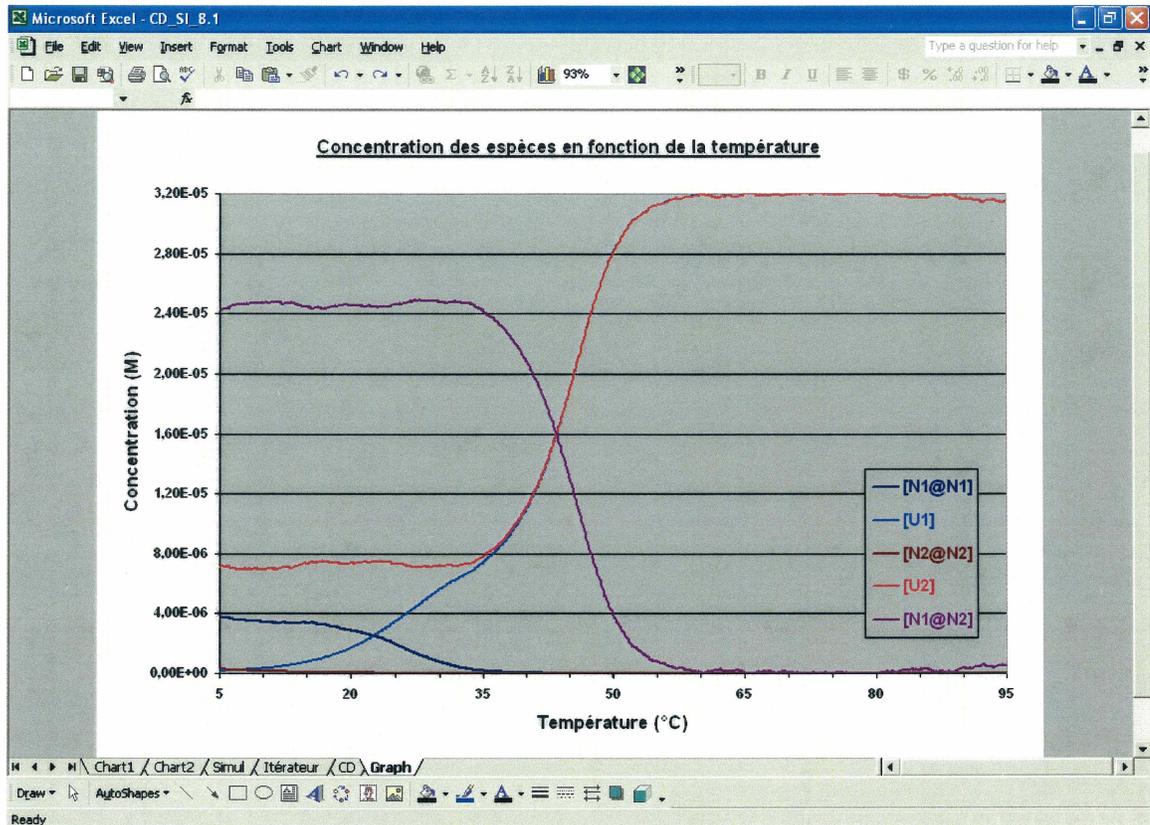


Figure 36 : Concentrations individuelles des espèces en cours de dénaturation

Saisie d'écran de l'interface du programme CD_SI.

4.3.4. Calcul des $K_D(T)$ s « réels » et des $K_D^*(T)$ s « apparents »

Connaissant les concentrations de chaque espèce en solution, de même que les concentrations initiales des protéines, le programme peut alors calculer deux différentes séries de valeurs de $K_D(T)$ pour le nouveau complexe c-Myc'SH/Max'SH. La première série, appelée K_D « apparent » ($K_D^*(T)$; éq. 8.0), est basée sur la concentration de protéine totale en solution. Ces valeurs sont nécessairement affaiblies des valeurs « réelles » de K_D puisqu'une certaine proportion de c-Myc'SH mais surtout de Max'SH

préfère se combiner sous la forme d'homodimère plutôt que d'hétérodimère. L'échange perpétuel entre les différents complexes possibles engendre donc des populations de c-Myc'SH et de Max'SH non complexées qui à leur tour, diminuent l'affinité « apparente » de c-Myc pour Max. Les valeurs de $K_D^*(T)$ sont représentatives de la réalité du système de c-Myc et Max, mais ne sont utiles que pour être comparées aux valeurs de $K_{D,c-Myc/Max}$ déterminées par d'autres groupes de recherche qui oignent de prendre en compte les effets de la compétition des homodimères sur les valeurs d'affinité réelles.

$$K_{D,c-Myc'SH/Max'SH}^* = \frac{([c - Myc'SH] + 2[c - Myc'SH / c - Myc'SH])([Max'SH] + 2[Max'SH / Max'SH])}{[c - Myc'SH / Max'SH]}$$

éq. 8.0

La deuxième série de valeur de K_D que le programme calcul, est fort plus intéressante puisqu'il s'agit des valeurs de $K_D(T)$ « réelles » du complexe c-Myc'SH/Max'SH (éq. 8.1). Celles-ci doivent être déterminées en considérant une proportion moindre de protéine, soit seulement la somme des concentrations de protéines libres c-Myc'SH, Max'SH et de complexes c-Myc'SH/Max'SH. Beaucoup moins intuitive *a priori*, il s'agit en réalité de la définition conventionnelle du K_D , qu'il est impossible d'isoler autrement que par une détermination complète des concentrations de toutes les espèces en jeu, de manière à pouvoir en dégager seulement les partenaires impliqués.

éq. 8.1

$$K_{D,c-Myc'SH/Max'SH} = \frac{[c - Myc'SH][Max'SH]}{[c - Myc'SH / Max'SH]}$$

Il est intéressant de mentionner ici que les concentrations de c-Myc'SH et de Max'SH ne sont plus équivalentes à tous instants au cours de la dénaturation, puisqu'elles fluctuent en fonction de la ténacité des complexes homodimériques. Ceci n'influence en rien la détermination des $K_D(T)$ s réels puisqu'ils sont, comparativement aux $K_D^*(T)$ s, « libres » d'accepter de telles inégalités tout en restant vrais. Le principal avantage de la détermination des $K_D(T)$ s réels est celui de pouvoir par la suite comparer ceux-ci directement avec les $K_D(T)$ s des homodimères déterminés à partir de solution pure. On pourrait ainsi renommer les $K_D(T)$ s « réels » comme « purs » et les $K_D^*(T)$ s « apparents » comme « impurs ». Dû au fait que les $K_D(T)$ s réels des homodimères ont été déterminés dans un référentiel qui n'impliquait pas la compétition de l'hétérodimère, seuls les jeux homogènes de $K_D(T)$ s réels peuvent être comparés entre eux et utilisés pour recalculer des populations. Les $K_D^*(T)$ s apparents, quant à eux, permettent seulement de déterminer la concentration de complexe c-Myc'SH/Max'SH sans permettre de déterminer celles de Max'SH et Max'SH/Max'SH ou de c-Myc'SH et c-Myc'SH/c-Myc'SH. C'est donc ici que notre méthode démontre sa puissance. Elle constitue en réalité la seule façon de paramétriser complètement un tel système, de manière à pouvoir par la suite le resimuler à volonté dans n'importe quelles conditions de température ou de concentration. Une telle resimulation implique cependant à nouveau des calculs avec des variables intriquées. Mais comme nous le verrons, une fois les valeurs de $K_D(T)$ réel calculées, cette étape se résout instantanément lorsqu'on recourt à un programme d'itérations multiples. La tâche de resimuler le système est donc le fardeau du troisième et dernier programme de notre suite, soit le programme appelé KD_I ou : « *KD Iterator* ».

4.4. Simulation de l'équilibre global à l'aide du programme KD_I

L'équilibre de c-Myc'SH/Max'SH est un système d'équations aux variables intriquées, dont l'isolement des unes par rapport aux autres est impossible. On ne peut isoler la concentration du complexe c-Myc'SH/Max'SH des concentrations de c-Myc'SH/c-Myc'SH et de Max'SH/Max'SH, puisqu'elles sont toutes interdépendantes vis-à-vis de la concentration de protéine totale. Il est donc impossible d'exprimer les concentrations de l'un ou l'autre des homodimères par une équation seulement fonction de c-Myc'SH/Max'SH, puisque la concentration de Max'SH/Max'SH dépend forcément à la fois de celle de c-Myc'SH/c-Myc'SH et de c-Myc'SH/Max'SH et *vice-versa*. Pour résoudre les concentrations d'espèces propres à un jeu de K_{DS} et des concentrations de protéine données, il nous faut donc exprimer le problème de manière à tenir uniquement compte de ce qu'il est possible de déduire des conditions finales du système à l'équilibre. Comme on sait qu'à l'équilibre les concentrations de toutes les espèces doivent obligatoirement respecter ensemble les trois K_{DS} réels ($K_{D,c-Myc'SH/c-Myc'SH}$, $K_{D,Max'SH/Max'SH}$, $K_{D,c-Myc'SH/Max'SH}$) et la loi de conservation de la masse, une façon simple d'exprimer cette condition est de poser qu'à l'équilibre, l'équation 9.0, représentant l'équilibre global intriqué, doit obligatoirement équivaloir « 1 » :

$$\frac{\sqrt{K_{D,c-Myc'SH/c-Myc'SH}} [c-Myc'SH / c-Myc'SH] \cdot \sqrt{K_{D,Max'SH/Max'SH}} [Max'SH / Max'SH]}{K_{D,c-Myc'SH/Max'SH} [c-Myc'SH / Max'SH]} = 1$$

éq. 9.0

4.4.1. Programme KD_I « KD Iterator »

Le programme KD_I attaque la tâche de déterminer les populations d'un jeu de K_D donné, de la même façon que le programme CD_SI aborde une dénaturation thermique, sauf qu'au lieu de tenter de simuler un signal CD, le programme tentera d'ajuster les concentrations des espèces pour faire égaliser à « 1 » l'équation de l'équilibre global intriqué. L'équation de l'équilibre global intriqué est obtenue à partir des expressions des K_{DS} réels des complexes individuels de la façon suivante. On devra adapter l'équation de l'équilibre de c-Myc'SH/Max'SH pour y inclure les compétitions engendrées par les K_{DS} des homodimères. Pour ce faire, on pose donc d'abord, l'expression simple du K_D réel de c-Myc'SH/Max'SH (éq. 8.1).

$$\text{éq. 8.1} \quad K_{D,c\text{-Myc'SH}/\text{Max'SH}} = \frac{[c\text{-Myc'SH}][\text{Max'SH}]}{[c\text{-Myc'SH}/\text{Max'SH}]}$$

Dans cette équation, les concentrations de c-Myc'SH libre ($[c\text{-Myc'SH}]$) et de Max'SH libre ($[\text{Max'SH}]$) sont uniquement dépendantes du K_D de l'hétérodimère, alors qu'elles devraient pouvoir également engendrer des homodimères. Pour remédier à la situation, on pose alors les K_{DS} « réels » des homodimères, d'où l'on isole les concentrations des partenaires libres.

Soit pour $[c\text{-Myc'SH}]$:

$$\text{éq. 9.1(a)} \quad K_{D,c\text{-Myc'SH}/c\text{-Myc'SH}} = \frac{[c\text{-Myc'SH}][c\text{-Myc'SH}]}{[c\text{-Myc'SH}/c\text{-Myc'SH}]} = \frac{[c\text{-Myc'SH}]^2}{[c\text{-Myc'SH}/c\text{-Myc'SH}]}$$

$$\text{éq. 9.1(b)} \quad [c - \text{Myc}'\text{SH}]^2 = K_{D,c-\text{Myc}'\text{SH}/c-\text{Myc}'\text{SH}} [c - \text{Myc}'\text{SH} / c - \text{Myc}'\text{SH}]$$

$$\text{éq. 9.1(c)} \quad [c - \text{Myc}'\text{SH}] = \sqrt{K_{D,c-\text{Myc}'\text{SH}/c-\text{Myc}'\text{SH}} [c - \text{Myc}'\text{SH} / c - \text{Myc}'\text{SH}]}$$

Et pour [Max'SH] :

$$\text{éq. 9.2(a)} \quad K_{D,\text{Max}'\text{SH}/\text{Max}'\text{SH}} = \frac{[\text{Max}'\text{SH}][\text{Max}'\text{SH}]}{[\text{Max}'\text{SH} / \text{Max}'\text{SH}]} = \frac{[\text{Max}'\text{SH}]^2}{[\text{Max}'\text{SH} / \text{Max}'\text{SH}]}$$

$$\text{éq. 9.2(b)} \quad [\text{Max}'\text{SH}]^2 = K_{D,\text{Max}'\text{SH}/\text{Max}'\text{SH}} [\text{Max}'\text{SH} / \text{Max}'\text{SH}]$$

$$\text{éq. 9.2(c)} \quad [\text{Max}'\text{SH}] = \sqrt{K_{D,\text{Max}'\text{SH}/\text{Max}'\text{SH}} [\text{Max}'\text{SH} / \text{Max}'\text{SH}]}$$

En substituant, dans l'équation du K_D de l'hétérodimère (éq. 8.1), les équivalences obtenues ci-haut des isolements de [c-Myc'SH] (éq. 9.1 (c)) et de [Max'SH] (éq. 9.2 (c)), on obtient l'équation de variables intriquées (éq. 9.3) :

$$K_{D,c-\text{Myc}'\text{SH}/\text{Max}'\text{SH}} = \frac{\sqrt{K_{D,c-\text{Myc}'\text{SH}/c-\text{Myc}'\text{SH}} [c - \text{Myc}'\text{SH} / c - \text{Myc}'\text{SH}]} \cdot \sqrt{K_{D,\text{Max}'\text{SH}/\text{Max}'\text{SH}} [\text{Max}'\text{SH} / \text{Max}'\text{SH}]}{[c - \text{Myc}'\text{SH} / \text{Max}'\text{SH}]}$$

éq. 9.3

Sous cette forme, cette équation ne nous est d'aucune utilité car peu importe les manipulations algébriques que l'on voudrait tenter, elle restera impossible à résoudre sans le recours à une itération. La résolution par itération elle par contre, devient aisée

lorsqu'on divise les membres de droite par la valeur du K_D de c-Myc'SH/Max'SH située à gauche. On retrouve alors l'équation 9.0 qui n'égale « 1 » que « *si et seulement si* » l'état d'équilibre global est satisfait! Il ne reste donc au programme KD_I qu'à itérer les concentrations des espèces par une titration virtuelle comme celle de la simulation par CD_SI.

En répartissant adéquatement les populations transitoires de [c-Myc'SH]_i* et [Max'SH]_i* en leurs formes monomérique et homodimérique selon leurs K_{DS} réels (voir : section 4.3.2. *Retro-calcul des concentrations de monomères et d'homodimères*) le programme raffine la solution exacte jusqu'à atteindre une précision de l'ordre de 10^{-15} M. En fournissant au programme une série de valeurs de concentration de protéine totale initiale [c-Myc'SH]_o et [Max'SH]_o, le programme peut recalculer les concentrations exactes de toutes les espèces qui devraient se trouver en solution pour un jeu donné de K_D réels.

On peut par exemple exporter de CD_SI vers KD_I, les trois valeurs de $K_D(T)$ réel des complexes à 37°C ($K_D(37^\circ\text{C})$), et faire varier seulement la concentration totale de c-Myc'SH pour analyser les effets d'une titration de c-Myc'SH sur le système global à 37°C. On peut également par la suite, à l'aide des intensités spectroscopiques extensives des différentes espèces, remultiplier celles-ci par les concentrations déterminées par KD_I et obtenir une simulation des résultats expérimentaux d'une titration de Max'SH par c-Myc'SH, lorsque analysée par dichroïsme circulaire à 37°C.

Une subtilité doit toutefois être prise en compte en ce qui a trait aux limites physiques initialement posées par KD_I pour l'itération de la concentration de complexe c-Myc'SH/Max'SH. En réalité, lorsque les concentrations des partenaires ne sont pas égale l'une à l'autre, la concentration de complexe c-Myc'SH/Max'SH ne peut être supérieure à la concentration du partenaire limitant. Pour résoudre ce problème et du même coup accélérer l'itération, on développe l'équation quadratique 9.7 (b) qui, résolue (éq. 9.8), permet de déterminer (en absence de toutes compétitions provenant des formes homodimériques) la concentration « initiale maximale » de complexe c-Myc'SH/Max'SH ($[c\text{-Myc'SH/Max'SH}]_o$) en fonction des concentrations totales initiales de c-Myc'SH ($[c\text{-Myc'SH}]_o$) et de Max'SH ($[Max'SH]_o$). Sachant qu'en absence d'équilibres compétiteurs :

$$\text{éq. 9.4} \quad K_{D,c\text{-Myc'SH/Max'SH}}[c\text{-Myc'SH/Max'SH}] = [c\text{-Myc'SH}][Max'SH]$$

Alors si :

$$\text{éq. 9.5} \quad [c\text{-Myc'SH}] = [c\text{-Myc'SH}]_o - [c\text{-Myc'SH/Max'SH}]$$

Et que :

$$\text{éq. 9.6} \quad [Max'SH] = [Max'SH]_o - [c\text{-Myc'SH/Max'SH}]$$

On pose alors que :

$$K_{D,c-Myc'SH/Max'SH}[c-Myc'SH/Max'SH] = ([c-Myc'SH]_o - [c-Myc'SH/Max'SH])([Max'SH]_o - [c-Myc'SH/Max'SH])$$

éq. 9.7 (a)

Et donc :

$$0 = [c-Myc'SH/Max'SH]^2 - ([c-Myc'SH]_o + [Max'SH]_o + K_{D,c-Myc'SH/Max'SH})[c-Myc'SH/Max'SH] + [c-Myc'SH]_o[Max'SH]_o$$

éq. 9.7 (b)

D'où l'on extrait la première racine quadratique (-) pour obtenir la concentration maximale initiale de c-Myc'SH/Max'SH, soit :

$$[c-Myc'SH/Max'SH]_o =$$

$$\frac{[c-Myc'SH]_o + [Max'SH]_o + K_{D,c-Myc'SH/Max'SH} - \sqrt{([c-Myc'SH]_o + [Max'SH]_o + K_{D,c-Myc'SH/Max'SH})^2 - 4[c-Myc'SH]_o[Max'SH]_o}}{2}$$

éq. 9.8

En déterminant de la sorte la concentration réelle maximum de complexe c-Myc'SH/Max'SH, le programme restreint l'étendue de la plage des concentrations à itérer et contourne les aberrations de calcul engendrant des concentrations négatives sous les racines carrées.

4.4.2. Équilibre de K_D et populations

Les populations de complexe homo et hétéro ne sont pas soumises aux mêmes lois de dépendance à la concentration des protéines libres. Par définition, les complexes homos sont formés de partenaires de même nature alors que les hétéros sont formés de partenaires distincts. Ainsi, pour des concentrations de protéine totale et des K_D identiques, les deux différents équilibres ne généreront pas des populations identiques de complexe homo et hétéro. Ceci est dû au fait qu'un complexe homo est dépendant d'une seule concentration de protéine mais selon le carré de celle-ci, alors qu'un complexe hétéro est dépendant de deux concentrations distinctes de protéine. Un des effets encourus est que : des concentrations de « protéine totale » égales à la valeur du K_D fourniront toujours un plus grand nombre de complexe homo qu'hétéro (voir Figure 37).

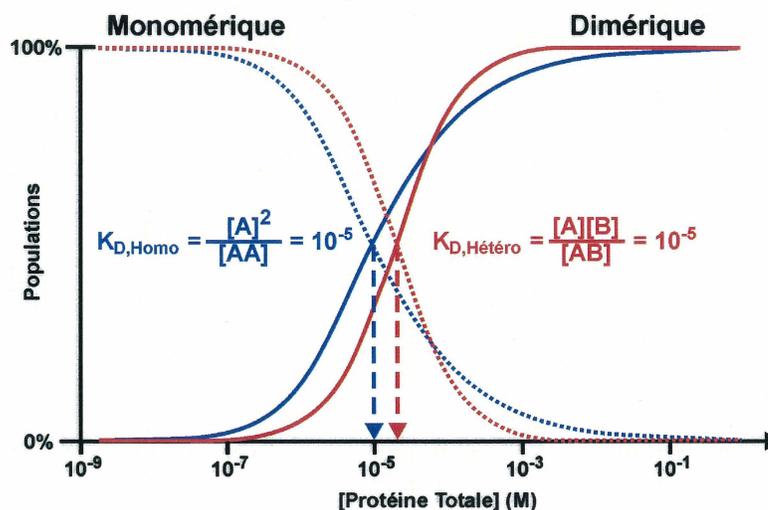


Figure 37 : Distributions de population engendrées par des K_D s homo et hétéro

Par contre, lorsque la concentration de protéine totale dépasse les valeurs de K_D , l'équilibre hétéro gagne du terrain et dépasse l'équilibre homo. Ainsi, lorsque les deux équilibres compétitionnent pour une même ressource, comme c'est le cas ici, les équilibres prendront avantages l'un sur l'autre selon la nature propre de leurs constituants et les concentrations de ces derniers. Pour des K_{DS} identiques à faible concentration, l'homodimère aura toujours une longueur d'avance pour s'accaparer les protéines libres plus fortement que l'hétérodimère. On ne peut donc plus parler de « *comparaisons de concentrations et de K_{DS}* » lorsqu'il s'agit d'évaluer l'affinité de complexes dissemblables de par leur nature, car leurs ordres réactionnels diffèrent. Finalement, on doit aussi prendre en compte que la partie n'est jamais jouée à « *armes égales* ». En réalité, il est physiquement impossible de réaliser (en solution) des conditions dans lesquelles l'homodimère et l'hétérodimère bénéficieront de quantités comparables de réactifs, puisque si l'on fournit 1 mole de protéine totale « Max » pour l'équilibre Max/Max, on se doit d'ajouter 1 mole de « c-Myc » pour l'équilibre c-Myc/Max, soit en fait 2 moles de protéine totale pour celui-ci.

Tous ces jeux de concentration de partenaire et de nature des K_{DS} rendent l'analyse des populations très inintuitive sans un robuste système de simulation des équilibres. Nous présenterons donc ici, pour le bénéfice du lecteur, quelques simulations modèles qui illustrent bien l'influence des différents paramètres « *concentrations* » et « *K_{DS}* » sur les équilibres de population de notre système c-Myc, Max. Dans un but de concision, nous n'aborderons ici que les combinaisons de facteurs les plus cruciales.

4.4.3. Influences des concentrations et des K_{Ds} homo et hétéro

La première particularité que l'on observe des simulations de l'équilibre global intriqué, est l'incidence des concentrations respectives de c-Myc et Max sur les populations relatives d'homodimère Max/Max (voir Figure 38 A) et d'hétérodimère c-Myc/Max (voir Figure 38 B). Les résultats de la simulation mettent en évidence l'effondrement de la population relative de complexe Max/Max (Figure 38 A) suite à une augmentation de la concentration de protéine c-Myc.

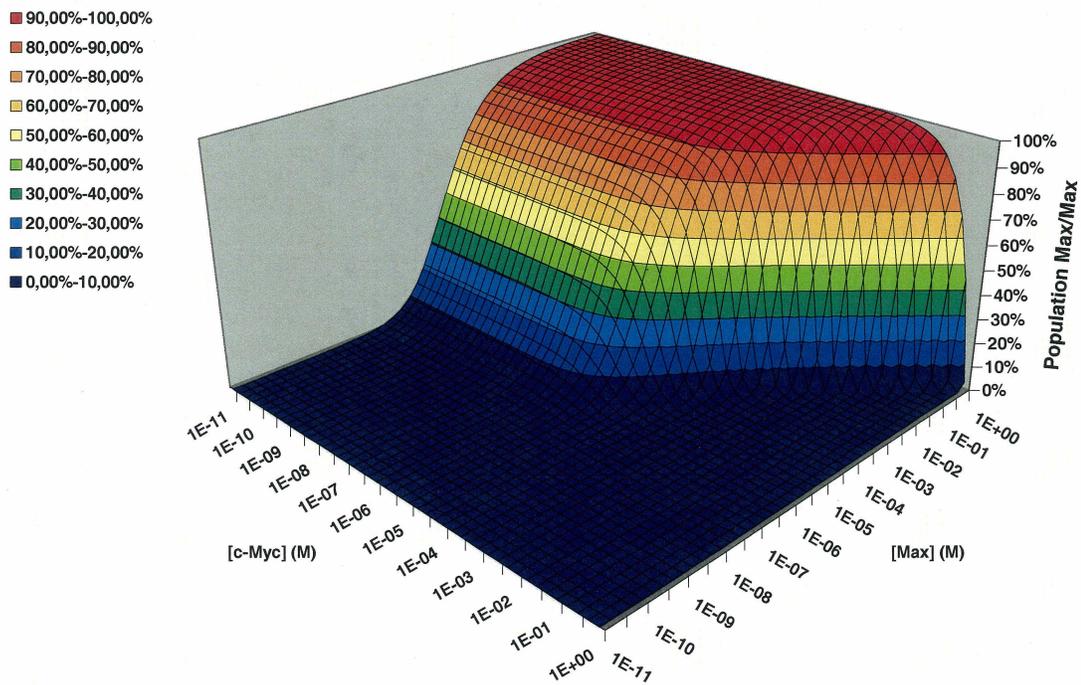


Figure 38 A) : Variations de population du complexe Max/Max en fonction des concentrations de c-Myc et Max Équilibre compétitif $K_{D, \text{Max/Max}} = 10^{-5}$; $K_{D, \text{c-Myc/c-Myc}} = 1$; $K_{D, \text{c-Myc/Max}} = 10^{-5}$.

Il s'agit ici de l'effet indirect du recrutement de partenaire Max par une population de protéine c-Myc libre dans le but de satisfaire l'équilibre global et former des hétérodimères. Aux concentrations négligeables de c-Myc, l'effet s'estompe et laisse entrevoir l'influence classique de la concentration de protéine Max sur sa propre population de complexe homodimère. Dans le cas des populations relatives d'hétérodimère (Figure 38 B), c'est la synergie des effets combinés des concentrations de c-Myc et Max qui retiendra notre attention.

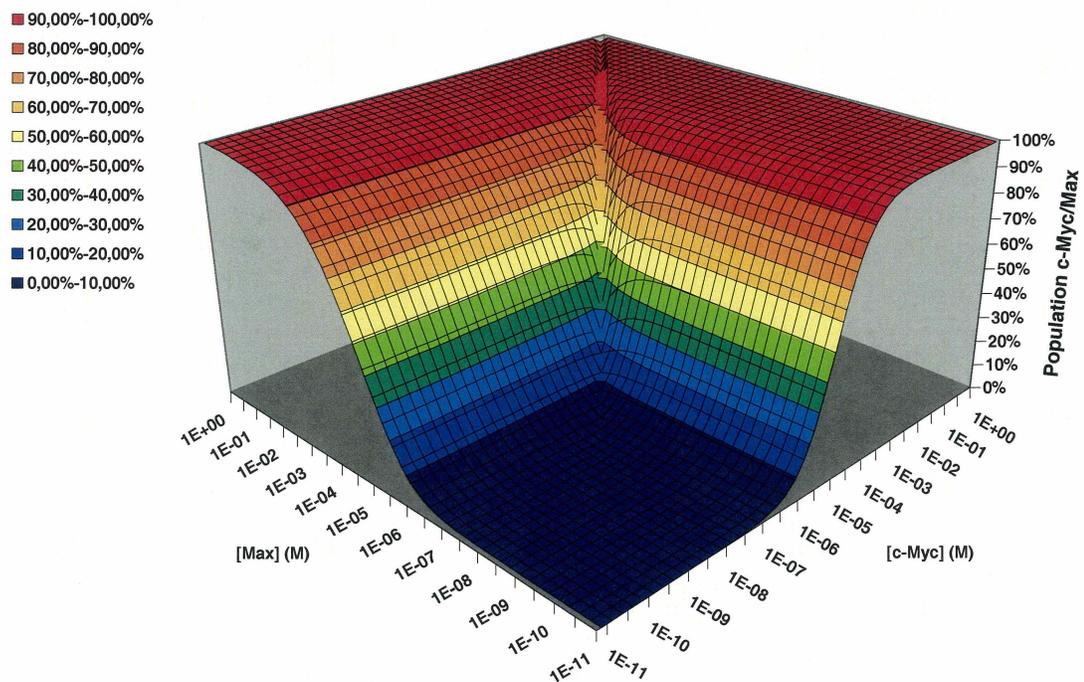


Figure 38 B) : Variations de population du complexe c-Myc/Max en fonction des concentrations de c-Myc et Max Équilibre compétitif $K_{D, \text{Max/Max}} = 10^{-5}$; $K_{D, \text{c-Myc/c-Myc}} = 1$; $K_{D, \text{c-Myc/Max}} = 10^{-5}$.

Bien que les conséquences de la réduction de l'un ou l'autre des partenaires soit facilement compréhensibles, les effets combinés d'une réduction simultanée des concentrations de deux partenaires ne sont pas aisément appréhendables d'autant plus que le système présente une plus grande susceptibilité aux fluctuations de concentration de Max que de c-Myc. Ce dernier point est dû au fait que le système simulé considère une affinité négligeable de c-Myc pour lui-même (c.-à-d. : $K_{D, c-Myc/c-Myc} = 1$). En réalité, dans le cas où « les deux partenaires » présentent des affinités homodimériques non-négligeables en plus d'exhiber une affinité hétérodimérique, seule une simulation rigoureuse parviendrait à éclaircir les populations respectives des différentes espèces en solution.

Comme nous venons de le voir, les fluctuations des concentrations de partenaires influencent les populations des complexes homos et hétéros de façon somme toute assez traditionnelle. Dans le but d'explicitier les phénomènes moins traditionnels de l'équilibre global intriqué, examinons donc à présent les implications des valeurs de K_D homo et hétéro parallèlement aux fluctuations de concentrations de protéine totale (c-Myc et Max équimolaire) sur les populations relatives des deux complexes. Pour éviter d'alourdir cet exposé, nous n'aborderons pas les conséquences de tous les paramètres les uns par rapport aux autres, mais seulement les simulations de cas simplifiés où seul l'un des partenaires (*ici Max*) présente une réelle affinité homo (c.-à-d. : $K_{D, Max/Max} = 10^{-5}$).

Examinons d'abord les effets sur les populations relatives d'hétérodimère, des valeurs du $K_{D, c-Myc/Max}$ (Figure 39 A) et du $K_{D, Max/Max}$ (Figure 39 B). Ces deux premiers cas

présentent des phénomènes similaires, mais engendrés par deux causes totalement distinctes, soit une fluctuation de la stabilité de l'hétérodimère (Figure 39 A) et une fluctuation de la stabilité de l'homodimère (Figure 39 B). Dans un premier temps, on observe que lorsque les affinités de l'hétérodimère et de l'homodimère Max/Max s'équivalent, une population majoritaire d'hétérodimère est formée. Ceci est dû au fait que l'équilibre c-Myc/Max bénéficie toujours de deux fois plus de protéine que l'équilibre Max/Max. Par contre, à forte concentration de protéine et pour des affinités de c-Myc/Max inférieures à celles de Max/Max, on perçoit qu'une domination de l'équilibre Max/Max mine la population d'hétérodimère et dévie l'équilibre global.

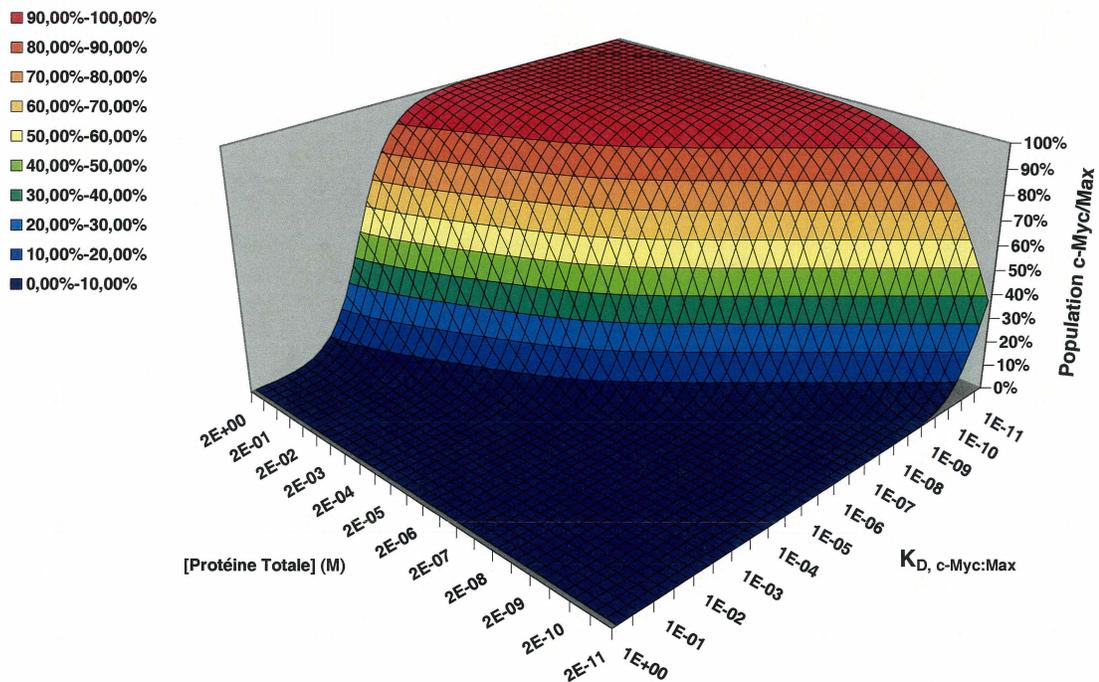


Figure 39 A) : Variations de population du complexe c-Myc/Max en fonction de la concentration de protéine totale et du $K_{D, c-Myc/Max}$ Équilibre compétitif $K_{D, Max/Max} = 10^{-5}$; $K_{D, c-Myc/c-Myc} = 1$.

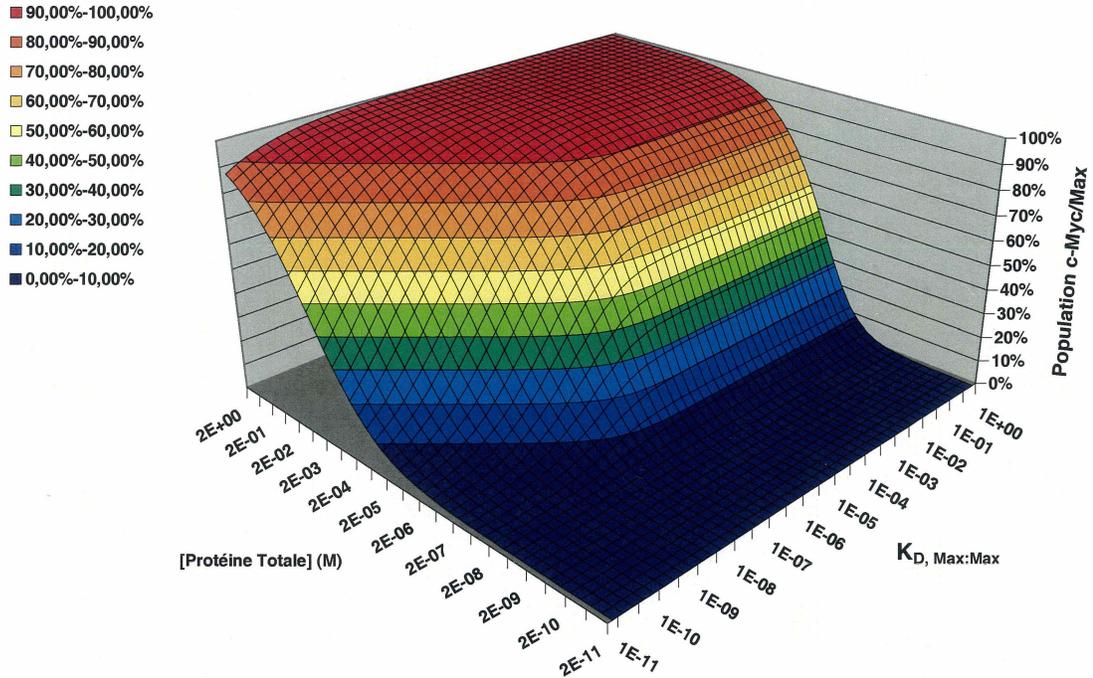


Figure 39 B) : Variations de population du complexe c-Myc/Max en fonction de la concentration de protéine totale et du $K_{D, Max/Max}$ Équilibre compétitif $K_{D, c-Myc/Max} = 10^{-7}$; $K_{D, c-Myc/c-Myc} = 1$.

Les observations précédentes nous ont pour l'instant permis de présenter l'équilibre global intriqué du point de vue de l'hétérodimère c-Myc/Max. Elles font principalement ressortir que pour des K_{D} s similaires et des concentrations de protéine équimolaires, la formation d'hétérodimère l'emporte sur la formation d'homodimère par inégalité des concentrations de réactif. Dans un contexte physiologique, ceci permet de faire ressortir les avantages de l'utilisation du complexe hétérodimère dans l'activation de mécanismes cellulaires. Comparativement au complexe homodimère, l'hétérodimère augmente sa population de « un complexe » pour chaque protéine c-Myc alors que l'homodimère requiert deux Max pour en faire autant. Dans cet ordre d'idée, nous avons pu observer

que, tel que prédit par la théorie, l'équilibre homo est bel et bien asservi par la concentration de protéine c-Myc *via* l'équilibre compétiteur c-Myc/Max.

Terminons cet exposé avec le point de vue de l'homodimère Max/Max afin de comprendre toute la subtilité qui se cache au sein d'un équilibre global intriqué, et qui influence les complexes dans leurs rôles réciproques d'effecteurs cellulaires. Les figures 40 A) et 40 B) présentent les effets sur les populations d'homodimère Max/Max des valeurs de K_D homo (Figure 40 B) et hétéro (Figure 40 A) en fonction des concentrations de protéine totale.

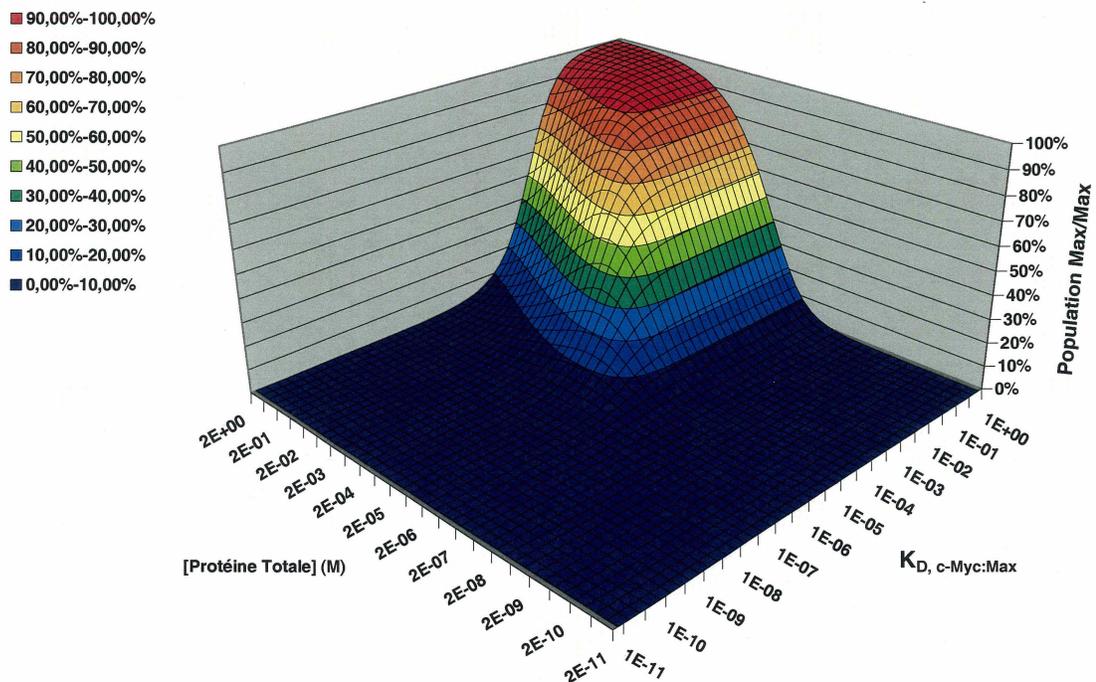


Figure 40 A) : Variations de population du complexe Max/Max en fonction de la concentration de protéine totale et du $K_{D, c-Myc:Max}$ Équilibre compétitif $K_{D, Max/Max} = 10^{-5}$; $K_{D, c-Myc/c-Myc} = 1$.

Dans les deux cas, on peut observer pour la première fois un phénomène tout à fait unique, soit une augmentation « transitoire » de la population de complexe homo proportionnelle à une diminution de la concentration de protéine totale. Ce phénomène est appréhendable de par les allures compétitives des variations de population des complexes homo et hétéro en fonction des concentrations de protéine totale (Figure 37, début chapitre).

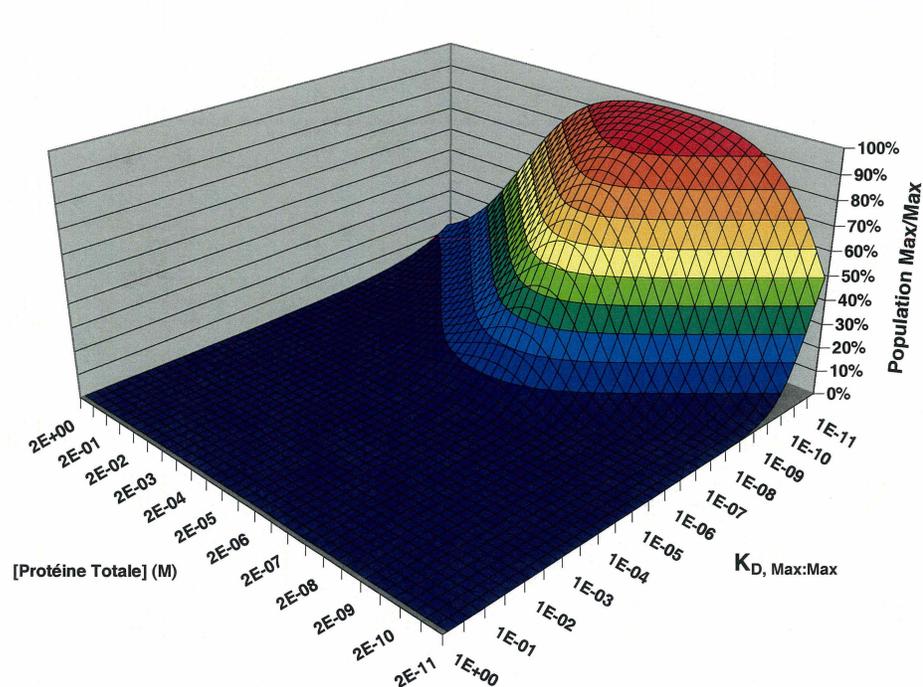


Figure 40 B) : Variations de population du complexe Max/Max en fonction de la concentration de protéine totale et du $K_{D, \text{Max/Max}}$ Équilibre compétitif $K_{D, \text{c-Myc/Max}} = 10^{-7}$; $K_{D, \text{c-Myc/c-Myc}} = 1$.

Il s'agit d'une domination de l'équilibre homo sur l'équilibre hétéro, dans les dernières limites de concentration à peine supérieures au K_D de Max/Max. Ce phénomène (moins

évident sur la Figure 40 A) puisque celle-ci ne réfère qu'à une seule valeur de K_D , $K_{D_{Max/Max}}$) est principalement gouverné par la différence existant entre les valeurs de K_D homo et hétéro. Plus l'affinité homo est grande par rapport à l'affinité hétéro, plus le système réagira fortement dans la transition complexes hétéros et complexe homos. Cependant, à outrance, le complexe hétéro est tout simplement annihilé et l'effet de génération de complexe homo est quasi imperceptible.

À notre avis, il ne fait nul doute qu'il s'agit ici d'un événement crucial pour le complexe Max/Max dans son rôle cellulaire de partenaire pivot obligatoire de c-Myc et Mad. Dans cette optique, une brève diminution des niveaux de protéine c-Myc et Max conjointement à une augmentation de la protéine Mad, suffirait pour compléter de façon élégante, le relargage et la dégradation de c-Myc tandis que la recentralisation de l'homodimère Max serait détournée pour une reprise d'activité avec Mad. Dans le scénario complémentaire, c-Myc prend la place de Mad et les partenaires s'échangent de nouveau, dans une même valse de reconnaissance moléculaire.

Nous voici rendu à l'étape de présenter l'article scientifique « *REVERSIBLE DIMERIZATION BETWEEN HUMAN C-MYC AND MAX B-HLH-LZ DOMAINS* » rédigé au cours de cette maîtrise, et soumis à la revue « *Journal of Molecular Biology* ».

5. CONTRIBUTION DES AUTEURS À L'ARTICLE

Les clonages des gènes encodant les domaines b-HLH-LZ des protéines c-Myc et Max ont été réalisés par M. Montagne et J.F. Naud. J'ai personnellement réalisé la transformation des différentes souches bactériennes de même que j'ai procédé à tous les essais d'induction de la protéine c-Myc. J'ai réalisé l'ensemble de la purification de toutes les protéines et complexes (c-Myc et Max) utilisées dans cet article. J'ai également produit et fait l'analyse par électrophorèse et dichroïsme circulaire de tous les échantillons. Seuls les échantillons de retardement sur gel (R. Lebel) et l'échantillon de résonance magnétique nucléaire de Max (S. Sauvé) ont été entièrement préparés et analysés par mes confrères. Les échantillons RMN de c-Myc et des complexes c-Myc et Max ont été analysés en collaboration avec S. Sauvé et P. Lavigne.

J'ai procédé seul à l'élaboration de l'électrophorèse « HHKOH » permettant la migration électrophorétique des protéines et complexes c-Myc et Max par leurs charges intrinsèques. Finalement, j'ai personnellement dérivés l'ensemble des équations thermodynamiques utilisés dans cet article en m'appuyant sur les travaux de mes collègues : F. Gagnon et J.F. Naud. J'ai effectué seul la programmation de la totalité des logicielles, CD_F, CD_SI et KD_I, et ai réalisé les simulations et effectué le traitement et l'analyse des données en collaboration avec P. Lavigne.

Les figures de cet article ont été préparées en collaboration avec J.F. Naud, S. Sauvé, R. Lebel et P. Lavigne. Les modélisations moléculaires sont de J.F. Naud et P. Lavigne.

**REVERSIBLE DIMERIZATION BETWEEN HUMAN
C-MYC AND MAX B-HLH-LZ DOMAINS**

Résumé

L'hétérodimérisation spécifiques ou reconnaissance moléculaire entre c-Myc et Max est un prérequis *sine qua non* des fonctions de c-Myc dans la croissance cellulaire, la prolifération et la tumorigénèse. Les domaines b-HLH-LZ C-terminaux des deux protéines spécifient cette hétérodimérisation. Lorsqu'en complexe, c-Myc/Max lient les séquences E-box dans les promoteurs de gènes cibles pour en activer la transcription et s'associent également avec d'autres facteurs de transcription, tel Miz-1, pour réprimer la transcription. Il est également rapporté que l'homodimère Max engendre des effets de répression sur la transcription des gènes cibles de c-Myc. Cependant, les déterminants contrôlant la reconnaissance moléculaire entre les b-HLH-LZ, et l'oligomère résultant (dimère vs tétramère) sont encore disputés. Nous rapportons ici la caractérisation de la reconnaissance moléculaire réversible et le repliement spécifique de l'hétérodimère c-Myc/Max en solution et en absence d'ADN. Nous montrons que la reconnaissance moléculaire entre c-Myc et Max est optimale à pH 4.8, ce qui confirme l'existence d'un pont salin stable et crucial entre des glutamates et une histidine à l'interface de dimérisation des domaines LZs. Nous démontrons à l'aide d'évidences par RMN que l'hétérodimère de b-HLH-LZ est correctement replié en absence d'ADN et est en équilibre avec l'état homodimère de Max/Max aux concentrations physiologiques. Nous montrons finalement que cet équilibre permet un échange dynamique entre les états c-Myc/Max et Max/Max. Tel que discuté, cet échange peut s'avérer crucial dans la terminaison des activités transcriptionnelles de c-Myc contrôlés par des agents cytostatiques, comme le TGF- β , qui simultanément réprime c-Myc et favorise l'expression des protéines Max et Mad.

Abstract

The specific heterodimerization or molecular recognition between c-Myc and Max is a *sine qua non* prerequisite for Myc's function in cell growth, proliferation and tumorigenesis. The C-terminal b-HLH-LZ domains of both proteins mediate this heterodimerization. Once in complex, c-Myc/Max binds E-box sequences in promoter of target genes to activate transcription and also associates with other transcription factors such as Miz-1 to repress transcription. It has also been reported that the Max homodimer have a repressive effect on the transcription of c-Myc target genes. However, the determinants controlling the molecular recognition between the b-HLH-LZ and the resulting oligomeric state (dimeric vs. tetrameric) are under debate. Here we report, the characterization of reversible molecular recognition and specific folding of the heterodimeric c-Myc/Max b-HLH-LZ in solution and in absence of DNA. We demonstrate that the molecular recognition between c-Myc and Max is optimal at pH 4.8, which confirms the existence of a crucial and stable salt bridge between glutamates and a histidine at the dimerization interface of the LZs. We show NMR evidence that the heterodimeric b-HLH-LZ is correctly folded in absence of DNA and is in equilibrium with the Max/Max homodimeric state at physiological concentration. We show that this equilibrium allows for an efficient switch from a c-Myc/Max to a Max/Max state. As discussed, this shall be crucial in the termination of c-Myc transcriptional activity triggered by cytostatic agents, such as TGF- β , that downregulate c-Myc and upregulate Max and Mad proteins.

Keywords: Molecular Recognition, c-Myc, Max, b-HLH-LZ, Thermodynamics, CD,
NMR

Introduction

The c-Myc protein is involved in many cellular processes including growth and proliferation^{1,2}. c-Myc is also implicated in the tumorigenesis of many human cancers when deregulated³⁻⁵. The diverse cellular functions of c-Myc are linked to its target genes and its ability to both, activate and repress transcription when heterodimerized with Max⁶. This molecular recognition between c-Myc and Max is mediated by their C-terminal b-HLH-LZ domains⁷.

As a heterodimer with Max, c-Myc can bind E-box DNA sequences in target gene promoters to activate transcription⁸. These target genes are implicated in cell growth and proliferation⁹⁻¹¹. Conversely, the c-Myc/Max heterodimer can also associate with other transcription factors such as Miz-1, Sp1 and Smads in promoters containing Inr element and repress transcription¹²⁻¹⁴. Cell cycle inhibitors such as p15, p21 and p27 are examples of genes amongst others that are repressed^{12,14}. The transactivation domain (TAD) of c-Myc, which contains conserved sequences (Myc-Box I and II), have been shown to recruit protein complexes with HAT activity. It also directly interact with Smad and Sp1 proteins, establishing the Myc-Box I and II as crucial regions for the regulation of c-Myc's transactivation and transrepression functions^{15,16}. However the biological activity of the TAD also depends on the heterodimerization between the b-HLH-LZ of c-Myc and Max^{6,14,17}. The heterodimeric b-HLH-LZ domain is required for the recruitment of c-Myc to the E-Box elements. Furthermore, Miz-1 interacting domain has been located in the HLH of c-Myc¹⁴. In other words, in absence of the b-HLH-LZ of

Max for heterodimerization, c-Myc is unfolded and cannot bind neither DNA nor Miz-1^{9,18}. Finally, the Max homodimer can have a repressive effect on the activation of the transcription and cell proliferation¹⁹. However, the structural determinants responsible for the reversible molecular recognition between the b-HLH-LZs of c-Myc and Max, the solution structure in absence of DNA and oligomeric state²⁰ of the heterodimeric b-HLH-LZ remain to be clarified¹⁸. Elucidation of these points is crucial for a structural and mechanistic understanding of the regulation of genes by this family of transcription factors²⁰⁻²².

Structural determinants important for the molecular recognition between Max and c-Myc have been demonstrated to be located in the LZ of c-Myc and Max²⁰⁻²³. A stable and buried salt bridge between His 81d on Max and Glu 410a and 417a on c-Myc was shown to play a crucial role in the specific heterodimerization^{21,23}. On the other hand, Glu 410a and 417a were proposed to destabilize the homodimeric state of the c-Myc LZ and to be relieved by Max's His 81d upon the formation of the stable macroscopic salt-bridge. Alternately, the His 81d was shown to destabilize: the Max homodimeric LZ²¹, b-HLHLZ^{24,25} and complete gene product²⁶. Glu 410a and 417a can be viewed as relieving the repulsions at the Max homodimeric LZ interface^{21,23}. Interestingly, all proteins known to heterodimerize with Max have a conserved acidic side chain in their LZ at the position corresponding to c-Myc Glu 410a. For example, on Mad1, this position is occupied by an Asp. We have recently shown that the b-HLH-LZ of Mad1 does not homodimerize at neutral pH. However, the simple mutation of this Asp to an Asn was sufficient to trigger the formation of a stable homodimer capable of DNA

binding under physiological conditions²⁷. Altogether, these data point at a crucial role for the acidic side-chain/His interplay for both destabilization in homodimer conformation and stabilization as heterodimer. However, in a recent study, it was suggested that the specific heterodimerization between the c-Myc and Max b-HLH-LZs (hereafter called c-Myc' and Max', respectively) relied on the interplay between a layer involving c-Myc Arg 423g and 424a and Gln 91g and Asn 92a on Max²⁰. In this view, the destabilization of the c-Myc homodimer by the Arg layer is relieved by a surface exposed H-bond between Arg 424a and Gln 91g to allow specific dimerization. In the same study, the X-ray structure of the C-terminal cross-linked heterodimer of c-Myc' and Max' (hereafter called c-Myc'SSMax') bound to an E-Box sequence has been determined. It has been observed that c-Myc'SSMax' bound to DNA (c-Myc'SSMax'/E-Box), packs as a head-to-tail dimer of heterodimers in the unit cell. Analytical ultra-centrifugation studies of the c-Myc'SSMax' free of DNA suggest that this construct sediments as a dissociable species with the molecular weight of a tetramer. The K_D of the latter was determined to be $90 \cdot 10^{-9}$ at 20°C and its organization has been ascribed to the tetramer observed in the unit cell of the crystalline form. Considering the fact that the nuclear concentration of c-Myc is estimated to lie in the range of 15 to 76 μ M in fibroblasts in culture and since the K_D of the cross-linked heterodimer is below the estimated nuclear concentrations, it has been suggested that the c-Myc/Max heterodimer functions *in vivo* as a head-to-tail dimer of heterodimers. However, one has to consider that the cross-link might induce the oligomerization. Furthermore, one has to address the possibility that in absence of a cross-link, other species than the heterodimer (e.g. Max'SH/Max'SH) can be populated. Since Max can have an antagonistic effect on

c-Myc¹⁹, the demonstration of the existence of a population of Max'SH/Max'SH in equilibrium with c-Myc'SH/Max'SH bears a profound mechanistic relevance. Finally, although others have characterized the heterodimerization of the b-HLH-LZ of viral²⁸ or murine Myc²⁹ with Max, no study reporting the detailed description of the heterodimerization between b-HLH-LZ of human Max and c-Myc has been published yet.

In this study, we reveal the molecular recognition between human c-Myc'SH and Max'SH in absence of DNA and demonstrate that this molecular recognition is optimal at pH 4.8. This confirms the role of the macroscopic salt bridge between Glu 410a and 417a of c-Myc and His 81d of Max in the molecular recognition of the b-HLH-LZ. We also show by solution state NMR that the dissociable heterodimeric b-HLH-LZ in absence of DNA is folded similarly to the homodimeric Max b-HLH-LZ³⁰. Under the conditions where the cross-linked c-Myc'SSMax' is reported to form a strong tetramer, we conversely show that the dissociable c-Myc'SH/Max'SH does not form a heterotetramer and is rather in equilibrium with monomeric (c-Myc'SH and Max'SH) and homodimeric (Max'SH/Max'SH) species. Using a novel numerical routine, we have determined the equilibrium concentration of all species and the K_D of c-Myc'SH/Max'SH between 5 and 95°C. The $K_D(20^\circ\text{C})$ and $K_D(37.5^\circ\text{C})$ were found to be $5.14 \cdot 10^{-7}$ and $3.53 \cdot 10^{-6}$, respectively. As shown and discussed, the reversible heterodimerization allows for an efficient concentration dependent switch from a c-Myc'SH/Max'SH heterodimer state to a Max'SH/Max'SH state. The biological implications of this switch are discussed in the context of the cytostatic effect of

hormones such as TGF- β ^{31,32} that induce an orchestrated downregulation of c-Myc and upregulation of the Max and Mad proteins.

Results and Discussion

We show in Figure 1, the alignment of the primary sequences of Max'SH and c-Myc'SH. Figure 2a-c depict the helical wheel diagrams of the homodimeric (a and b) and heterodimeric (c) LZs, describing unfavorable (solid arrows) and favorable (dashed arrows) interhelical interactions, that have been proposed^{21,22} to be important for the molecular recognition between c-Myc and Max. Upon inspection of the Max homodimeric and c-Myc/Max b-HLH-LZ structures, it appears that no specific interaction resides at the interface of the HLH domains. However, it is not clear which interaction(s) or molecular determinant(s) at the interface of the LZ domains play(s) a prominent role in the molecular recognition between the b-HLH-LZ. We provide in the following sections the characterization of the molecular recognition and its prominent molecular determinants.

Molecular recognition between c-Myc'SH and Max'SH

We present in Figure 3a (red) the CD spectrum of c-Myc'SH recorded at pH 6.8, 20°C and a concentration of 32 μ M in monomer units. The shallow minimum at 222 nm and stronger minimum observed around 205 nm are indicative of an unfolded protein with a low helical content. This is compatible with the notion that at neutral pH, in absence of Max, c-Myc'SH exists in an unfolded state with residual helical content. This is also in accordance with the well-recognized facts that neither the LZ²¹ nor the complete gene product¹⁸ is able to homodimerize at neutral pH. Moreover, we have noted the poor

solubility of the construct, especially at pH 6.8. This is intriguing since the pI of the construct is close to 10. This suggests that insoluble states of c-Myc'SH exist at neutral pH. In order to further characterize the tertiary and quaternary structure of c-Myc'SH, we have performed a temperature-induced denaturation under the same conditions (Figure 3b, red). The denaturation depicts a very shallow transition superimposed onto a monotonic increase in CD signal. This is suggestive of a gradual denaturation of residual secondary and quaternary structures. In contrast, the CD spectrum obtained for Max'SH (Figure 3a, blue) at 20°C is indicative of a stable α -helical dimer with the minima at 208 nm and 222 nm. The ellipticity at 222 nm corresponds to *circa* 80 % of the helical content expected for a dimeric b-HLH-LZ with the b-region unfolded²⁵. The temperature-induced denaturation also reveals a cooperative transition as expected for a stable dimeric b-HLH-LZ (Figure 3b, blue). The thermodynamical characterization of the dissociable Max b-HLH-LZ homodimer has been described in detail elsewhere and has revealed a $K_D(20^\circ\text{C})$ of $1.02 \cdot 10^{-6}$ which corresponds to a population of dimeric Max'SH of roughly 80 % at 20°C²⁵.

In order to expose the molecular recognition event, we have also recorded the CD spectra of an equimolar mix of c-Myc'SH and Max'SH both at 32 μM of monomer, for a maximum possible heterodimer concentration of 32 μM (Figure 3a, purple). As exemplified by Vinson and collaborators, if specific heterodimerization occurs, the CD spectrum of the mixture will depict more helical structure than the theoretical sum of the spectra of both monomers (Figure 3a, grey)²⁹. As judged by the difference between the experimental spectra of the mixture and the theoretical sum, an important amount of

secondary structure (α -helix) is induced when c-Myc'SH and Max'SH are mixed together. The ellipticity at 222 nm observed is -70 mdeg. It is almost exactly twice that observed for the spectrum of Max'SH/Max'SH (-34 mdeg), which was recorded at half the concentration). This suggests that the majority (*circa* 80 %), but not all of c-Myc'SH forms a specific and stable quaternary structure with Max'SH and that Max'SH/Max'SH and unfolded species of both b-HLH-LZs are in equilibrium at room temperature. In order to further investigate the molecular recognition, we have recorded the temperature denaturation of the same sample (Figure 3b, purple). Again, the larger absolute value in ellipticity, but furthermore the apparent T_m of the mixture compared to the theoretical counterpart, indicates the formation of additional secondary structure part of a new quaternary complex with a larger apparent stability. This is the hallmark of molecular recognition and specific heterodimerization.

The c-Myc'SH/Max'SH complex is dimeric

It has been observed that the cross-linked version (C-terminal of the LZ) of the heterodimeric b-HLH-LZ (c-Myc'SSMax') bound to DNA packs as a head-to-tail dimer of dimer in the crystalline form²⁰. In absence of DNA, this cross-linked dimer apparently sediments as a dissociable specie with the molecular weight of a tetramer. The reported K_D for this tetramer is $90 \cdot 10^{-9}$ at 20°C. This has led to the proposal that this tetrameric state is the functional form of the heterodimeric b-HLH-LZ in the nucleus of non-transformed and proliferative cells. In order to ascertain the oligomeric state of the non-cross-linked c-Myc'SH/Max'SH, we have proceeded with an

electrophoretic mobility shift assay (EMSA) with the premise that if the c-Myc'SH and Max'SH associates as a strong heterotetramer, it will form a stable DNA complex that should shift an E-box probe with the apparent MW of a tetramer. On the other hand, if c-Myc'SH/Max'SH is a heterodimer, it should modify the electrophoretic mobility of an E-box with an apparent molecular weight similar to Max'SH/Max'SH, which is a *bona fide* dimer (Figure 4a, lane 3). Note that the c-Myc'SSMax'/E-Box and c-Myc'SH/Max'SH/E-Box complexes (lane 4 and 5) have a slightly higher apparent MW than the Max'SH/Max'SH/E-Box complex. One can also notice that the cross-linked species shift a larger amount of probe (Lanes 2 and 4). This is expected since cross-linking the dimer will have a stabilizing effect on both the protein and the DNA complex²⁵. Is the slightly higher apparent molecular weight of the heterodimeric complex corresponds to a tetrameric molecular weight or is it due to an intrinsic difference on the overall charge (and electrophoretic mobility) of the homodimeric and heterodimeric complexes? In this regard, it has to be noted that c-Myc'SH has a larger net positive charge than Max'SH ($Z_{c-Myc'SH(pH\ 7.00)} = 7.15$ vs. $Z_{Max'SH(pH\ 7.00)} = 4.51$). As a consequence, the net negative charge of the c-Myc'SH/Max'SH/E-box and c-Myc'SSMax'/E-Box are expected to be smaller than that of the Max'SH/Max'SH/E-box and consequently should have a lesser electrophoretic mobility. We validate this assertion on Figure 5 where the electrophoretic mobility of c-Myc'SSc-Myc' and Max'SSMax' in urea and in absence of SDS are displayed. Note that the electrophoresis has been run in a reverse electric field or a cationic migration. One can see that c-Myc'SSc-Myc' has a larger electrophoretic mobility than that of Max'SSMax' as expected from the higher density of positive charges on c-Myc'SH. Therefore, since the

overall shapes of all the dimeric complexes are expected to be the same, one can rationalize the lesser electrophoretic mobility of the heterodimeric DNA complexes on the basis of a lesser negative charge density rather than on a MW or oligomeric state difference. In order to further prove this point we have measured the electrophoretic mobility of a mandatory c-Myc'SSc-Myc'/Max'SSMax' tetrameric DNA complex (Figure 4, lane 6). The assembly and structure are described in Figure 4b and c respectively. On its own, c-Myc'SSc-Myc' displays very little helical content (Figure 4b, red). However, as displayed in Figure 4b, when mixed with Max'SSMax' (blue), c-Myc'SSc-Myc' forms a heteromeric complex (purple) with an ellipticity corresponding to approximately 80% of a fully helical tetrameric b-HLH-LZ with the b-region unfolded. Here again, this suggests that the mandatory heterotetramer is in equilibrium with Max'SSMax' and c-Myc'SSc-Myc'. On Figure 4a, lane 6, we can distinguish two bands, one with the molecular weight of Max'SSMax'/E-box complex (lane 2) and the other at a higher molecular weight that corresponds to c-Myc'SSc-Myc'/Max'SSMax'/E-Box. Taken altogether these results demonstrate that in our conditions, the c-Myc'SH/Max'SH/E-box and c-Myc'SSMax'/E-Box complexes have molecular weights of dimeric rather than tetrameric complexes.

c-Myc'/Max' has a similar fold than the Max'/Max' homodimeric b-HLH-LZ in absence of DNA.

The NMR solution structure of the c-Myc/Max heterodimeric LZ was previously solved²³. The heterodimeric LZ was shown to adopt the classical parallel double-

stranded α -helical coiled-coil fold. However, the characterization of the structure of the heterodimeric HLH in absence of DNA remains to be done. Indeed, the only structure of a b-HLH or b-HLH-LZ in absence of DNA known to date is the structure of the homodimeric b-HLH-LZ of Max³⁰. This structure reveals that in absence of DNA the helices in the HLH are arranged similarly than in the structure DNA complexes with the exception that the b-region is mostly unfolded and the loop experiences significant motions on the ns and ps time-scale. In order to provide insights into the structure of the heterodimeric HLH domains in absence of DNA, we have characterized the non-covalent ¹³C,¹⁵N-c-Myc'SH/¹²C,¹⁴N-Max'SH complex by multi-nuclear and multi-dimensional NMR spectroscopy.

The unfolded nature of c-Myc'SH is further demonstrated in Figure 6a (red) where the ¹H-¹⁵N-HSQC of ¹³C,¹⁵N-c-Myc'SH at pH 6.8 and 37°C is shown. The poor dispersion of the correlations (cross peaks) is the signature of a near homogeneous magnetic environment for all ¹H-¹⁵N spin-pairs as expected in mostly unfolded protein. In opposition, we present the ¹H-¹⁵N-HSQC of an equimolar mixture of ¹³C,¹⁵N-c-Myc'SH and ¹²C,¹⁴N-Max'SH at pH 6.8, 37°C (Figure 6a, purple). One can observe a single set of well dispersed ¹H-¹⁵N correlations. This confirms that the majority of c-Myc'SH is involved in a single and stably folded quaternary complex under these conditions. Subsequently, we have recorded a ¹H-¹³C-HSQC of the ¹³C,¹⁵N-c-Myc'SH alone and in presence of ¹²C,¹⁴N-Max'SH. The rationale of running these experiments is the following: The ¹H of the δ -methyl groups of the Leu 44 side-chains in the HLH domain of the Max homodimer are significantly upfield shifted (chemical shifts <0.4 ppm)³⁰.

This is caused by a ring current effect from the Phe 43 side-chains. Indeed both Leu and Phe side-chains form a cluster where the Leu of one monomer packs on the Phe ring of the other monomer and *vice versa* at the very interior of the hydrophobic core HLH bundle³². In the HLH of c-Myc, a Leu (396) and a Phe (374) are also conserved at the corresponding positions and the same cluster is observed in the crystal structure of the c-Myc'SSMax'/DNA complex²¹. Therefore, if the heterodimeric HLH-LZ is folded likewise in absence of DNA, the side-chain Leu 64 on Max'SH should pack against the ring of Phe 374 on c-Myc'SH and the side-chain of Leu 396 on c-Myc'SH should pack against the ring of Phe 43 on Max'SH. Consequently the ¹H-¹³C-HSQC of the ¹³C,¹⁵N-c-Myc'SH in presence of unlabeled Max'SH should display two upfield-shifted ¹H δ -methyls resonances. Inspection of the structure of the c-Myc'SSMax'/DNA complex²¹ reveals that no other aromatic ring is potentially involved in stable tertiary or quaternary interactions with methyl groups. We present on Figure 6b the ¹H-¹³C-HSQC of ¹³C,¹⁵N-c-Myc'SH alone (red) and in presence of unlabeled Max'SH (purple). One can see two significantly upfield-shifted ¹H resonances, supporting our hypothesis that the two δ -methyls of Leu 396 of c-Myc'SH are indeed packed against Phe 43 of Max'SH (Figure 6c). In this experiment, the two δ -methyls of Leu 64 of Max'SH are invisible. In order to show that Leu 64 of Max'SH also packs against the Phe 374 of c-Myc'SH, we have recorded the aromatic ¹³C-edited NOESY-HSQC of the same sample. In this experiment only the NOEs involving ¹H attached to aromatic ¹³C are observed. Therefore, if the two δ -methyls of Leu 64 of Max'SH are in close proximity to Phe 374 of c-Myc'SH, NOE's between the δ^1 H of Leu 64 should appear in ¹³C aromatic planes. We present, on Figure 6d (upper panel) a Phe aromatic ¹³C plane (132 ppm) where NOEs between aromatic

Phe side-chain ^1H and δ -methyl ^1H are observed. In order to prove that these ^1H resonances belong to Leu 64, we display the upfield region of the 1D spectrum (Figure 6d, lower panel). Since no ^{13}C decoupling was applied, the ^1H resonances of δ -methyl ^1H of Leu 396 of c-Myc'SH are split while those of Leu 64, which are not coupled, have a larger apparent peak height. As observed in Figure 6d, the latter have chemical shifts matching those in the aromatic ^{13}C -edited NOESY-HSQC plane. The intramonomer NOEs from the ^1H of Phe 43 ring to the δ -methyl ^1H Leu 396 of c-Myc'SH are only observed at very low contour levels, indicating that they are at a longer distance than the inter-monomer Phe/Leu pair. This is also the case in the solution structure of the homodimeric b-HLH-LZ of Max³⁰. In summary, the NMR data presented here demonstrates that, in solution and in absence of DNA, c-Myc'SH/Max'SH adopt a dimeric b-HLH-LZ conformation similar to that of the Max homodimeric b-HLH-LZ.

Structural determinants responsible for the molecular recognition

It is accepted that the molecular recognition involves the destabilization of the c-Myc homodimer to ensure that monomeric state is available for heterodimerization with a transient monomeric Max²⁰⁻²³. While the residues responsible for the reversible nature of Max dimerization have been unraveled²⁴⁻²⁶, the residues responsible for the destabilization of the c-Myc homodimer and those involved in the specific interaction(s) leading to a stable heterodimer are under debate. Indeed two (non-exclusive) sets of side-chains located in the LZ domains of c-Myc and Max have been proposed to dictate specificity of heterodimerization. It has been observed that the LZ domain of c-Myc is

unable to form a homodimer. Two pairs of Glu side-chains (409g, 410a and 416g, 417a) located at the interface of the LZ (Figure 2a) have been proposed to be responsible for the destabilization of the homodimeric LZ^{21,23}. The specific heterodimerization between the LZ of c-Myc and Max has been demonstrated^{21,22} and the existence of a specific salt-bridge between His 81d on Max and Glu 410a and 417a on c-Myc in the solution structure of the heterodimeric LZ confirmed²³. This interaction is also observed to occur in the crystal structure of the DNA bound heterodimeric b-HLH-LZ (Figure 7a, b). On the other hand, it was suggested that consecutive interfacial Arg side-chains on c-Myc, Arg 423g and 424a (Figure 2a), were responsible for the destabilization of its homodimer²⁰. This destabilization is proposed to be relieved by a surface exposed H-bond between Arg 424a and Gln 91g of Max to allow specific dimerization (Figure 7a, c). This interaction is observed in one of the heterodimer in the unit cell of the c-Myc'SSMax'/DNA crystal structure²⁰ and in members of the ensemble of solution NMR structures of the heterodimeric LZ²³. A question now arises. Do both sets of residues contribute equally or is there one that plays a more prominent role?

In order to answer this question, we have studied the molecular recognition between c-Myc'SH and Max'SH as a function of pH (Figure 8a, b, c and d) at 20°C. We have chosen this temperature in order to reproduce, with Max'SH, conditions as close as possible to those that prevail *in vivo* and where p21 Max can homodimerize²⁶ and compete with the c-Myc/p21 Max heterodimer. If the molecular recognition relies only on the interplay between Arg 423g and 424a on c-Myc'SH and Gln 91g on Max'SH, no pH dependence should be observed between pH 2.8 and 6.8 because the protonation

state of the Arg side-chains will not change in this pH range. However, if the Glu 410a and 417a / His 81d cooperation plays a critical role, then molecular recognition can be predicted to be optimal at a pH between the pK_a of the His (*circa* 6.8) and the Glu's (*circa* 4.1). Indeed at such a pH, the fully protonated interfacial His81d destabilizes the Max homodimer²⁵ and Glu 410a and 417a will still be charged and destabilizing the c-Myc'SH homodimer. This shall optimize the probability of interaction between c-Myc'SH and Max'SH and consequently the formation of the heterodimer. In the heterodimer, the charged His 81d will be stabilized by Glu 410a and 417a (*vice versa*) through the formation of the stable salt-bridge and provide the stabilization free energy for molecular recognition. Conversely at lower and higher pH (e.g. 2.8 and 6.8), one should start to minimize molecular recognition through the population of c-Myc homodimer and Max homodimer *via* protonation of Glu's and deprotonation of the His, respectively. Moreover, the lack of stable interfacial salt-bridge at the interface of the heterodimeric species will also contribute to a possible reduction of the molecular recognition.

We present in Figure 8a, the Far-UV CD spectra of c-Myc'SH, Max'SH and c-Myc'SH+Max'SH at 20°C and pH 2.8. We also present on Figure 8b the corresponding temperature denaturations. One can see that the helical content of c-Myc'SH is increased compared to pH 6.8 (Figure 3a, red), indicating the stabilization of secondary structure most probably by removing medium and long range repulsions between acidic side-chains in secondary and quaternary structures. Conversely and as reported before²⁵, the helical content of Max'SH and stability are decrease at pH 2.8 due to unfavorable

quaternary interactions caused by the charged His 81d. More importantly and in contrast to pH 6.8 (Figure 3a), the arithmetic sum of the spectra of c-Myc'SH and Max'SH (Figure 8a, grey) is only slightly above the experimental spectrum of the c-Myc'SH+Max'SH mix (purple) indicating, as suspected, a lack of significant molecular recognition. The lack of molecular recognition is also further demonstrated by the perfect superimposition of the temperature denaturation of the mixture (purple) and the arithmetic sum (grey) above 20°C (Figure 8b). The matching Far-UV CD spectra and denaturations at pH 4.8 are shown in Figure 8c and d, respectively. As seen, the ellipticity of both c-Myc'SH and Max'SH are similar to those at pH 2.8. But, as expected, the ellipticity of the c-Myc'SH+Max'SH is lower (higher helical content) than the one observed at pH 6.8. This demonstrates optimal molecular recognition. Alternately, we present on Figure 8e, the Far-UV CD spectra of c-Myc'SH+Max'SH at pH 2.8 (red), 4.8 (purple) and 6.8 (blue). It can be clearly noted that the helical content is optimal at pH 4.8, between the pK_as of Glu 410a/417a and His 81d. This lends solid credence to the existence of a more prominent role for the Glu 409g, 410a and Glu 416g, 417a/His 81d cooperation than the interplay between Arg 423g and 424a on c-Myc'SH and Gln 91g on Max'SH. This can be further rationalized as follows: It has been shown that after Asp, Glu is the second most destabilizing residue at position a in LZ and three times more destabilizing than Arg³³. This implies that the two Glu g-a pairs (409g/410a and 416g/417a) should play a more prominent role than the single Arg 423g/424a g-a pair (Figure 2a). The repulsions between Arg 423g/424a at the interface are likely relieved at the interface of the heterodimer with the absence of neighboring positive charges and where the Arg side-chains can be solvated by water and/or by the

formation of an H-bond between Arg 424a c-Myc and Gln 91g Max. On the other hand, one has to remind the microscopic nature of this interaction compared to the buried salt bridge that has a macroscopic character. Microscopic and solvent exposed interactions can either be made or broken in the folded state of proteins as evidenced from the NMR structure of the heterodimeric LZ of c-Myc/Max (*vide supra*). In opposition, macroscopic interactions like buried and stable salt bridges generally contribute more to the stabilization free energy of the folded state of proteins because they are linked to folding^{34,35}. Taken together, the data presented in this section advocates for a predominant role for c-Myc Glu 410a and 417a and Max His 81d residues in the molecular recognition between c-Myc'SH and Max'SH.

Equilibrium populations and K_D of the c-Myc'SH/Max'SH heterodimer

So far, we have observed that at room temperature the helical content of an equimolar mix between c-Myc'SH and Max'SH is only 80% of what is expected for fully folded b-HLH-LZ in absence of DNA. Therefore it appears that unfolded species are populated. In fact, the dissociation of c-Myc'SH/Max'SH and consequent accumulation of free and unfolded c-Myc'SH is expected to be promoted by the formation of Max'SH/Max'SH (Figure 9). Since the concentrations of c-Myc and Max vary in different cellular contexts (e.g. proliferation vs. differentiation) and that the Max/Max and c-Myc/Max can have opposite effects on transcription, the characterization of the equilibrium populations of all species (and shifts in equilibrium population as a function of concentration), is of significant and mechanistic relevance to c-Myc transcriptional

activity. In this context, we have developed a numerical routine allowing for the determination of the equilibrium concentration (populations) of all species from the temperature denaturation of a c-Myc'SH+Max'SH equimolar mix monitored by CD (see Materials and Methods).

We present in Figure 10a the simulated temperature denaturation (black line) of the c-Myc'SH+Max'SH equimolar mix at a total protein concentration of 64 μ M, along with the contributions of each species to the total $CD_{222nm}(T)$ signal. The concentrations of all species as a function of temperature are presented in Figure 10b. As anticipated, below 37°C and around 20°C, the concentration Max'SH/Max'SH (dark blue) and c-Myc'SH (light red) are significant and the population of folded (helical) species is around 80%. In addition, note that the concentration of c-Myc'SH is larger than that of Max'SH (light blue) due to the lack of population of c-Myc'SH/c-Myc'SH (dark red).

The determination of the equilibrium concentrations of all species allows one to evaluate the $K_{D,c-Myc'SH/Max'SH}$ (equation 1) even in presence of concurrent equilibriums (i.e. $K_{D,c-Myc'SH/c-Myc'SH}$ and $K_{D,Max'SH/Max'SH}$). This would have been otherwise impossible.

equation 1

$$K_{D,c-Myc'SH/Max'SH} = \frac{[c-Myc'SH] \cdot [Max'SH]}{[c-Myc'SH/Max'SH]}$$

In addition, Figure 10b allows one to determine the apparent K_D of c-Myc'SH/Max'SH ($K_{D,c-Myc'SH/Max'SH}^*$). The $K_{D,c-Myc'SH/Max'SH}^*$ is given by equation 2:

equation 2

$$K_{D,c-Myc'SH/Max'SH}^* = \frac{([c-Myc]_T - 2[c-Myc'SH/c-Myc'SH] - [c-Myc'SH/Max'SH]) \cdot ([Max]_T - 2[Max'SH/Max'SH] - [c-Myc'SH/Max'SH])}{[c-Myc'SH/Max'SH]}$$

where $[c-Myc]_T$ and $[Max]_T$ stand for the total concentrations of c-Myc and Max (monomeric units), respectively. Note that in absence of concurrent equilibrium (e.g. above 37.5°C), the two K_D s become equivalent. The $K_{D,c-Myc'SH/Max'SH}^*$ (20°C) and (37.5°C) we determined are $2.27 \cdot 10^{-6}$ and $3.62 \cdot 10^{-6}$ respectively while the $K_{D,c-Myc'SH/Max'SH}$ (20°C) and (37.5°C) are $5.14 \cdot 10^{-7}$ and $3.53 \cdot 10^{-6}$ respectively (Table 1). Note the difference between the two K_D s at 20°C. This difference becomes negligible at 37.5°C.

As sketched in Figure 9, mixing c-Myc'SH and Max'SH at a total protein concentration of 64 μ M allows for the significant population of Max'SH/Max'SH, which in response leads to the accumulation of c-Myc'SH. It can therefore be anticipated that decreasing the total concentration will lead to an increase in the relative populations of Max'SH/Max'SH and c-Myc'SH. Using the K_D values listed in Table 1, we have simulated the population of all species as a function of the total concentration (Figure 11a; see Materials and Methods for details). One can see that lowering the total concentration from 64 to 16 μ M should lead to a maximal population of

Max'SH/Max'SH' and to a concomitant accumulation of c-Myc'SH. Using the molar CD_{222nm} signal ($MCD_{222nm}(20^{\circ}C)$) of each species and the population presented in Figure 11a, we have predicted the CD_{222nm} signal (solid line) of equimolar c-Myc'SH+Max'SH mixes from 68 to 6 μ M (Figure 11b). We also present experimental values (open circles). The righteousness of the predicted concentration dependence of the CD_{222nm} signal support the populations of all the different species simulated in Figure 11a. This result is of functional and mechanistic relevance for the regulation and termination of c-Myc/Max transcriptional activities upon down regulation of the concentration of c-Myc following cytostatic signals. Indeed it can be anticipated that the down-regulation of c-Myc will be followed by an efficient shift from a c-Myc/Max to a Max/Max both free and DNA bound (Figure 9). Similarly, it can also be expected that increasing the nuclear concentration of Max above that of c-Myc will be accompanied by an increase of DNA free heterodimer and a concomitant increase in free c-Myc available to be degraded. It is also worth noting the population of free Max'SH available to heterodimerize with the Mad proteins also increases. Interestingly, a downregulation of c-Myc and an upregulation of Max and Mad protein are respectively observed during TGF- β treatment of fibroblasts³¹ and differentiation of intestinal epithelial cells³². Our results provide mechanistic and thermodynamical insights of the chronological steps leading to the termination of c-Myc transcriptional activity.

Conclusion

In this contribution we report the structural and thermodynamical characterization of the molecular recognition between the b-HLH-LZ domain of c-Myc (c-Myc'SH) and Max (Max'SH) in solution. We show using CD and solution state NMR that the dissociable c-Myc'SH/Max'SH heterodimer has stable heterodimeric b-HLH-LZ fold comparable to the Max/Max b-HLH-LZ free of DNA, with a $K_{D,c-Myc'SH/Max'SH}$ of $5.14 \cdot 10^{-7}$ at 20°C and $3.53 \cdot 10^{-6}$ at 37.5°C. We demonstrate that the molecular recognition is optimal at pH 4.8 and describe the prominent role for c-Myc Glu 410a and 417a and Max His 81d in this process. These side-chains form a macroscopic inter-helical salt-bridge at the interface of the c-Myc'SH/Max'SH LZ domains as demonstrated elsewhere²³. This is in disagreement with another suggestion that the molecular recognition between c-Myc'SH and Max'SH relies solely on c-Myc Arg 423g and 424a and Max Gln 91g also located in the LZ domains²⁰. As discussed, due to the microscopic nature of the interaction made by these residues in the heterodimer, they cannot account by themselves for the inability of c-Myc'SH to form a stable homodimer and stabilize the heterodimer. We reveal that when the b-HLH-LZ of c-Myc and Max are free to dissociate they do not form a tight heterotetrameric complex as reported elsewhere for the cross-linked species²⁰. The heterodimeric c-Myc/Max b-HLH-LZ is rather in equilibrium with homodimeric Max/Max b-HLH-LZ and monomeric species at near physiological concentrations (μ M).

Our thermodynamic characterization allows one to predict that the down-regulation of c-Myc observed upon TGF- β treatment³¹ and during differentiation³² will be efficiently followed by a switch from a prevailing E-Box bound c-Myc/Max state to a DNA bound Max/Max state. This can contribute to increase the amount of free c-Myc available to be degraded and lead to the repression of transcription of c-Myc target genes and the activation of the transcription of cytostatic genes such as p21. This shift would not be as efficient if the c-Myc/Max were to exist as a strong heterotetramer in the nucleus. Finally, it is worth noting that Max and subsequently Mad are upregulated by TGF- β ³¹ and during the differentiation of the intestinal epithelium³². All these evidence point to a likely active role of the Max homodimer in the regulation and termination of c-Myc transcriptional activity by promoting the accumulation of monomeric c-Myc and the dissociation of c-Myc/Max. Subsequently once in stoichiometric excess the homodimer could compete with the c-Myc/Max heterodimer for E-box sequence genes to transiently repress transcription before the Mad proteins modifies the chromatin around the promoters target genes and irreversibly shut down c-Myc activated transcription.

Acknowledgments

This research was funded by a grant from the Canadian Institutes of Health Research (CIHR). P.L. is Junior2 scholar of the Fonds de la Recherche en Santé de Québec (FRSQ).

Materials and Methods

Construction of pET3a expression plasmid encoding for the b-HLH-LZ domains of c-Myc and Max

The b-HLH-LZ of c-Myc (residue 353N to residue 436N) was amplified by a polymerase chain reaction (PCR) from pSP72-c-Myc (kindly provided by C. Asselin, department of Anatomy and Cell Biology, Université de Sherbrooke, Quebec, Canada,) which contain the third exon with the 3' untranslated region using the following oligonucleotides 5'-d(CGGAAGGATCCCATATGGTCAAGAGGCGAACACACAACGTC)-3' and 5'-d(CGGGATCCGTTCCGTAGCTGTTCAAGTTTCAACTG)-3', as 5' and 3' primers which contain *NdeI* and *BamHI* restriction sites, respectively, for cloning into the pET3a plasmid (Novagen). No Stop codon was used. This resulted in a construct with C-terminal extension of GSGC, which is named c-Myc'SH throughout the text. The amplification and cloning of Max'SH has been described elsewhere²⁵.

Expression and purification of c-Myc'SH and Max'SH

The pET3a containing c-Myc'SH was transformed in *E. Coli* BL21 Codon + (Novagen). Single colonies were picked and grown overnight as pre-cultures of 10 ml of LB medium or M9 minimal medium for the production of unlabeled or ¹³C, ¹⁵N-uniformly labeled protein³⁰, respectively. Then, 1L of the desired medium was inoculated. The culture was incubated at 37°C until the D.O._{595nm} reached 0.6. Protein expression was

induced overnight at 37°C by adding IPTG (0.6 mM). Cells were harvested by centrifugation at 3 500 g and frozen at -20°C until purification. Bacteria were lysed in 10 ml of c-Myc lysis buffer (50 mM HEPES-NaOH (pH 7.50), 100 mM NaCl, 10 mM MgCl₂, 5 mM DTT, 0.35 mg/ml lysozyme) and incubated 30 minutes at 37°C. Triton X-100 was added to the solution to obtain a final concentration of 1% then, cells were homogenized by sonication on ice. DNAase I (0.04 mg/mL) was added and the solution was incubated at 37°C for 60 minutes. Inclusion bodies were collected by a centrifugation at 30 000 g for 30 minutes at 4°C and resuspended with 15 ml of inclusion bodies buffer (6 M urea, 100 mM sodium acetate (pH 5.00), 0.5 M GuHCl and 25 mM DTT). After complete solubilization, 15 ml of 2 M urea was added to dilute the solution, followed by a centrifugation at 30 000 g for 30 minutes at 4°C). Supernatant was loaded on a HiTrap SP HP column (Amersham) preconditioned with Buffer A (50 mM sodium acetate (pH 5.00)) and B (buffer A with 5 M NaCl). The column was washed with 5 volume of U75 buffer (4 M urea, 50 mM Tris-HCl (pH 7.50)). Elution of c-Myc was performed with a gradient (25 mM/ml) from 100% of A to 100% of B. Max'SH was expressed and purified as described elsewhere²⁵.

Preparation of the cross-linked c-Myc'SSMax' heterodimer

The disulfide-linked heterodimer was prepared by air oxidation of equal molar quantities of c-Myc'SH and Max'SH at pH 8.8 in presence of 30 mM of Cu⁺⁺ (CuSO₄) in a Tris-HCl buffer (100 mM) containing 6M GuHCl. The heterodimer was purified from the aqueous reaction mixture by HPLC on a homemade reverse-phase C12 column

using a gradient of acetonitrile disole in water using tri-fluoroacetic acid as a counter-ion.

Electrophoretic mobility shift assays.

The dsDNA was created by heating oligonucleotides containing E-Box sequence (underline) 5'-d(CCCCCAACACGTGTTGCCTGA)-3' and 5'-d(TCAGGCAACACGTGTTGGGGG)-3' at 90°C and cooling the solution slowly to room temperature. All complexes, 0.5 µM of proteins in monomer units with 0.5 µM of dsDNA, were incubated at room temperature for an hour in 20 mM Tris-HCl (pH 8.0) and 50 mM KCl. To reduced complexes, 25 mM DTT was added. The complexes were loaded on a polyacrylamide gel (6% in 1X TBE; 89 mM Tris, 89 mM boric acid, 1.9 mM EDTA) and run in 0.5 X TBE at 100 V. The gels were stained in ethidium bromide and visualized under UV light to reveal the DNA and DNA complexes.

Cationic PAGE, His-HEPES-KOH

The cationic electrophoresis conditions were adapted from the work of Chrambach and Jovin³⁶. In more details, free base histidine was used as trailing ion and potassium hydroxide as leading ion, in a N-(2-Hydroxyethyl)piperazine-N'-2-ethane sulfonic acid (HEPES) base buffer. We used a 4% acrylamid bis-acrylamid (29:1) staking gel (125 mM HEPES-KOH, pH 7.75, 6 M urea) and a 12% acrylamid bis-acrylamid (29:1) resolving gel (375 mM HEPES-KOH, pH 7.00, 6 M urea). The polymerizations were

catalyzed with ammonium persulfate and N,N,N',N'-Tetramethylethylenediamine (Temed). The Anode Buffer was made of 25 mM HEPES and 175 mM histidine free base and the Cathode Buffer was composed of 200 mM HEPES-KOH (pH 7.00). Loading buffer (62.5 mM HEPES-KOH (pH 7.75), 30% glycerine and 6 M urea) was added to the samples with a ratio of 1:1. Horse heart cytochrome C was used to follow the migration. Electrical cable polarity was inverted and the migration was conducted using constant voltage 200V and 18 mA for about 2 hours.

Circular dichroism spectropolarimetry

Circular dichroism measurements were performed on a Jasco J-810 spectropolarimeter equipped with a Jasco Peltier-type thermostat. The instrument was calibrated with an aqueous solution of d-10- (+)-camphor-sulfonic acid at 290.5 nm. Samples were loaded into quartz cells with a path length of 0.1 cm. Far-UV CD spectra were recorded at the desired temperature from 200 nm to 250 nm by averaging three scans at 0.1 nm intervals. Sample of proteins from the stock solutions were diluted to the desired concentration in 50 mM KH_2PO_4 , 50 mM KCl, 5 mM DTT (pH was adjusted using 10N KOH or 6N HCl). Before measurement, the samples were incubated for 24 hours at room temperature to ensure reduction and equilibrium. Temperature denaturations were performed from 5°C to 95°C with a heating rate of 1°C/min. The unfolding was recorded at 222 nm.

NMR spectroscopy

Both protein sample of $^{13}\text{C},^{15}\text{N}$ -cMyc'SH and $^{13}\text{C},^{15}\text{N}$ -cMyc'SH/ $^{12}\text{C},^{14}\text{N}$ -Max'SH were dissolved in a 50 mM KH_2PO_4 , 50 mM KCl buffer and 50 mM DTT at a concentration of approx. 2.5 mM in dimer units. All NMR experiments were performed at 37°C on a Varian unity INOVA 600 MHz spectrometer equipped with an indirect detection, triple resonance ($^1\text{H}/^{13}\text{C}/^{15}\text{N}$) probe and Z-axis pulsed-field gradient capabilities. Side-chain resonances of ^1H δ -methyl groups were assigned using ^1H - ^{13}C -HSQC, 3D ^{13}C -separated NOESY HSQC and ^{13}C -separated aromatic NOESY-HSQC experiments³⁷. A mixing time of 200 ms was used. All pulse sequences were taken from the Varian ProteinPack library based on the pulse sequences of Lewis Kay and co-workers³⁸. The spectra were references relative to water in the 1H dimension and according to Wishart et al.³⁹ for indirect dimensions. All spectra were processed using NMRPipe⁴⁰ and analyzed with NMRview⁴¹.

Determination of equilibrium concentrations of c-Myc'SH, c-Myc'SH/c-Myc'SH, Max'SH, Max'SH/Max'SH and c-Myc'SH/Max'SH from the temperature denaturation of an equimolar mixture between c-Myc'SH and Max'SH monitored by CD.

The equilibrium concentrations are determined from the temperature denaturation of the c-Myc'SH+Max'SH mixture ($\text{CD}_{222\text{nm},\text{c-Myc'SH+Max'SH}}(\text{T})$) using the temperature-dependent molar $\text{CD}_{222\text{nm}}$ signal ($\text{MCD}_{222\text{nm}}(\text{T})$) of each specie and the $K_D(\text{T})$ of the

homodimers. The $MCD_{222nm}(T)$ of c-Myc'SH/c-Myc'SH, c-Myc'SH, Max'SH/Max'SH and Max'SH are given by the temperature dependent baselines $CD_{222nm}(T)$ (in mdeg) of the native (homodimeric state) and denatured (monomeric state) respectively. The actual baselines used are displayed in Figure 3 (dashed lines). They are simply divided by the concentration (in molar units) in order to obtain the $MCD_{222nm}(T)$ values (mdeg(T)/M). The baselines and the $K_D(T)$ values of homodimers were obtained from the simulation of the corresponding temperature denaturation curves (5.0 to 95.0°C with 0.1° increments), assuming a two-state mechanism as described elsewhere²⁵⁻²⁷ (see also footnotes in Table 1). The $CD_{222nm}(T)$ for a solution of 32 μ M (dimer units) of “pure” c-Myc'SH/Max'SH heterodimer is given by the temperature dependent baseline of the native state of the cross-linked c-Myc'SSMax' heterodimer (Figure 9a, grey curve).

Once all the $MCD_{222nm}(T)$ and $K_D(T)$ values of homodimers are known, the concentration of the different species can be iterated using equation 3 and the following steps. The starting point of all iterations is by default 32 μ M (100%) of c-Myc'SH/Max'SH so the corresponding $CD_{222nm}(T)$ value of the c-Myc'SSMax' heterodimer base-line. Next, [c-Myc'SH/Max'SH] is decreased to 0 in 0.4 μ M intervals assuming that the dissociation is irreversible. For example, decreasing c-Myc'SH/Max'SH concentration from 32 to e.g. 28 μ M releases a non-equilibrium concentration of 4 μ M of free c-Myc'SH ([c-Myc'SH]*) and 4 μ M of free Max'SH ([Max'SH]*).

equation 3

$$\begin{aligned}
 CD_{c\text{-Myc}'SH+Max'SH} &= [c - Myc'SH] \cdot MCD_{c\text{-Myc}'SH} \\
 &+ [c - Myc'SH/c - Myc'SH] \cdot MCD_{c\text{-Myc}'SH/c\text{-Myc}'SH} \\
 &+ [Max'SH] \cdot MCD_{Max'SH} \\
 &+ [Max'SH/Max'SH] \cdot MCD_{Max'SH/Max'SH} \\
 &+ [c - Myc'SH/Max'SH] \cdot MCD_{c\text{-Myc}'SH/Max'SH}
 \end{aligned}$$

Knowing the homodimeric $K_D(T)$ values, new pseudo-equilibrium concentrations of Max'SH, Max'SH/Max'SH, c-Myc'SH, c-Myc'SH/c-Myc'SH can be determined with the constraints that $[Max'SH]^*$ and $[c\text{-Myc}'SH]^*$ become the “total” available concentration (in monomer units) of Max and c-Myc species respectively. Therefore, the pseudo-equilibrium concentration of Max'SH is given by the root of the following quadratic equation:

equation 4

$$0 = 2 \cdot [Max'SH]^2 + K_{D,Max'SH/Max'SH} \cdot [Max'SH] - K_{D,Max'SH/Max'SH} \cdot [Max'SH]^*$$

and the pseudo-equilibrium concentration of Max'SH/Max'SH is given by:

equation 5

$$[Max'SH/Max'SH] = \frac{[Max'SH]^* - [Max'SH]}{2}$$

The pseudo-equilibrium concentrations of c-Myc'SH species are determined likewise and the theoretical $CD_{222nm,c-Myc'SH+Max'SH}(T)$ value is then calculated with equation 3. The two [c-Myc'SH/Max'SH] values that gave $CD_{222nm,c-Myc'SH+Max'SH}(T)$ above and below the experimental value respectively become the upper and lower limits for the next step of minimisation. The new interval of concentrations is given by the difference between these upper and lower limits of [c-Myc'SH/Max'SH] divided by 4. This process is continued until convergence, i.e. when the [c-Myc'SH/Max'SH] interval becomes less than 1 fM (10^{-15} M).

Simulation of the population of all species as a function of the total protein concentration.

Knowing the $K_{D,c-Myc'SH/Max'SH}(T)$ values, equilibrium populations of all different species can be simulated as a function of total protein concentration, using the above routine to numerically solve equation 6:

equation 6

$$1 = \frac{\sqrt{K_{D,c-Myc'SH/c-Myc'SH}} \cdot [c - Myc'SH/c - Myc'SH] \cdot \sqrt{K_{D,Max'SH/Max'SH}} \cdot [Max'SH/Max'SH]}{K_{D,c-Myc'SH/Max'SH} \cdot [c - Myc'SH/Max'SH]}$$

Equation 6 can be derived recognizing that

$$[Max'SH] = \sqrt{K_{D,Max'SH/Max'SH}} \cdot [Max'SH/Max'SH] \quad \text{and}$$

$$[c - Myc'SH] = \sqrt{K_{D,c-Myc'SH/c-Myc'SH}} \cdot [c - Myc'SH/c - Myc'SH], \text{ and replacing } [Max'SH] \text{ and}$$

$$[c - Myc'SH] \text{ in } K_{D,c-Myc'SH/Max'SH} = \frac{[c - Myc'SH] \cdot [Max'SH]}{[c - Myc'SH/Max'SH]} \text{ to obtain equation 7:}$$

equation 7

$$K_{D,c-Myc'SH/Max'SH} = \frac{\sqrt{K_{D,c-Myc'SH/c-Myc'SH}} \cdot [c - Myc'SH/c - Myc'SH] \cdot \sqrt{K_{D,Max'SH/Max'SH}} \cdot [Max'SH/Max'SH]}{[c - Myc'SH/Max'SH]}$$

Equation 7 can be rearranged into equation 6.

Molecular modeling

All the molecular modeling was done using the InsightII suite (Acceleris) running on an Octane SGI workstation. The protein and DNA coordinates of Myn:E-Box complex (murine Max) were used (PDB accession no. **1AN2**) for the models in Figure 4 and 9. The coordinates of the HLH from the c-Myc'SSMax'/E-box complex (PDB accession no. **1NKP**) were used for the model in Figure 6. The molecular rendering has been done using the programs Ribbons⁴² (figures 4, 7 and 9) and PyMOL⁴³ (figure 6).

References

- 1- Eilers, M., Schirm, S. and Bishop, J. M. (1991). The Myc protein activates the transcription of the α -prothymosin gene. *EMBO J.* **10**, 133-142.
- 2- Henriksson M. and Lüsher B. (1996). Proteins of the Myc network: essential regulators of cell growth and differentiation. *Advan. Cancer Res.* **68**, 109-182.
- 3- Baudino, T.A., McKay, C., Pendeville-Samain, H., Nilsson, J. A., Maclean, K.H., White, E. L., Davis, A.C., Ihle, J.N. and Cleveland, J. L. (2002). c-Myc is essential for vasculogenesis and angiogenesis during development and tumor progression. *Genes Dev.* **16**, 2530-2543.
- 4- Pelengaris, S., Kahn, M. and Evan, G.I. (2002). Suppression of Myc-induced apoptosis in β cells exposes multiple oncogenic properties of Myc and triggers carcinogenic progression. *Cell* **109**, 321-334.
- 5- Felsher, D. W. and Bishop, J.M. (1999). Transient excess of Myc activity can elicit genomic instability and tumorigenesis. *Proc. Natl. Acad. Sci. USA* **96**, 3940-3844.
- 6- Blackwood, E. M., Lüsher, B. and Eisenman, R. N. (1992). Myc and Max associate in vivo. *Cell* **6**, 71-80.

- 7- Blackwood, E. M. and Eisenman, R. N. (1991). Max : a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Nature* **251**, 1211-1217.
- 8- Grandori, C., Cowley, S. M., James, L. P. and Eisenman, R.N. (2000). The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu. Rev. Cell Dev. Biol.* **16**, 653-99.
- 9- Li, Z., Calcar, S.V., Qu, C., Cavenee, W.K., Zhang, M.Q. and Ren. B. (2003). A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. *Proc. Natl. Acad. Sci. USA* **100**, 8164-8169.
- 10- Waters, C.M., Littlewood, T.D., Hancock, D.C., Moore, J.P. and Evan, G.I. (1991). C-Myc protein expression in untransformed fibroblasts. *Oncogene* **6**, 797-805.
- 11- Orian, A., van Steensel, B., Delrow, J., Bussemaker, H.J., Li, L., Sawado, T., Williams, E., Loo, L.W.M., Cowley, S.M., Yost, C., Pierce, S., Edgar, B.A., Parkhurst, S.M. and Eisenman, R.N. (2003) Genomic binding by the *Drosophila* Myc, Max, Mad/Mnt transcription factor network. *Genes Dev.* **17**, 1101-1114.

- 12- Brenner, C., Deplus, R., Didelot, C., Lorient, A., Viré, E. De Smet, C., Gutierrez, A. Danovi, D., Bernard, D., Boon, T. , Pelicci, P. G., Amati, B., Kouzarides, T., de Launoit, Y. Di Croce, L., and Fuks, F. (2005). Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J.* **24**, 336-346.
- 13- Seoane, J., Pouponnot, C., Staller, P., Schader, M., Eilers, M. and Massague, J. (2001). TGF- β influences Myc, Miz-1 and Smad to control the CDK inhibitor p15^{INK4b}. *Nature Cell Biology* **3**, 400-408.
- 14- Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Moroy, T., Bartek, J., Massague, J., Hanel, F. and Eilers, M. (2001). Repression of p15^{INK4b} expression by Myc through association with Miz-1. *Nature Cell Biol.* **3**, 392-399.
- 15- Bahram, F., von der Lehr, Cetinkaya, C. and Larsson, L.G. (2000). C-Myc hotspot mutations in lymphomas result in inefficient ubiquitination and decrease in proteasome-mediated degradation. *Blood* **95**, 2104-2110.
- 16- Stone, J., de Lange, T., Ramsay, G., Jakobovits, E., Bishop, J.M., Varmus, H, and Lee, W. (1987). Definition of regions in human c-myc that are involved in transformation and nuclear localization. *Mol. Cell Biol.* **7**, 1697-1709.

- 17- Mao, D., Watson, J., Yan, P., Barsyte-Lovejoy, D., Khosravi, F., Wong, W., Farnham, P., Huang, T., Penn, L. (2003). Analysis of Myc Bound Loci Identified by CpG Island Arrays Shows that Max Is Essential for Myc-Dependent Repression. *Curr. Biol.* **13**, 882-886.
- 18- Adhikary, S. and Eilers, M. (2005). Transcriptional regulation and transformation by Myc proteins. *Nature Reviews. Molecular Cell Biology* **6**, 635-645.
- 19- Gu, W., Cechova K., Tassi, V. and Dalla-Favera, R. (1993). Opposite regulation of gene transcription and cell proliferation by c-Myc and Max. *Proc. Natl. Acad. Sci. USA* **90**, 2935-2939.
- 20- Nair, S.K. and Burley, S.K. (2003). X-ray structures of Myc-Max and Mad-Max recognizing DNA: Molecular bases of regulation by proto-oncogenic transcription factors. *Cell* **112**, 193-205.
- 21- Lavigne, P., Kondejewski, L. H., Houston, M. E. Jr., Sönnichsen, F.D., Lix, B. Sykes, B.D., Hodges, R.S. and Kay, C.M. (1995). Preferential heterodimeric parallel coiled-coil formation by synthetic Max and c-Myc Leucine zippers: A description of putative electrostatic interactions responsible for the specificity of heterodimerization. *J. Mol. Biol.* **254**, 505-520.

- 22- Muhle-Goll, C., Nilges, M., and Pastore, A. (1995). The leucine zippers of the HLH-LZ proteins Max and c-Myc preferentially form heterodimers. *Biochemistry* **34**, 13554-13564.
- 23- Lavigne, P., Crump, M. P., Gagné, S. M., Hodges, R. S., Kay, C. M. and Sykes, B. D. (1998). Insights into the mechanism of heterodimerization from the ¹H-NMR solution structure of the c-Myc-Max heterodimeric Leucine Zipper. *J. Mol. Biol.* **281**, 165-181.
- 24- Tchan M.C. and Weiss A.S. (2001). Asn78 and His81 form a destabilizing locus within the Max HLH-LZ homodimer. *FEBS Lett* **509**, 177-180.
- 25- Naud, J. F., Gagnon, F., Wellinger, R., Chabot, B. and Lavigne, P. (2003). Improving the Stability of the LZ domain of Max increases the stability of its B-HLH-LZ:E-Box complex. *J. Mol. Biol.* **326**, 1577-1595.
- 26- Naud, J.F., McDuff, F.-O., Sauvé, S., Montagne, M., Webb, B. A., Smith, S. P., Chabot, B. and Lavigne, P. (2005). Structural and thermodynamical characterization of the complete p21 gene product of Max. *Biochemistry* **44**, 12746-12758.

- 27- Montagne, M., Naud, J.-F., McDuff, F.-O and Lavigne, P. (2005). Towards the elucidation of the structural determinants responsible for the molecular recognition between Mad1 and Max. *Biochemistry* **44**, 12860-12869.
- 28- Fieber, W., Schneider, M. L., Matt, T., Krautler, B., Konrat, R., and Bister, K. (2001). Structure, function, and dynamics of the dimerization and DNA-binding domain of oncogenic transcription factor v-Myc. *J. Mol. Biol.* **307**, 1395-410.
- 29- Krylov, D., Kasai, K., Echlin, D. R., Taparowsky, E. J., Arnheiter, H. and Vinson, C. (1997). A general method to design dominant negatives to B-HLHZip proteins that abolish DNA binding. *Proc. Natl. Acad. Sci. USA* **94**, 12274-12279.
- 30- Sauv e, S., Tremblay, L. and Lavigne, P. (2004). The NMR Solution Structure of a Mutant of the Max b/HLH/LZ Free of DNA: Insights into the Specific and Reversible DNA Binding Mechanism of Dimeric Transcription Factors. *J. Mol. Biol.* **342**, 813-832.
- 31- Siegel, P.M., Shu, W. and Massague, J. (2003). Mad upregulation and Id2 repression accompany transforming growth factor (TGF-beta)-mediated epithelial cell growth suppression. *J. Biol. Chem.* **278**, 35444-35450.

- 32- Mariadason, J.M., Courtney, N.M., L'Italien, K.E., Zhuang, M., Smartt, H.J.M., Heerdt, B.G., Yang, W., Corner, G.A., Wilson, A.J., Arango, D., Augenlicht, L.H. (2005). Gene expression profiling of intestinal epithelial cell maturation along the crypt-villus axis. *Gastroenterology* **128**, 1081-1088.
- 33- Wagschal, K., Tripet, C. B., Lavigne, P., Mant, C. M. and Hodges, R. S. (1999). The role of position a in determining the stability and oligomerization state of α -helical coiled-coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Sci.* **8**, 2312-2329.
- 34- Lavigne, P., Sönnichsen, F. D., Kay, C. M. and Hodges, R. S. (1996). Interhelical salt bridges, coiled-coil stability, and specificity of dimerization. *Science* **271**, 1136-1138.
- 35- Sharp, K.A. and Englander S.W. (1994). How much is a stabilizing bond worth ? *Trends Biochem. Sci.* **19**, 526-529.
- 36- Chrambach, A., Jovin, T.M. (1983). Selected buffer systems for moving boundary electrophoresis on gels at various pH values, presented in a simplified manner. *Electrophoresis* **4**, 190-204.
- 37- Clore, G.M. & Gronenborn, A.M. (1998). Determining structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* **16**, 22-34.

- 38- Kay, L.E. (1995). Field gradient techniques in NMR spectroscopy. *Curr. Opin. Struct. Biol.* **5**, 674-681.
- 39- Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. & Sykes, B.D. (1995). ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135-140.
- 40- Delaglio, F., Grzesiek, S., Vuister, G., W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293.
- 41- Johnson, B.A. and Blevins, R.A. (1994). NMRview: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603-614.
- 42- Carson, M. (1987). Ribbon Models of Macromolecules. *J. Mol. Graph.* **5**, 103-106.
- 43- DeLano, W.L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, San Carlos, CA, USA.

Figure captions

Figure 1. Sequence alignment of the human c-Myc (c-Myc'SH) and human Max (Max'SH) b-HLH-LZ. Note the C-terminal GSGC extensions.

Figure 2. Interhelical electrostatic interactions potentially involved in the molecular recognition between c-Myc'SH and Max'SH. Helical wheel diagrams showing (a) Interfacial repulsions (solid arrows) between two pairs of Glu side-chains (409g, 410a, 416g, 417a) and one pair of Arg side-chains (423g and 424a) in the LZ domains of a putative c-Myc'SH/c-Myc'SH dimer. (b) Interfacial repulsion between potentially charged His side-chains (81d) at the interface of the Max'SH homodimeric LZ. (c) Specific salt-bridge between Max His 81d and c-Myc Glu 410a and 417a^{19,22} and charged H-Bond between c-Myc Arg 424a and Max Gln 91g²¹ (dashed arrows). They are both proposed to be key interactions for specific heterodimerization.

Figure 3. Molecular recognition between c-Myc'SH and Max'SH in absence of DNA. (a) Far-UV CD spectra (20°C) of c-Myc'SH (red), Max'SH (blue), the arithmetic sum of both spectra (grey) and the c-Myc'SH+Max'SH mix (purple). The concentrations of c-Myc'SH and Max'SH were 32 μM in monomeric units and the total concentration of the mixture was 32 μM in dimer units. (b) Temperature denaturation curves of c-Myc'SH (red), Max'SH (blue), their arithmetic sum (grey) and the temperature denaturation curve of the c-Myc'SH+Max'SH mix (purple). Shown in dashed lines are the baselines for the unfolded (upper lines: c-Myc'SH (light red) and Max'SH (light blue)) and folded

(lower lines: c-Myc'SH/c-Myc'SH (dark red) and Max'SH/Max'SH (dark blue)) states. See footnote 1 of Table 1 for further detail on the baseline of c-Myc'SH/c-Myc'SH.

Figure 4. Confirmation of the dimeric state of c-Myc'SH/Max'SH by EMSA (a). Lane 1. Free double-stranded E-Box DNA. Lane 2. Max'SSMax'/E-Box complex. Lane 3. Max'SH/Max'SH/E-Box complex. Lane 4. c-Myc'SSMax'/E-Box complex. Lane 5. c-Myc'SH/Max'SH/E-Box complex. Lane 6 shows two DNA complexes resulting from the incubation of equimolar quantities of c-Myc'SSc-Myc' and Max'SSMax' in presence of double-stranded E-Box DNA. All complexes were formed at a total protein concentration of 0.5 μ M of protein with a molar excess (100%) of double-stranded E-Box DNA. (b) Far UV-CD spectra of c-Myc'SSc-Myc' (red,), Max'SSMax' (blue) and c-Myc'SSc-Myc'+ Max'SSMax' (purple) recorded at pH 6.8 and 20°C at 32, 32 and 64 μ M in total concentration (monomeric units) respectively. (c) Sketch describing the formation of the obligatory c-Myc'SSc-Myc'/Max'SSMax' heterotetramer and the c-Myc'SSc-Myc'/Max'SSMax'/E-Box complex. Both the c-Myc'SSc-Myc'/Max'SSMax' and Max'SSMax' and their corresponding DNA complexes are formed while c-Myc'SSc-Myc' remains unfolded. c-Myc and Max are rendered in red and blue respectively.

Figure 5. Electrophoretic mobility of c-Myc'SSc-Myc' and Max'SSMax' at different concentrations measured by cationic PAGE (in absence of SDS) in presence of 6 M urea at pH 7.00. The electrophoresis was performed in reverse polarity or in cationic migration. c-Myc'SSc-Myc' is observed to have a larger electrophoretic mobility than

Max'SSMax' by virtue of its larger charge/mass ratio ($1.10 \cdot 10^{-22}$ C/g vs. $7.08 \cdot 10^{-23}$ C/g at pH 7.00). Cyt C = cytochrome C. Cyt C is used as a marker.

Figure 6. c-Myc'/Max' has a similar fold than the Max'/Max' homodimeric b-HLH-LZ in absence of DNA. (a) ^1H - ^{15}N -HSQC of ^{13}C , ^{15}N -c-Myc'SH (red) and ^{13}C , ^{15}N -c-Myc'SH+unlabeled-Max'SH (purple). (b) ^1H - ^{13}C -HSQC of ^{13}C , ^{15}N -c-Myc'SH (red) and ^{13}C , ^{15}N -c-Myc'SH+unlabeled-Max'SH (purple). The two up-field shifted ^1H resonances (boxed) are assigned to c-Myc'SH Leu 396 δ -methyl ^1H . (c) Highlight of the HLH domain of the c-Myc'SSMax'/E-Box (PDB accession no. **1NKP**)²⁰ displaying the δ -methyl of Leu 396 (purple side-chain) of c-Myc' (red ribbon) packing against Phe 43 (purple side-chain) of Max' (blue ribbon) and the δ -methyl of Leu 64 (grey side-chain) of Max' (blue ribbon) packing on Phe 374 (grey side-chain) on c-Myc' (red ribbon). (d) Upper panel: NOE's between Max'SH Leu 64 δ -methyl ^1H and c-Myc'SH Phe 374 aromatic ^1H s (grey) in a plane (132 ppm) of an aromatic ^1H - ^{13}C NOESY-HSQC. Lower panel: upfield region of the non-decoupled 1D ^1H spectrum of the ^{13}C , ^{15}N -c-Myc'SH/ ^{12}C , ^{14}N -Max'SH is also presented (black spectrum). Note that it is the sharper (non-coupled) δ -methyl ^1H s of Leu 64 that are involved in the NOE's with the ^{13}C -attached aromatic ^1H s of Phe 374.

Figure 7. Superposition of the NMR structure of the heterodimeric c-Myc/Max LZ (PDB accession no. **1A93**: yellow side-chains)²³ onto one of the heterodimeric LZ of the c-Myc'SSMax'/E-Box crystal structure (PDB accession no. **1NKP**: grey side-chains)²⁰. Note in (a) the similarity between the conformations of the side-chains and of the

backbones (red = c-Myc, blue = Max). (b) An interfacial salt-bridge involving Max His 81d is observed in both structures. (c) A charged H-Bond between Arg 424a on c-Myc'SH and Gln 91g on Max'SH is observed in the crystal structure.

Figure 8. Effect of pH on molecular recognition between c-Myc'SH and Max'SH in absence of DNA. (a) and (b). Far-UV CD spectra (20°C) and temperature denaturations of c-Myc'SH (red), Max'SH (blue) and c-Myc'SH+Max'SH (purple) at pH 2.8. The arithmetic sum of the spectra and temperature denaturations of c-Myc'SH and Max'SH are shown in grey. (c) and (d). Far-UV CD spectra (20°C) and temperature denaturations of c-Myc'SH (red), Max'SH (blue) and c-Myc'SH+Max'SH (purple) at pH 4.8. The arithmetic sum of the spectra and temperature denaturations of c-Myc'SH and Max'SH are shown in grey. (e). Far-UV CD spectra (20°C) of c-Myc'SH+Max'SH at pH 2.8 (red), 4.8 (blue) and 6.8 (purple). The concentrations of c-Myc'SH and Max'SH were 32 μ M in monomeric units and the total concentration of the mix (c-Myc'SH+Max'SH) was 64 μ M in monomer units.

Figure 9. Equilibriums between the monomeric Max'SH and c-Myc'SH and their dimeric states free and bound to an E-Box DNA sequence. Since c-Myc'SH cannot homodimerize significantly, the formation of Max'SH/Max'SH promotes its accumulation. It can be expected from the law of mass action that, lowering the concentration of c-Myc'SH will lower the population of c-Myc'SH/Max'SH and will increase the population of Max'SH/Max'SH (red arrows). As discussed in the text, we

propose that these equilibrium shifts are critical in the regulation and termination of c-Myc transcriptional activity. c-Myc and Max are rendered in red and blue respectively.

Figure 10. Equilibrium populations of all b-HLH-LZ species. (a). Simulation of the temperature denaturation curve of an equimolar mixture between c-Myc'SH and Max'SH at pH 6.8 and total protein concentration of 64 μ M. The experimental and simulated denaturation curves are in green and black respectively. Light and dark blue curves correspond to the simulated contribution of Max'SH and Max'SH/Max'SH. Light and dark red curves correspond to the simulated contribution of c-Myc'SH and c-Myc'SH/c-Myc'SH. The contribution of c-Myc'SH/Max'SH to the simulated CD_{222nm} signal is given by the purple. Shown in grey is the temperature denaturation of the cross-linked heterodimer: c-Myc'SSMax'. The line corresponds to the temperature depended CD_{222nm} signal of 32 μ M of a population of 100% of folded heterodimer. Note that the helical content of the cross-linked heterodimer is more important than that of the equimolar mixture between c-Myc'SH/Max'SH due to the population of unfolded species. (b). Simulated and equilibrium concentrations of all the species as a function of temperature. Color-coding identical to (a).

Figure 11. (a). Simulation of the population of all species as a function of the total protein concentration (c-Myc'SH+Max'SH) at 20°C (see Materials and Methods for details). The population of each species is given by their concentration divided by the total concentration. One can notice that from 64 to 16 μ M, the concentration of c-Myc'SH/Max'SH decreases while the concentration Max'SH/Max'SH reaches a

maximum. (b). Simulation of the CD_{222nm} signal (mdeg) as a function of concentration. The CD values are calculated from the concentrations displayed in (a) and the $MCD_{222nm}(20^{\circ}C)$ of each specie. Shown in open circle are the experimental values for 8, 16, 32 and 64 μM in total protein concentration.

Table 1. Dissociation constants (K_D) of c-Myc'SH/c-Myc'SH, Max'SH/Max'SH and c-Myc'SH/Max'SH at pH 6.8 and 20 or 37.5°C.

	$K_D(20^\circ\text{C})$	$K_D(37.5^\circ\text{C})$
c-Myc'SH/c-Myc'SH ^{1,2}	$1.08 \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$
Max'SH/Max'SH ²	$1.02 \cdot 10^{-6}$	$8.06 \cdot 10^{-4}$
c-Myc'SH/Max'SH ³ (<i>apparent</i>)	$2.27 \cdot 10^{-6}$	$3.62 \cdot 10^{-6}$
c-Myc'SH/Max'SH ³	$5.14 \cdot 10^{-7}$	$3.53 \cdot 10^{-6}$

¹. Accurate determination of the K_D of c-Myc'SH/c-Myc'SH is complicated in many regards. While the baseline for the unfolded state is straightforward to define, the baseline for a native dimeric state is more uncertain. Indeed, residual dimerization is negligible and the nature of the rare dimeric state ambiguous. Since the LZ of c-Myc at neutral pH is very unlikely to dimerize and that HLH domains can dimerize even if the LZ is unfolded²⁵, we have ascribed the residual homodimeric state to a folded dimeric HLH. This corresponds to 50% of the optimal molar ellipticity attainable for a b-HLH-LZ in absence of DNA²⁶. Therefore the K_D reported for c-Myc'SH/c-Myc'SH has to be taken as approximate. On the other hand, it is evident that the dimerization is negligible and the impact on the determined K_D of c-Myc'SH/Max'SH is therefore minor.

². The K_D for the homodimeric species have been determined as described in references 25-27.

³. The K_{DS} for c-Myc'SH/Max'SH have been determined as described in details in the materials and methods.

Figure 1

	BASIC REGION	HELIX 1	LOOP	HELIX 2	LEUCINE ZIPPER	LINKER		
	30	40	50	60	70	80	90	100
Max	ADKRAHNNALERKRR	DHIKDSFHSLRDSVP	SLQG-EKAS	RAQILDKATEYIQY	MRRKNH	HQDIDDLKRONALLEQQVRALEK		GSGC
c-Myc	NVKRRRTHNVLERQRR	NELKRSFFALRDQIP	ELENNEKAP	KVVILKKATAYILS	VQAE	EOKLISEEDLLRKRREQLKHKLEQLRN		GSGC
	361	371	381	392	402	412	422	432

Figure 2

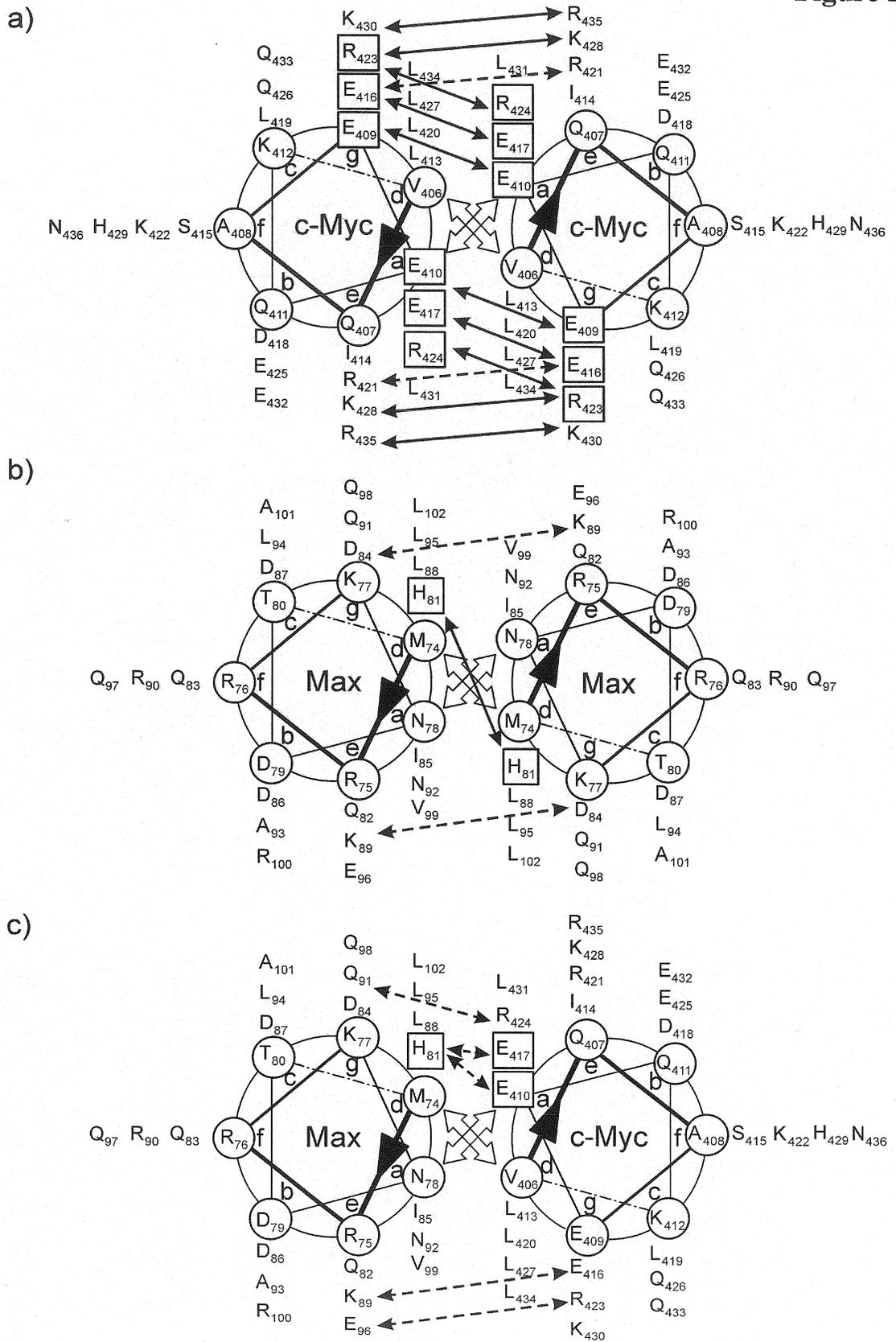
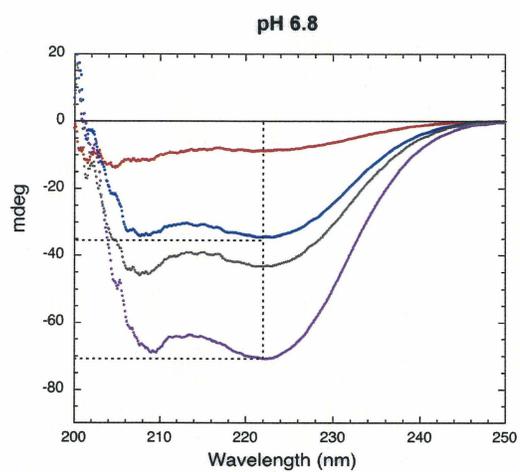


Figure 3

a)



b)

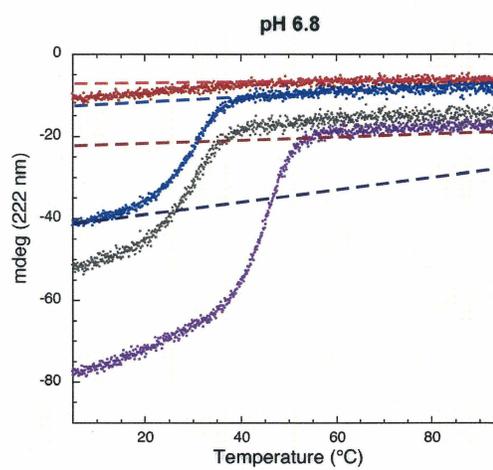


Figure 4

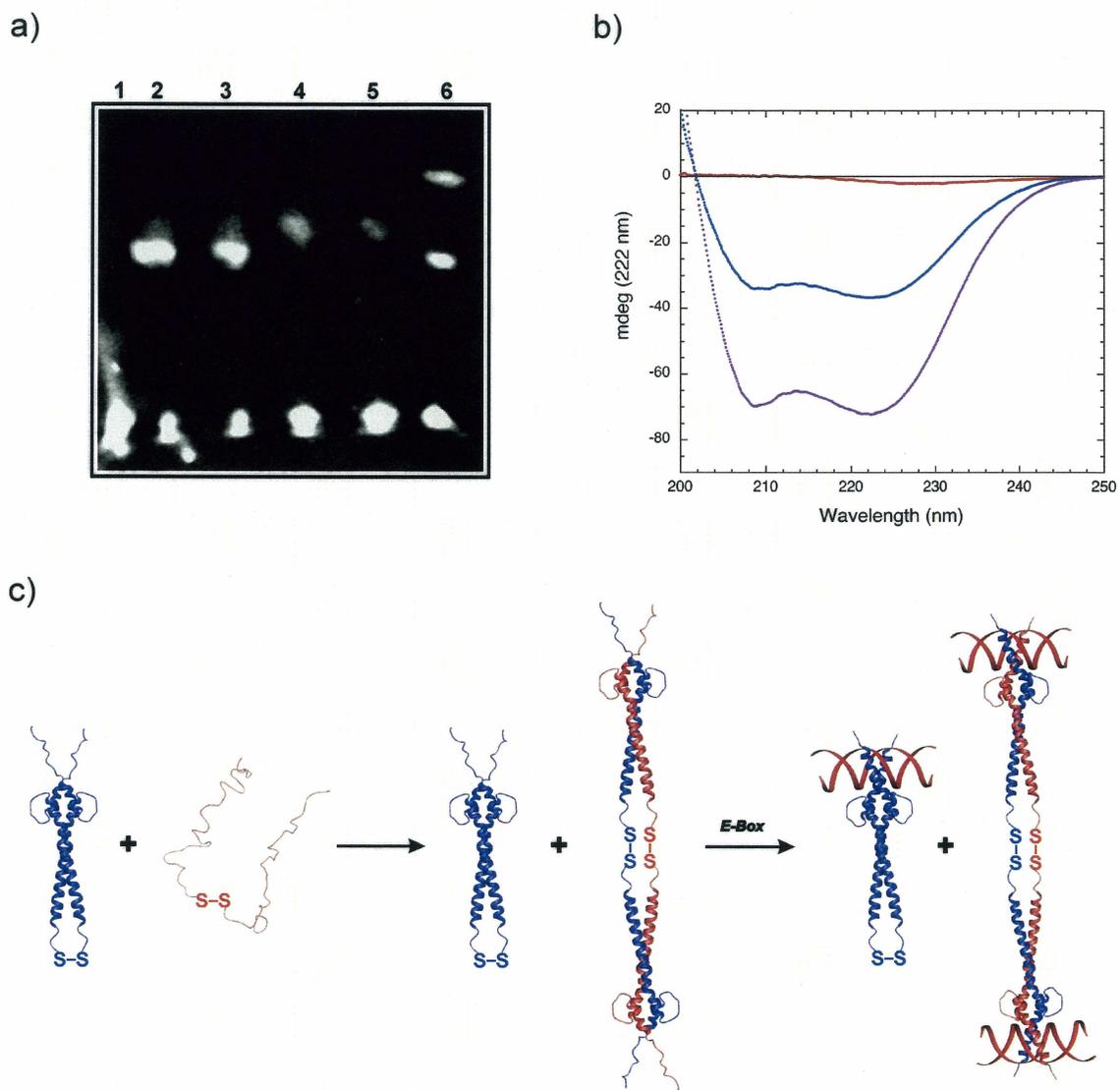


Figure 5

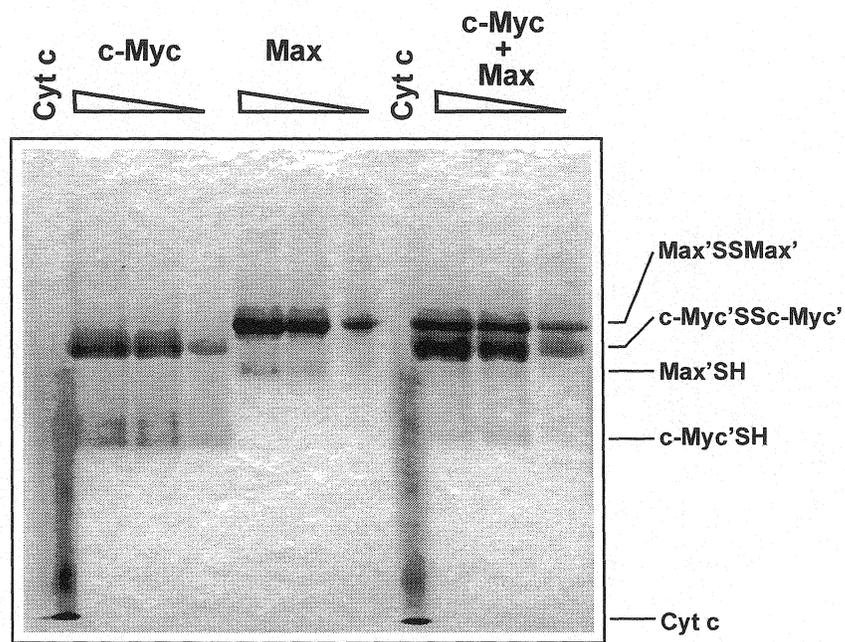


Figure 6

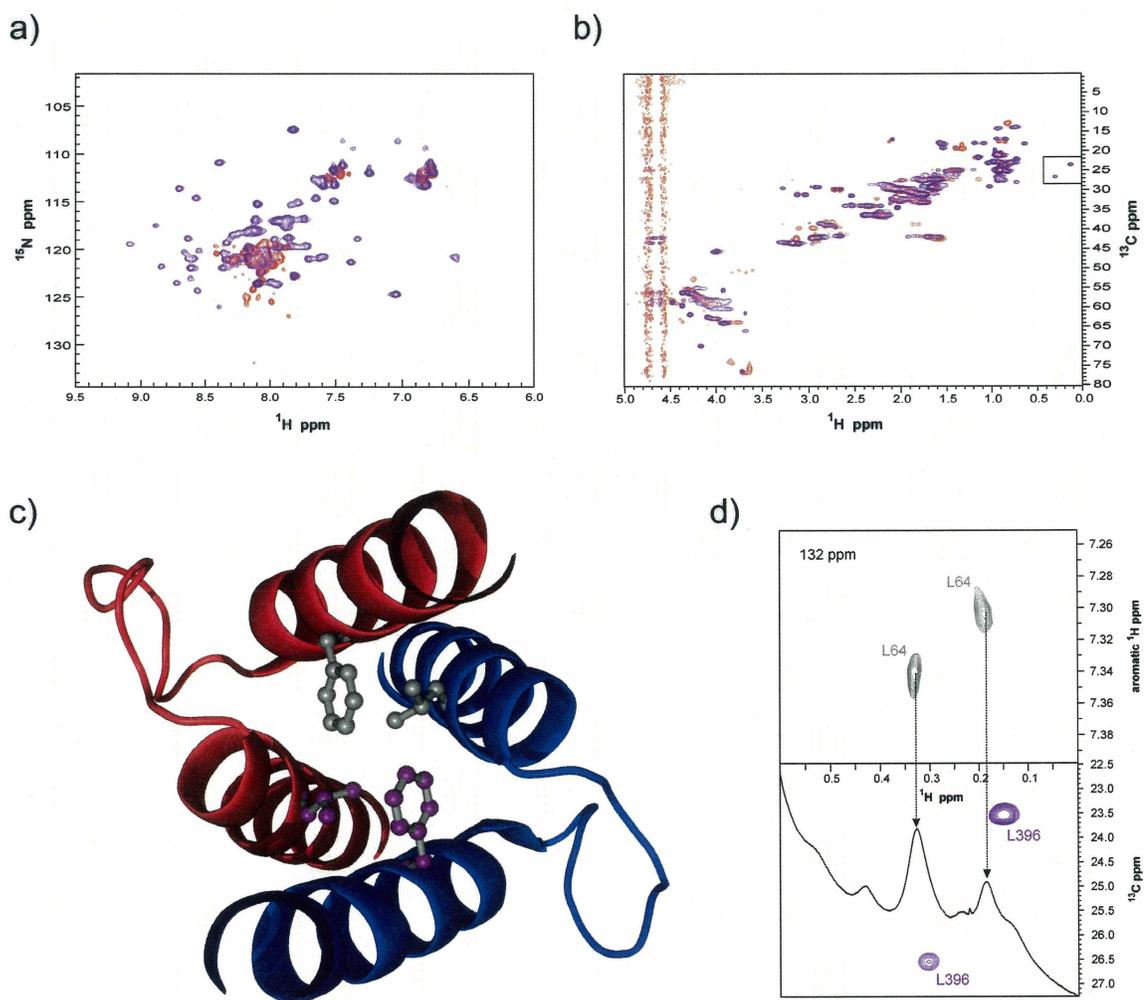


Figure 7

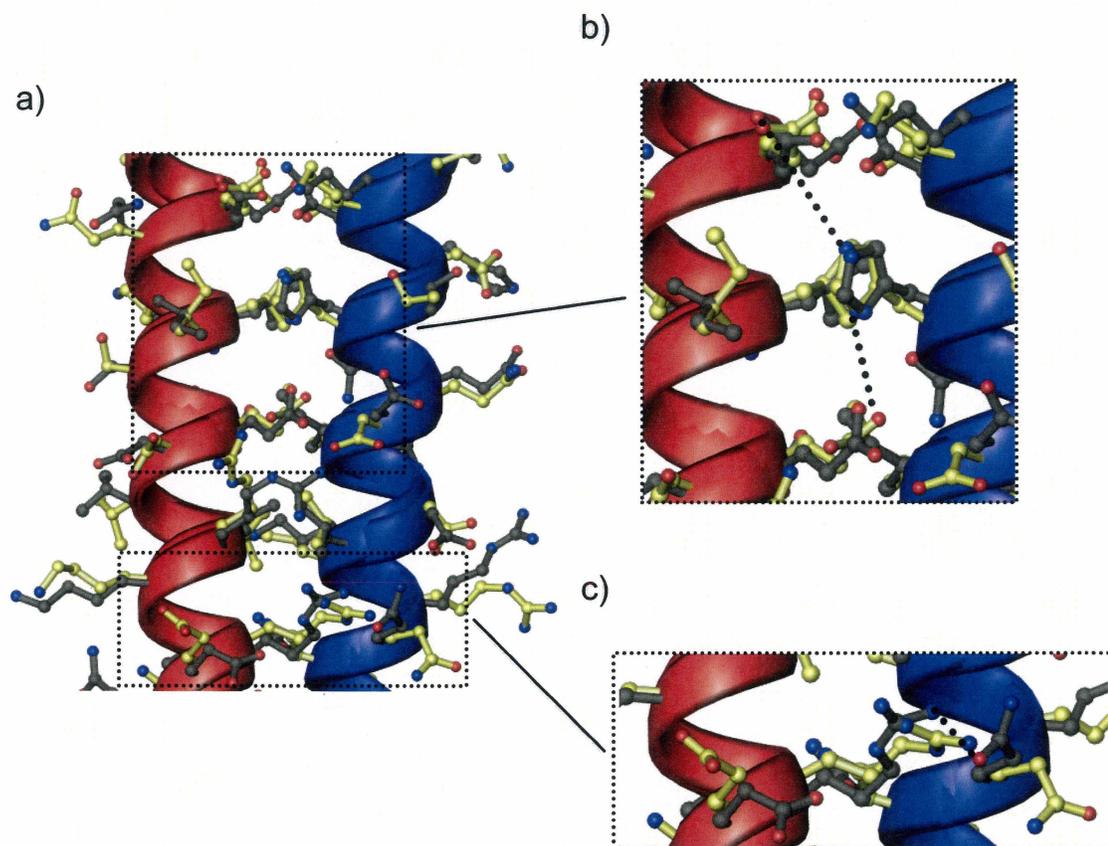


Figure 8

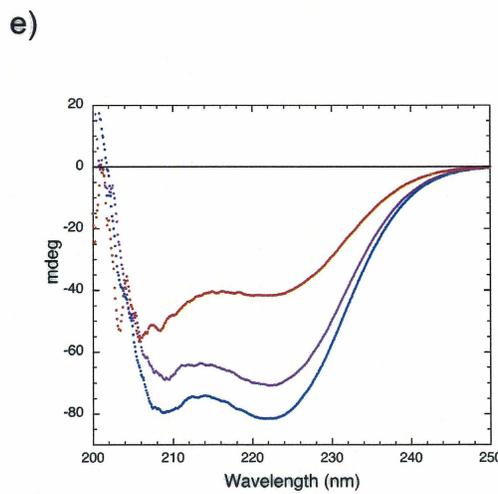
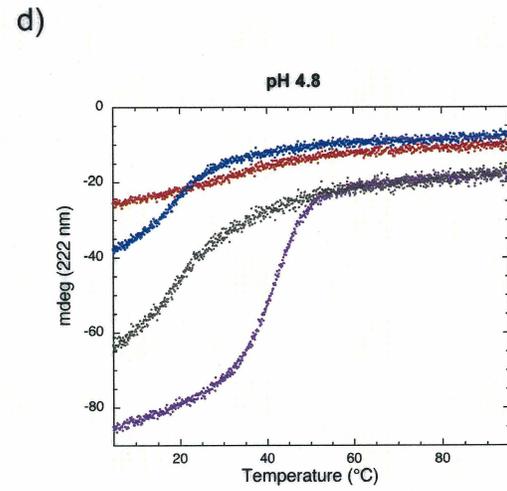
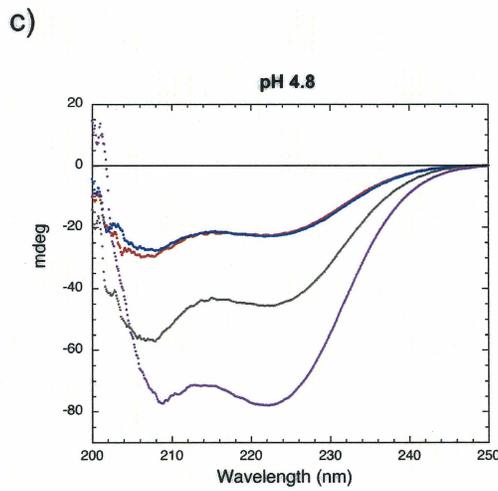
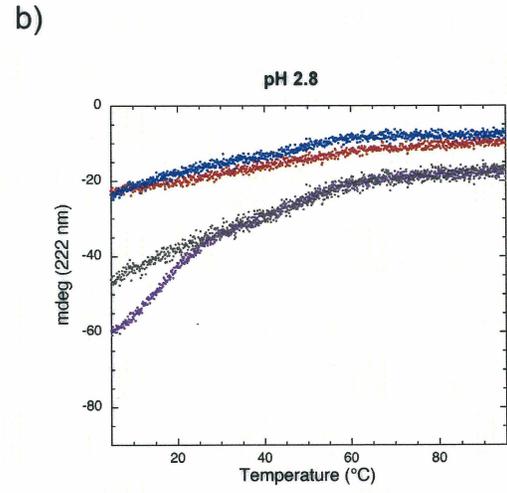
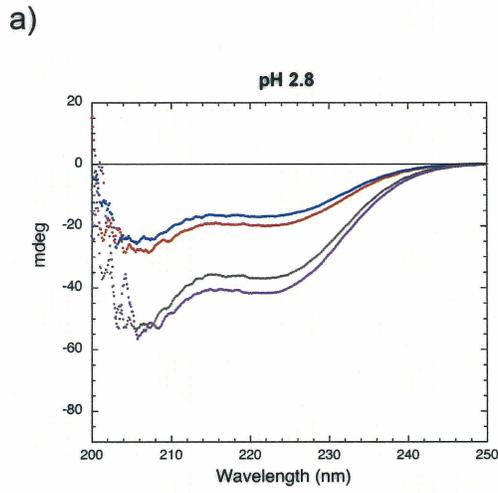


Figure 9

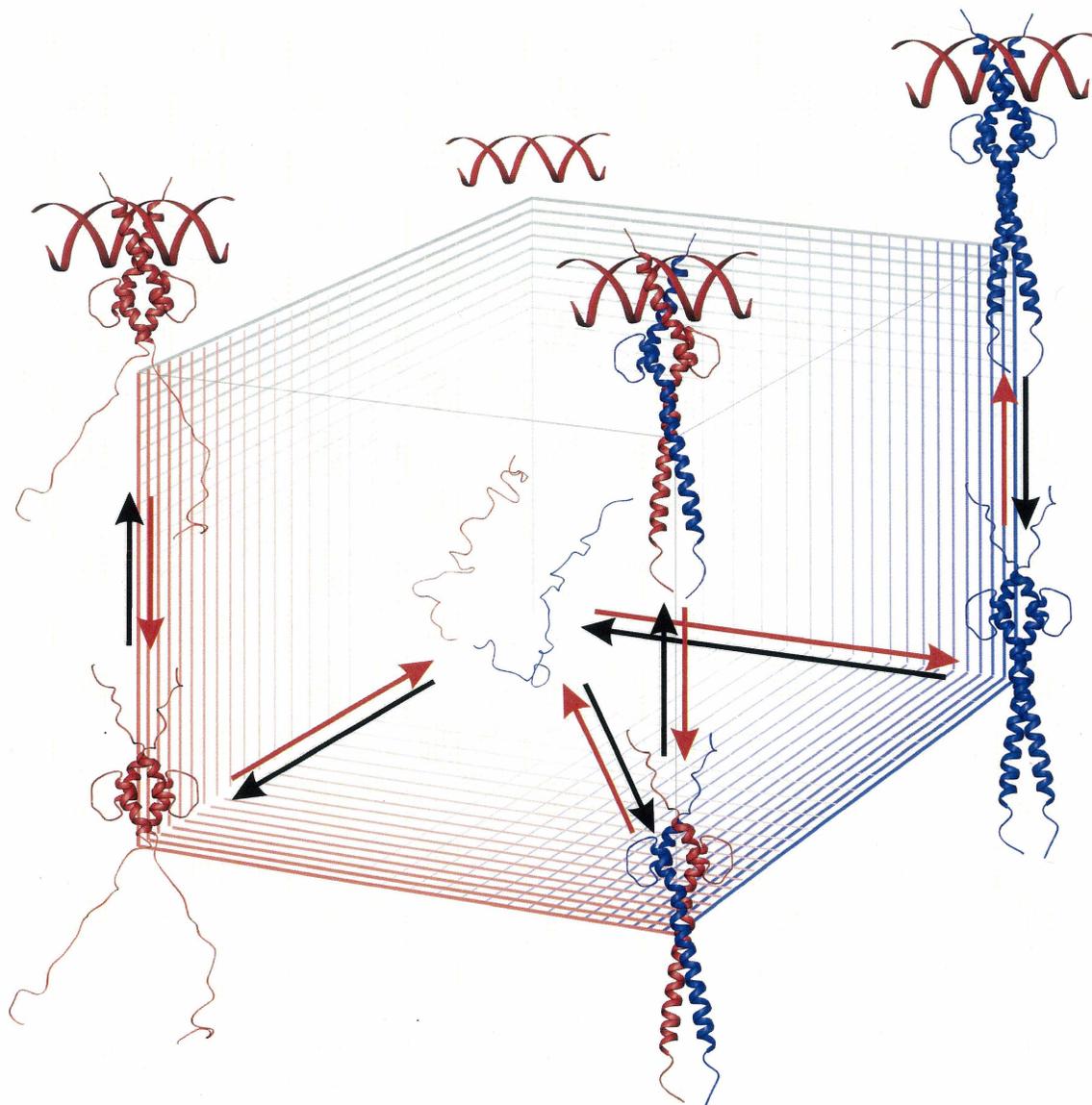
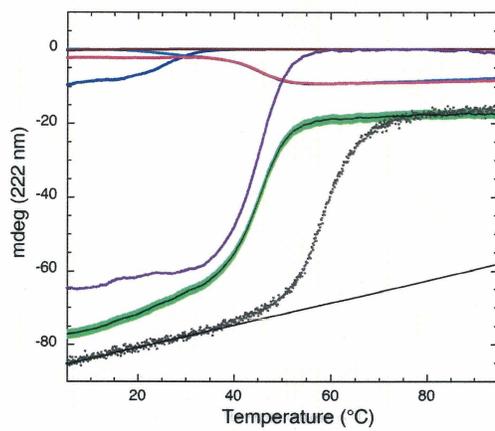


Figure 10

a)



b)

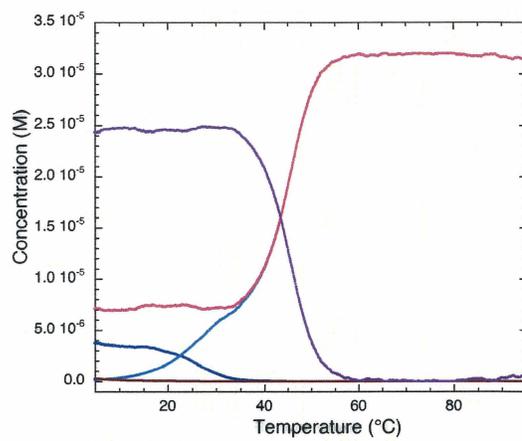
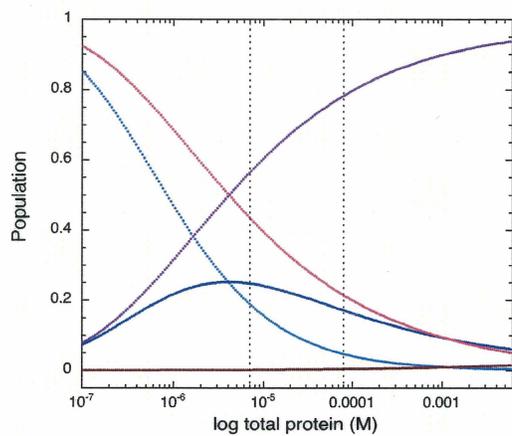
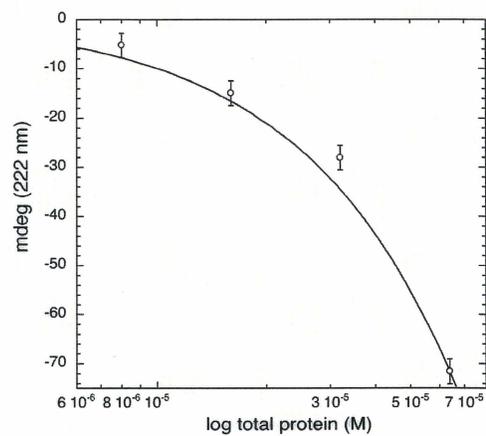


Figure 11

a)



b)



DISCUSSION ET CONCLUSION

6. LES OBJECTIFS DE L'ÉTUDE

L'objectif de notre étude était de caractériser la reconnaissance moléculaire des domaines b-HLH-LZ des partenaires c-Myc et Max en solution en absence d'ADN. Notre étude visait à caractériser le complexe du point de vue structural et thermodynamique, de manière à en comprendre les forces et les faiblesses, et éventuellement développer des inhibiteurs spécifiques de cette reconnaissance moléculaire. Dans cette perspective, nous envisagions de parvenir à défavoriser la formation du complexe c-Myc/Max et sa liaison aux sites E-box situés dans les promoteurs de gènes impliqués dans la prolifération cellulaire et la tumorigenèse.

Dans un premier temps, nous avons réalisé le clonage, l'expression et la purification des domaines b-HLH-LZ des deux partenaires c-Myc et Max, sous la forme de protéine recombinante en *E. coli*. Cette approche nous a permis de produire des mélanges hétérodimères/hétéroisotopiques c-Myc-¹³C-¹⁵N/Max-¹²C-¹⁴N, rendant possible l'étude structurale du complexe par résonance magnétique nucléaire et l'investigation thermodynamique par dichroïsme circulaire. Tel qu'appréhendé à partir de travaux antérieurs (LAVIGNE *et al.*, 1995, 1998), nous avons observé une dépendance au pH dans la reconnaissance des domaines b-HLH-LZ de c-Myc et Max. L'ellipticité moléculaire du complexe c-Myc/Max à pH neutre étant inférieure à celle attendue pour une population à 100% dimérique d'un complexe b-HLH-LZ typique, nous avons

investigué l'influence du pH sur le système et avons observé une ellipticité maximale au pH légèrement acide de 4,80. Cette étape cruciale de notre démarche nous a permis d'appréhender l'existence de populations de monomères et de complexes homodimères conjointement en solution avec l'hétérodimère. Ceci nous a permis d'entrevoir la complexité de ce système *a priori* simple, mais *a posteriori* très articulé.

Dans le but de caractériser convenablement notre système et en déterminer les populations, nous avons développé trois différents outils d'analyse computationnelle (CD_F, CD_SI et KD_I) dont les calculs sont essentiellement basés sur des valeurs obtenues d'expériences spectroscopiques par dichroïsme circulaire. Pour appuyer certaines assomptions inhérentes aux calculs, nous avons caractérisé par RMN la formation du complexe des b-HLH-LZ de c-Myc/Max en solution, en absence d'ADN.

7. CARACTÉRISATION STRUCTURALE ET RECONNAISSANCE MOLÉCULAIRE DE C-MYC ET MAX

Les résultats présentés dans notre étude confirment que l'hétérodimérisation de c-Myc et Max n'est pas optimale à pH neutre mais plutôt à un pH légèrement acide (pH 4,80). Ceci corrèle bien avec la déstabilisation de l'homodimère Max/Max en pH acide rapportée par NAUD *et al.*, (2003) et supporte l'hypothèse selon laquelle la triade « E410_{cMyc}, H81_{Max}, E417_{cMyc} » joue un rôle prépondérant comme déterminant de la reconnaissance moléculaire de ces partenaires. Tout comme l'histidine H81_{Max} pour Max, les glutamates E410_{cMyc} et E417_{cMyc} situés à l'interface du leucine zipper de c-Myc, sont suspectés être responsables de la déstabilisation du complexe homodimère c-Myc/c-Myc aux pHs supérieurs à leurs pKas, en occasionnant des répulsions électrostatiques à l'interface de dimérisation. Ces répulsions, au sein des homodimères, pouvant être neutralisées par la formation de ponts salins à l'interface du complexe hétérodimère, la reconnaissance de c-Myc et Max devrait théoriquement être optimale au pH moyen des pKas des glutamates ($\approx 4,30$) et de l'histidine ($\approx 6,00$), soit un pH d'environ $(4,30 + 4,30 + 6,00)/3 \approx 4,87$ (article Figure 8 (e)).

En ce qui à trait à l'homodimérisation de c-Myc, nos résultats démontrent que l'homodimère c-Myc/c-Myc dépend lui aussi du pH et ne se forme qu'à pH très acide (2,80), soit sous les valeurs de pKa des acides glutamiques et aspartiques. À pH neutre, l'homodimérisation de c-Myc est tout simplement inexistante et la protéine exhibe une forte tendance agrégative en absence de Max. Pour étayer davantage la thèse de la triade

« E-H-E », on note une perte quasi-totale de la reconnaissance moléculaire de c-Myc/Max à pH 2,80, soit lorsque les glutamates de c-Myc sont complètement protonés et incapables de stabiliser la charge de l'histidine sur Max.

Nos résultats démontrent que les domaines b-HLH-LZ de c-Myc et Max se reconnaissent *in vitro* en absence d'ADN et lient l'ADN des E-box sous une forme dimérique. Les évidences de la caractérisation structurale par RMN du complexe en absence d'ADN (article Figure 6), confirme que le complexe c-Myc/Max présente une structure similaire à celles déjà publiées de dimères de protéines du type b-HLH-LZ (NAIR et BURLEY, 2005; SAUVÉ *et al.*, 2004; BROWNLIE *et al.*, 1997; FERRÉ-D'AMARÉ *et al.*, 1993). Bien que nos résultats soient insuffisants pour écarter complètement l'hypothèse d'un tétramère rapporté par NAIRE et BURLEY (2005), nous considérons que le modèle supporté par les auteurs est insoutenable sous nos évidences. Il semble à notre avis, que l'observation de ce tétramère puisse résulter d'un artefact dû à la présence du pont disulfure et à l'utilisation de sels chaotropes dans la préparation des complexes hétérodimères spécifiques. Dans les faits, comme nous démontrons qu'à pH neutre l'hétérodimère c-Myc/Max se trouve en échange avec la forme homodimérique Max/Max, il n'est pas exclus que l'hétérodimère spécifique c-Myc-SS-Max se retrouve lui aussi en équilibre avec un pseudo tétramère *via* Max/Max et/ou encore par le biais d'un réarrangement tête-à-queue de deux hétérodimère spécifiques. Comme les hétérodimères de leur étude furent préparés en présence de fortes concentrations d'hydrochlorure de guanidine (GuHCl) avant d'être désalés, il semble évident que diverses sous-populations de tétramère puissent s'être formées sans qu'il

s'agisse pour autant de complexe tel que ceux retrouvés dans la maille du cristal utilisé pour la diffraction de rayons-X.

Enfin, bien que la formation d'hétérodimère des b-HLH-LZs c-Myc et Max bénéficie naturellement d'un avantage statistique sur l'homodimérisation de Max en solution à pH neutre, l'hétérodimère subit tout de même une compétition non négligeable de la part de l'homodimère Max. Dans l'ensemble, nos résultats démontrent que bien que l'association des domaines b-HLH-LZ de c-Myc et Max ne soit pas optimale aux conditions physiologiques, le complexe hétérodimère est tout de même fonctionnel, replié convenablement et capable de lier les cibles E-box. Tel que discuté plus loin, on peut penser que la dynamique d'échange qui prend place entre c-Myc/Max et Max/Max confère aux facteurs de transcription une capacité de réponse plus rapide et plus fluide dans leur rôle de modulateur des événements cellulaires.

8. THERMODYNAMIQUE DE L'ÉQUILIBRE GLOBAL INTRICUÉ DE C-MYC ET MAX

Les populations de complexes Max/Max et c-Myc/Max étant sensibles à des interactions avec d'autres partenaires comme Mad, Miz-1 et l'ADN *in vivo*, le modèle présenté ici mérite certaines réserves d'interprétation. L'ensemble des paramètres régissant l'équilibre global intriqué et leurs influences respectives souvent non linéaires, rendent son analyse complexe et sa réinterprétation physiologique très délicate. Dans cet ordre d'idée, à lui seul, l'impact des liaisons des complexes à l'ADN laisse entrevoir des équilibres passablement différents *in vivo*, vu les interdépendances démontrées au sein de l'équilibre global intriqué. Comme, à la fois l'homodimère Max/Max et l'hétérodimère c-Myc/Max peuvent lier l'ADN, mais avec des affinités différentes, il est probable que l'asservissement des populations par ces nouvelles interactions *in vivo* engendre un réarrangement complet de l'équilibre global intriqué. Cela dit, notre modèle permet pour la première fois d'esquisser réellement les implications des différents paramètres sous-jacents de cet équilibre global intriqué et d'en amorcer une ébauche plus complète. Par contre, en ce qui a trait aux affinités réelles des partenaires c-Myc et Max complets *in vivo*, la question doit malheureusement rester ouverte pour l'instant. Une caractérisation complète du système est nécessaire et implique de rendre compte des affinités d'association du complexe c-Myc/Max avec le double brin d'ADN parallèlement à l'association de l'homodimère Max/Max avec ce même brin d'ADN. Il va sans dire que l'homodimère Mad/Mad et l'hétérodimère Mad/Max devraient eux

aussi être pris en compte dans une modélisation de l'équilibre « général » intriqué : c-Myc/Max/Mad/E-box.

À titre d'exemple concret, si l'affinité homodimérique plus grande de Max complet (alias p21 Max) laisse présager une population transitoire p21/p21 plus importante *in vivo*, on ne doit pas négliger le fait que l'effet observé pour p21 comparativement au b-HLH-LZ seul, est le fruit d'une diminution de la valeur du « k_{off} » $\left(K_D = \frac{k_{\text{off}}}{k_{\text{on}}} \right)$ (NAUD *et al.*, 2005). Cette diminution de la vitesse de dissociation est, rappelons-le, produite par un blindage électrostatique du domaine basique *via* une région acide en N-terminal de p21 Max. Cet effet, bien que réellement favorable à l'homodimère p21/p21, n'est pas restrictif à l'homodimère seul, puisqu'il devrait théoriquement favoriser aussi, dans une moindre mesure, une diminution du k_{off} pour l'hétérodimère c-Myc/p21. De plus, on note qu'une région acide similaire à celle de p21 se retrouve également sur la protéine c-Myc complète (ADHIKARY et EILERS, 2005), avec hypothétiquement les mêmes implications stabilisantes pour le complexe c-Myc/p21. Dans cette nouvelle perspective, les populations à l'équilibre ne sont peut-être que singulièrement influencées par ces nouveaux attributs, mais la réponse de l'équilibre global aux fluctuations de concentration elle, est assurément exacerbée par ce raffermissement bilatéral des affinités homo et hétéro.

L'étude approfondie de l'équilibre global intriqué de c-Myc et Max nous à tout de même permis d'entrevoir que : « à concentration équivalente de partenaire », la population d'homodimères Max est maximale aux environs de 10^{-6} M et qu'elle décroît

proportionnellement à l'augmentation ou la réduction des concentrations des partenaires au-delà ou en-deçà de 10^{-6} M. Dans un contexte physiologique (où la concentration de Max reste à peu près constante par rapport aux variations de concentration de c-Myc et Mad) ceci peut permettre une réponse plus rapide de ce système associatif/dissociatif, en préparant une certaine quantité d'homodimère Max/Max, transitoirement liée à des sites E-box.

Par ailleurs, les résultats de nos simulations permettent de comprendre que, pour contrer la reconnaissance moléculaire de c-Myc/Max, on ne doit pas envisager de développer un « super chélateur de Max » puisque l'astringence du système vient principalement d'une affinité c-Myc:Max et non d'une trop grande disponibilité de Max. Dans un cas de surexpression dérégulée de c-Myc, l'hétérodimère gagnera toujours la bataille dans les équilibres de complexes puisqu'il profite d'une dépendance aux concentrations des partenaires selon un ordre 1 et non selon le carré comme c'est le cas des homodimères. Ainsi, même dans le cas d'une cellule « normale » en prolifération (avec des concentrations de c-Myc estimées à 10^{-6} M), on devrait, pour chélater Max et réduire à moins de 10% la population de c-Myc/Max, développer une protéine d'une affinité équivalente à un K_D homo de 10^{-10} pour Max wt, et qui n'exhibe aucune affinité homo pour elle-même, ni hétéro pour c-Myc (voir Figure 39 B).

On peut également envisager de décupler l'avantage déjà conféré par l'homodimère Max/Max en lui juxtaposant un chélateur de c-Myc. De fait, une protéine ou un peptide ayant la possibilité de priver l'équilibre global en protéine c-Myc permettrait de relâcher

la tension exercée sur l'homodimère Max/Max et rendrait éventuellement ce dernier disponible pour une hétérodimerisation avec les protéines Mad. Un candidat liant c-Myc avec une affinité de l'ordre de 10^{-9} et exprimé dans l'ordre du micro molaire enclencherait une débâcle du complexe c-Myc/Max puisqu'il s'unirait à l'effet déjà délétère de l'homodimère Max/Max. En tirillant l'équilibre chacun de leur côté, ces alliés pourraient réduire à moins de 10% la population d'hétérodimère c-Myc/Max (résultats non publiés). Cela dit, il s'agit du cas d'une cellule exprimant des quantités normales de c-Myc, et le peptide ne devrait pas pouvoir lier Max et Mad, en plus de ne fournir à c-Myc aucune possibilité de lier l'ADN, au risque de ne simplement déplacer le problème dans une autre dimension de l'équilibre « général » intriqué.

Une dernière tentative consisterait, à l'aide d'une petite molécule, à interférer directement dans l'affinité de c-Myc/Max pour dévaluer leur attrait réciproque. Dans ce cas, pour engendrer une réduction de la population de c-Myc/Max à moins de 10% (toujours aux concentrations cellulaires normales) il faudrait synthétiser une molécule capable d'augmenter le $K_{D, c-Myc/Max}$ à 3×10^{-5} (voir Figure 39B) et qui n'interférerait pas avec l'homodimère Max/Max.

On comprend donc à présent pourquoi les stratégies relevant d'une surexpression de la protéine mutante MaxVL (un super Homo Max) dans des essais *in cellulo*, ne rendent pas compte des effets escomptés. En plus d'un manque de spécificité pour Max wt, les affinités relatives de MaxVL pour lui-même et pour c-Myc (NAUD *et al.*, 2005 résultats non publiés), contribuent seulement à déplacer le problème dans une autre

sphère de complexité. La présence de nouveaux résidus hydrophobes à l'interface du leucine zipper de MaxVL ne contribue pas à déstabiliser efficacement le complexe c-Myc/MaxVL. Ainsi, MaxVL projette l'équilibre global intriqué sur un nouveau palier de compétition avec des affinités décuplées par rapport aux complexes wt. Dans cet ordre d'idée, une surexpression de MaxVL dans des cellules surexprimant déjà c-Myc, ne peut avoir d'autres effets que de fournir plus de munitions à l'hétérodimère.

Toutes ces stratégies comportent peut-être un avantage commun caché, soit celui d'éventuellement engendrer de grandes instabilités génétiques, voire des avortements de division cellulaire. Toutefois, comme l'hétérodimérisation bénéficie toujours d'un avantage statistique par rapport aux complexes homos, et que les mécanismes d'association de c-Myc/Max, Max/Max et de Mad/Max sont très similaires, il apparaît improbable de parvenir un jour à synthétiser des molécules capables de rompre la totalité des complexes c-Myc/Max sans interférer d'aucune manière avec les autres complexes et produire des effets dévastateurs sur les cellules saines avoisinantes d'une tumeur. Le système hétérodimère représente un choix extrêmement judicieux de la cellule pour assurer l'efficacité de son mécanisme de transactivation. À faible concentration de c-Myc, le système profite d'une abondance de Max pour engendrer rapidement des complexes hétéros, tandis qu'à saturation de c-Myc, c'est au tour de Max de bénéficier de l'abondance de c-Myc. Jamais le complexe ne se trouve réellement défavorisé, à moins bien sûr que les expressions des deux partenaires soient simultanément réprimées, ce qui n'arrive jamais. Dans les cas pathologiques de surexpression de c-Myc, il apparaît

improbable de parvenir un jour à bloquer complètement et spécifiquement la formation de l'hétérodimère c-Myc/Max.

Bien que plutôt défaitistes, toutes ces considérations d'ordre mécanistique n'auraient pu être clairement explicitées sans le recours à une caractérisation fidèle de l'équilibre global intriqué de c-Myc et Max, et à une détermination précise des constantes d'affinités réelles des complexes homos et hétéro. Notre conception intuitive de pareils systèmes ne peut servir de base rationnelle à l'élaboration de stratégies de recherche et de traitements. Ce faisant, la caractérisation thermodynamique des complexes formés de la reconnaissance moléculaire des domaines b-HLH-LZ des protéines humaines c-Myc et Max représente un percée cruciale dans la caractérisation et l'étude thermodynamique de complexes multimères de macromolécules en solution. Notre ensemble d'outils computationnels a permis, pour la première fois à notre connaissance, d'analyser et de simuler les différentes populations d'un équilibre d'homodimères et d'hétérodimère intriqués. De plus, ces outils permettent de prédire les bouleversements moléculaires occasionnés par des variations de concentration de l'un et/ou de l'autre partenaires, ou encore de considérer des altérations des K_{DS} des différents complexes. Les résultats issus de ces simulations permettent de mieux saisir la plasticité de ce type d'équilibre macromoléculaire et d'en préciser davantage les enjeux. Nos outils permettent enfin d'envisager des criblages d'inhibiteurs de reconnaissance moléculaire, avec le pouvoir de discriminer exactement les effets sur chacune des différentes cibles en solution et sans le besoin de recourir à des spectroscopies complexes et coûteuses.

En intégrant à notre suite de programmes les résultats des plus récents travaux sur la reconnaissance spécifique de complexe tri et tétramérique (MONTAGNE *et al.*, 2005), de même que sur la spécificité de la reconnaissance de cibles ADN (NAUD *et al.*, 2005; SAUVÉ *et al.*, 2007), il devient possible d'envisager la résolution de problèmes thermodynamiques beaucoup plus complexes avec une simplicité similaire. On pourrait ainsi éventuellement traiter l'intégration de Mad et/ou de Miz-1 au système de c-Myc/Max, ou encore caractériser les liaisons concomitantes des différents complexes à des cibles ADN multiples, en présence de différents inhibiteurs. On doit finalement noter que nos outils ne sont nullement restrictifs à l'analyse du système c-Myc/Max/Mad et peuvent tout aussi bien être utilisés pour caractériser des systèmes différents avec d'autres techniques spectroscopiques.

Dans l'ensemble, nos résultats lèvent le voile sur les interdépendances au sein du réseau de facteurs de transcription Myc/Max/Mad et suggèrent que le développement d'une thérapie génique visant à empêcher la reconnaissance moléculaire de l'hétérodimère c-Myc/Max par le biais d'agonistes ou d'antagonistes peptidiques représente une avenue complexe et malheureusement très peu prometteuse.

9. CONSIDÉRATIONS THÉRAPEUTIQUES

Bien qu'il s'agisse de la pierre d'assise de la plupart des mécanismes cellulaires, l'analyse thermodynamique de système d'équilibres intriqués relève encore aujourd'hui du défi biophysique. Plus que jamais, des complexes macromoléculaires multi-protéiques et multifonctionnels sont chaque jour révélés par les biochimistes et biologistes moléculaires du monde entier. Étant nous-mêmes confrontés à un tel problème avec le système de c-Myc et Max, il nous est apparu indispensable de développer enfin des outils d'analyse permettant de traiter ces systèmes complexes de manière efficiente et précise. Nous croyons que la biologie moléculaire et la pharmacologie moderne ne peuvent poursuivre leur déploiement en réduisant éternellement les interrupteurs moléculaires à de simples systèmes à deux états « ON/OFF ». Dans les faits, c'est justement l'analyse approfondie de ces équilibres intriqués qui permet d'en révéler toute la splendeur et le raffinement. On observe alors à quel point certains systèmes peuvent ou ne peuvent pas, doivent ou ne doivent pas pouvoir être altérés dans leurs fonctions.

On attribue aujourd'hui à c-Myc le rôle de dépositaire de la reconnaissance des sites d'activation de la traduction des gènes impliqués dans la prolifération cellulaire. Considérant la croissance d'un organisme aussi « pluricellulaire » qu'un être humain, on comprend à ce simple énoncé que le rôle de c-Myc « ne doit pas » pouvoir être modulé aisément. Une cellule, quelle qu'elle soit, ne doit pas pouvoir enclencher un cycle de division cellulaire et soudainement être arrêtée à mi-parcours pour une raison anodine. Il

ne s'agit ici pas de décider si « oui ou non » on transforme le glucose en glycogène ou *vice versa*, il s'agit habituellement de décider du sort d'une descendance cellulaire possiblement ultra différenciée et extrêmement nombreuse. Le système de la reconnaissance moléculaire de c-Myc et Max doit en conséquence être excessivement bien régulé et couvrir toute la gamme des événements moléculaires de façon très précise. De cette simple analyse contextuelle, on découvre les implications sur le devenir de toute protéine c-Myc au cœur d'une cellule : elle doit ou bien « agir », ou bien « périr » et être rapidement et efficacement éliminée. Le sort de la division cellulaire ne repose alors plus sur c-Myc, mais bien sur un mécanisme simple et rigide, la régulation de l'expression de c-Myc.

Bien entendu, c'est ici que le bas blesse, dans une majorité de cancer, c-Myc n'est pas en faute de « trop bien reconnaître Max », mais plus simplement d'être trop bien exprimé. La « surexpression » d'un facteur de transcription activateur de la prolifération cellulaire, ne peut intuitivement conduire à autre chose qu'à la prolifération cellulaire. Il n'y a donc aucune surprise dans la suite d'événements observés dans la biologie moléculaire du cancer : dérégulation et expression de c-Myc, expression des gènes cibles, prolifération cellulaire. On anticipe déjà que quiconque voudrait, dans un but thérapeutique, interférer dans le processus de reconnaissance moléculaire entre c-Myc et Max, se butera implacablement sur l'un des systèmes les plus rigoureux que comporte une cellule évoluée. Ce « n'est pas », et ça « ne doit pas » être le lot de c-Myc de décider si « oui ou non » il doit traduire des gènes. Son travail doit être celui de reconnaître « quels gènes » doivent être traduits. Le mécanisme, une fois déclenché,

« ne doit pas » pouvoir être arrêté en cours de route, sinon au risque de conduire à une instabilité génétique et automatiquement au suicide de la cellule, l'apoptose, avec les pertes énergétiques effarantes que cela implique pour l'individu. Si c-Myc est présent dans une cellule, il y aura donc forcément traduction et enclenchement de la prolifération. Pour stopper la prolifération en visant c-Myc, il faut éliminer toutes traces de c-Myc, ou de Max, ou encore, engendrer une instabilité génétique globale. Bien entendu, il existe d'autres cibles dérivées potentielles, comme les protéines du complexe TRAPP, les polymérase ou la liaison à l'ADN, mais du point de vue de c-Myc/Max la solution est simple, « noir ou blanc ».

Des suites de notre analyse, il apparaît donc évident que d'autres avenues doivent être envisagées. La présence de c-Myc au sein d'une cellule cancéreuse ne doit, à notre avis, plus être perçue comme un mal, mais plutôt comme un bien dont on doit apprendre à se servir. Toutes les cellules cancéreuses n'ont d'autre choix que de produire c-Myc pour parvenir à se multiplier. Ainsi, c-Myc représente un excellent marqueur de la prolifération cellulaire et doit être utilisé à cette fin. Si les implications de c-Myc dans le pouvoir apoptotique sont bien réelles (ROTTMAN et LÜSHER, 2006), les cellules cancéreuses présentant des dérégulations de l'expression de c-Myc ont aussi en elles le pouvoir de s'éliminer d'elles-mêmes. Il apparaît donc essentiel non pas d'envisager une inhibition de l'expression de c-Myc, ni même une défavorisation de la reconnaissance moléculaire de c-Myc et Max, mais bien seulement une déstabilisation génétique globale de la cellule.

Une des cibles les plus prisées d'une telle stratégie est sans conteste les télomères, dépositaires officiels de la longévité cellulaire. Les télomères, au nombre de 92 par cellules (2 fois 2 paires de 23 chromosomes), représentent le « Talon d'Achille » de toute cellule cancéreuse. Perpétuée par une enzyme elle-même sous le contrôle de c-Myc (la sous-unité catalytique de hTert)(SAKAMURO et PRENDERGAST, 1999), ils soutiennent la stabilité génétique des chromosomes en agissant comme « bouts de lacet ». Il s'agit probablement du nombre de cibles le plus restreint qu'on puisse imaginer dans une thérapie génique. On pourrait ce faisant envisager une thérapie dans laquelle on utiliserait un gène encodant une endonucléase spécifique aux télomères, et dont le partenaire obligatoire serait c-Myc. L'expression de cette chimère (*e.g.* Doigt-de-Zinc-Endonucléase-b-HLH-LZ) pourrait elle-même être régulée par un promoteur sous le contrôle de c-Myc/Max *via* un site E-box. De cette façon, l'expression de la chimère au cœur des cellules en prolifération, l'ouverture de la chromatine des télomères et le mécanisme d'élimination des cellules par apoptose seraient tous commandés par la protéine à l'origine même du désordre cellulaire : c-Myc. Une telle stratégie nécessite encore de préciser un mécanisme de délivrance du nouvel ADN exogène aux sites des tumeurs, mais présente (du moins en apparence) l'avantage de découpler les sécurités du mécanisme d'activation.

Pour la construction de la chimère, il apparaît évident que l'utilisation (en C-terminal) des domaines HLH-LZ de Max wt (sans le domaine basique) représente le choix le plus judicieux pour la tâche de la reconnaissance moléculaire avec c-Myc. Pour ce qui est de la partie reconnaissance des télomères, on peut envisager de juxtaposer un domaine

basique *de novo* à la chimère, et utiliser des motifs « doigt de zinc » (MILLER *et al.*, 2007; SZCZEPEK *et al.*, 2007) à l'extrémité N-terminale. Reste alors à inclure au cœur de la future enzyme un domaine exhibant une activité endonucléase spécifique aux séquences des télomères et la construction est complète.

MATÉRIEL ET MÉTHODES

10. PRODUCTION DU b-HLH-LZ DE C-MYC

10.1. Clonage et expression du b-HLH-LZ de c-Myc

10.1.1. Clonage du b-HLH-LZ de c-Myc

L'ADN complémentaire encodant le domaine b-HLH-LZ de c-Myc (résidu 353N à 436N) a été amplifié par PCR à partir de pSP72-c-Myc (gracieusement fourni par le Dr. Claude Asselin, département d'Anatomie et de Biologie Cellulaire, Université de Sherbrooke, Québec, Canada) et inséré dans un plasmide pET-3a de Novagen entre les sites de clonage NdeI et BamHI, en aval d'un promoteur T7. Le produit de ligation a ensuite été utilisé pour transformer diverses souches de bactéries *Escherichia coli* compétentes et possédant un système endogène d'induction de la polymérase du phage T7 (e.g. Induction à l'IPTG de l'élément DE3 dans les souches BL21 DE3 pLysS).

10.1.2. Expression du b-HLH-LZ de c-Myc

La souche BL21-AI (AI) d'Invitrogen fut sélectionnée pour la surexpression en milieu de culture Luria Bertani (LB : 10 g Bacto tryptone, 5 g Yeast extract, 10 g NaCl, par litre), par induction à l'aide de 0,2 % P/V L-arabinose, à 37°C, pour 24 heures. La souche

AI s'est par contre révélé inefficace pour la surexpression en milieu de culture minimal M9 (Milieu M9 : 100 ml de sln. Sels M9 10X (sln. Sels M9 10X : 5 g NaCl, 30 g KH_2PO_4 , 31.77 g Na_2HPO_4), 20 ml sln. 20% P/V Glucose (^{12}C ou ^{13}C), 5 ml sln. 20% P/V NH_4Cl (^{14}N ou ^{15}N), 1 ml de 1M MgSO_4 , 900 ml $\text{H}_2\text{O}_{\text{dd}}$ autoclavée), nécessaire à la production de protéine uniformément marquée ^{13}C , ^{15}N , et essentielle aux études de résonance magnétique nucléaire. Une raison proposée pour expliquer la perte d'expression en milieu minimal est l'inhibition de l'opéron *araBAD* par le glucose du milieu de culture. Comme l'expression de notre vecteur est directement régie par l'expression de la polymérase T7, elle-même induite par l'intermédiaire d'un promoteur *araBAD*, sensible au L-arabinose, la production de la polymérase (*et ce faisant l'expression de notre gène*) est soumise à la capacité du L-arabinose à lier l'inducteur de l'opéron *araBAD*. Hors, tout comme l'opéron *Lac* bien connu, l'*araBAD* est soumis à une inhibition compétitive en présence de glucose (MIYADA *et al.*, 1984).

Une autre hypothèse fut aussi envisagée, celle-ci concernant la toxicité relative du fragment b-HLH-LZ de c-Myc qui, rappelons-le, est un facteur de transcription. La surexpression de c-Myc à l'aide de la souche AI en milieu LB se soldait par la production de corps d'inclusion à l'intérieur desquels la protéine est emmagasinée systématiquement. Ceci ayant pour effet de protéger à la fois la protéine de la dégradation et la bactérie de l'intoxication. L'induction de protéine en milieu minimal représente alors un défi, puisque la bactérie, limitée dans sa production de corps d'inclusion dû à la pauvreté du milieu, devra dégrader la protéine pour éviter l'intoxication.

À la lumière de ces deux hypothèses, nous avons donc procédé à de nouveaux essais de surexpression pour identifier une souche habilitée à produire c-Myc en milieu M9. Nous avons identifié la souche BL21-CodonPlus (C+) de Stratagene, qui exprime la protéine de façon constitutive sous forme de corps d'inclusion. L'expression appréciable mais toutefois moins importante de la souche C+ par rapport à la souche AI en milieu LB, nous persuada de ne l'utiliser que pour les productions nécessitant le milieu M9.

L'étalement sur pétris d'un frottis à partir de stocks glycérol de souches bactériennes transformées ayant présenté certaines irrégularités dont la perte d'expression, nous avons eu recours à la conservation des souches bactériennes sous la forme de « Boosters » produits à partir d'une culture bactérienne non-induite ($DO_{600nm} = 0,6$) à laquelle est ajouté 15% de glycérol v/v avant d'être aliquotée à 1,5 ml et gelé à $-80^{\circ}C$ pour utilisations futures. Au besoin, les « Boosters » sont dégelés et utilisés selon un ratio 1.5 ml / litre de milieu de culture. Le principal avantage est de permettre une inoculation efficiente à tout coup sans avoir à recourir à un étalement sur pétris et à une pré-culture. Les boosters étant à usage unique, ceci permet d'éviter les cycles de gel dégel et les contaminations des précieux « stocks glycérol » conventionnels.

Une fois autoclavés, les milieux de culture utilisant la souche AI sont additionnés d'ampicilline (sel de sodium) à raison de 50 mg/l (Sln. AMP 1000X (50 mg/ml) : 2,5 g de sel d'ampicilline, 50 ml H_2O_{dd}) et les milieux utilisant la souche C+ doivent être supplémentés à la fois d'ampicilline à raison de 50 mg/l et de chloramphénicol à 34 mg/l

(Sln. CAM 1000X (34 mg/ml) : 1,7 g de chloramphénicol, 50 ml éthanol 70% v/v) pour assurer la conservation du plasmide C+.

La croissance des cultures AI et C+ est réalisée à 37°C avec agitation à 250 rpm. À une densité optique de 0,6 ($DO_{600nm} = 0,6$) la culture AI est induite à l'aide de 0,2% p/v de L-arabinose (Sln. L-Ara 100X (20g/100 ml) : 10 g de L-arabinose, 50 ml H₂O_{dd}), puis l'incubation est poursuivie pour un minimum de 12 heures. La souche C+ quant à elle, ne requiert aucune induction mais peut tout de même être supplémentée de 600 mM d'Isopropyl-β-D-thiogalactopyranoside (IPTG) sans aucun effet délétère observé (Sln. IPTG 100X (14,3mg/ml) : 0,714 g d'IPTG, 50 ml H₂O_{dd}). Vu le coût associé aux milieux servant aux marquages isotopiques ¹³C et ¹⁵N, cette mesure de sûreté fut tout de même préconisée.

Une fois l'induction achevée, la culture bactérienne est récoltée par centrifugation à 10 000 rpm, 4°C, pendant 5 min dans un rotor SLA-1500. Le milieu de culture est centrifugé séquentiellement (4 X 250 ml), pour combiner 1 litre de culture par bouteille.

10.1.3. Laemmli à bactéries

Pour l'optimisation des inductions et la sélection des souches bactériennes, nous avons dû revoir notre façon de réaliser les migrations électrophorétiques d'extraits crus de bactérie. c-Myc s'exprimant en corps d'inclusion, ses inductions étaient souvent

difficiles à distinguer par SDS-PAGE. Il fallait s'assurer de lyser et solubiliser convenablement les énormes complexes protéiques que sont les corps d'inclusion et en permettre l'électrophorèse. Les simples gels-dégels conventionnels produisant des pistes de migration souvent sales et diffuses, nous avons envisagé d'utiliser un sel chaotrope pour faciliter la solubilisation des protéines. L'hydrochlorure de guanidine (GuHCl) fut d'abord envisagé mais avec les effets d'augmenter la force ionique des échantillons de façon telle que leurs migrations devenaient encore plus troubles. L'urée ne s'ionisant pas en solution, nous avons donc ajouté 4M urée à la recette conventionnelle de solution Laemmli 2X pour échantillons (125 mM Tris-HCl, pH 6,80, 3% ^{P/v} Lauryl Sulfate (SDS), 20% ^{v/v} glycérol, 20% ^{v/v} β -mercaptoéthanol, 0,02% ^{P/v} Bleu de bromophénol) et avons appelé cette solution « Laemmli à bactéries ».

L'optimisation du protocole de lyse bactérienne s'effectua par essais et erreurs, et permit de définir le ratio optimum : 75 μ l de Laemmli à bactérie / culot obtenu de la centrifugation de 1 ml de culture bactérienne à $DO_{600nm} = 0,6$. La densité optique des cultures bactériennes variant durant les cinétiques d'induction, une simple « règle de trois » permet de relativiser tous les échantillons pour une même concentration sur gel. La procédure de gel-dégel s'effectue, de préférence, à l'aide d'un bloc chauffant à 100°C et du bain d'azote liquide à -180°C. 6 cycles assurent le démantèlement complet des parois, des membranes bactériennes et des corps d'inclusion. Pour faciliter le travail de l'urée et du SDS, il est préférable de vortexer les tubes à tous les deux cycles et de laisser chauffer les échantillons pendant 3 minutes à 100°C à la toute fin, pour assurer le travail complet du β -mercaptoéthanol. Les échantillons doivent par la suite être

centrifugés 60 secondes à 13 000 rpm à l'aide d'une mini-centrifugeuse pour précipiter l'ADN, les parois bactériennes et les membranes. Le SDS et l'urée assurent la solubilité des protéines durant ce processus qui élimine la majorité des interférents de migration. Finalement, 12,5 µl de surnageant est chargé sur un « mini-gel » conventionnel de 0,75 mm d'épaisseur.

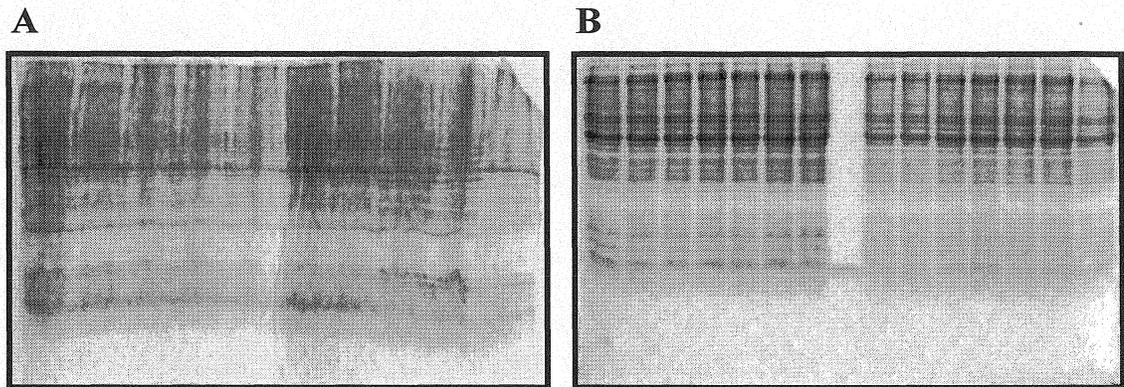


Figure 41 : SDS-PAGE d'extrait crus de culture bactérienne A) Laemmli traditionnel, B) Laemmli à bactérie.

10.2. Isolement du b-HLH-LZ de c-Myc

Bien que connus depuis déjà plus de 20 ans, les études *in vitro* sur les protéines Mycs ne débutèrent que tout récemment dû aux difficultés d'obtenir une protéine stable et soluble. Il aura fallu surmonter de nombreux problèmes sous jacent à la nature même de c-Myc avant de pouvoir l'isoler. L'insolubilité et l'agrégation, de même que la déficience des expressions liée à la toxicité du transcrit de gène, furent les principaux obstacles. Les premières études publiées furent ainsi réalisées soit à partir de synthèses peptidiques du sous-domaine LZ (MUHLE-GOLL *et al.*, 1995; LAVIGNE *et al.*, 1995; LAVIGNE *et al.*, 1998) ou du fragment b-HLH-LZ de v-Myc, un membre de la famille particulièrement soluble (FIEBER *et al.*, 2001). Aucun protocole n'ayant été rapporté concernant la purification de c-Myc, on dut s'en remettre aux notions de bases et à l'expertise développée antérieurement dans le laboratoire pour la purification du b-HLH-LZ de Max. Cette quête longtemps restée infructueuse fût résolue à l'aide de protocoles de lyse, de centrifugations différentielles et des principes de chromatographies échangeuses de cations.

Les expressions en milieu minimal n'ayant, à ce moment, toujours pas porté fruit, les études RMN restèrent en suspens. Huit mois après le début des travaux, la souche C+ fut finalement identifiée pour permettre le marquage isotopique et c'est à ce moment (janvier 2003) que fût rapporté pour la première fois une étude portant sur le fragment b-HLH-LZ de c-Myc obtenu par génie recombinant (NAIR et BURLEY, 2003). En conditions dénaturantes, l'équipe de Stephen Burley utilisait l'affinité à une colonne de

nickel, venant d'une queue His-Tag ajoutée à la protéine. L'artéfact de clonage est par la suite retiré par digestion enzymatique et la protéine repurifiée par échange cationique. Cette stratégie, quoi que très efficace, présente le défaut de négliger la compréhension des aptitudes propres de la protéine.

Quoi que notre protocole de purification fût très performant pour les expressions de la souche AI en milieu LB, il s'est avéré un cuisant échec lors de la purification des cultures C+ en milieu minimal M9. Il fallut investir d'avantage de travail pour parvenir à isoler la protéine à partir des expressions en milieu M9 et ce n'est qu'après une année complète de travail que furent obtenues les premières récoltes de protéines uniformément marquées ^{13}C et ^{15}N . Cette démarche particulièrement longue, a permis d'élaborer le seul protocole connu à ce jour de purification du b-HLH-LZ de c-Myc sans recours à un artéfact de clonage. Elle conféra au laboratoire une expertise unique dans la compréhension du maniement de cette protéine au comportement malaisé et permis l'accès direct à la purification d'un autre membre de cette famille de facteur de transcription soit la protéine Mad (MONTAGNE *et al.*, 2005). Des rendements de purification de l'ordre de 90% estimés par SDS-PAGE, assurèrent un approvisionnement suffisant de protéine exprimée à raison de 20-25 mg/l pour la souche AI et de 10-15 mg/l pour les cultures uniformément marquées ^{13}C , ^{15}N .

La base de notre protocole de purification utilise les avantages d'une expression en corps d'inclusions, selon la méthode de Vuillard et Freeman, (VUILLARD *et al.*, 1995) c'est pourquoi il s'est avéré inadapté lors de la transition entre les souches AI et C+, les deux

colonies n'expriment pas les mêmes proportions de constituants cellulaires dans les différents milieux LB et M9.

10.2.1. Lyse et isolement des corps d'inclusion

La lyse cellulaire s'effectue à l'aide de lysozyme en conditions non dénaturantes et peu détergentes afin de retirer les constituants cellulaires solubles, de ne pas resolubiliser les corps d'inclusions et de ne pas entraver le travail ultérieur de la DNaseI. Les bactéries culottées sont d'abord resuspendues dans le tampon « Lyse c-Myc » (50 mM HEPES-NaOH pH 7,50, 100 mM NaCl, 10 mM MgCl₂, 5 mM DTT, 350 µg/ml Lysozyme) à raison de 10 ml / culot de 1 litre de culture bactérienne.

Une fois resolubilisés, les culots sont rassemblés dans une même bouteille et incubés avec agitation à 37°C; 50 rpm; 15 min, pour permettre au lysozyme de blanc d'œuf d'attaquer le peptidoglycan de la paroi bactérienne. Après incubation, 100 µl de détergent Triton X100 sont ajoutés pour chaque culot de 1 litre de culture. Les parois affaiblies et les membranes solvatées, les bactéries sont alors lysées par sonication sur glace à puissance moyennement élevée, jusqu'à homogénéisation. Le pH 7.5 du tampon assure la précipitation de la protéine advenant le démantèlement de certains corps d'inclusion.

Une fois la solution homogénéisée, 400 µl de DNaseI pancréatique bovine (27-0516-01 Amersham 20 mg) (1mg/ml dans 50% glycérol) sont ajoutés pour chaque culot de 1 litre de culture initialement lysé. Le lysa bactérien est alors incubé de 60-180 min; 37°C; 50 rpm. Le temps nécessaire à la digestion de l'ADN génomique pouvant varier, la digestion optimum est obtenue avec un maximum de temps de réaction. La protéine n'étant pas resolubilisée, elle ne souffre que très peu des protéases libérées dans le lysat et la purification profite d'une dégradation maximale de l'ADN et des protéines solubles.

Le lysat est alors redistribué dans des tubes à centrifugation (35 ml) de type SS34, de manière à ne contenir par tube que la quantité équivalent à 1 litre de culture initialement apprêtée (environ 12 ml). Ceci permet de profiter des avantages de la force centrifuge distribuée au plein volume du tube à l'étape suivante. La centrifugation est réalisée à 19 500 rpm, (30 000 G), pour 30 minutes à 4°C. Les corps d'inclusion constituent le culot avec certains complexes de haut poids moléculaire comme les parois bactériennes, les ribosomes et l'ADN génomique non dégradé. Les lipides, les protéines solubles, les acides aminés, les sucres et les acides nucléiques constituent le surnageant. Conséquemment, une fois la centrifugation complétée, le surnageant peut être jeté.

10.2.2. Solubilisation des corps d'inclusion et élimination des hauts poids moléculaires.

Les corps d'inclusion devant être dissous pour permettre la récupération de la protéine, chaque culot est resolubilisé avec 15 ml de tampon « Bull Cracker Part A » (6 M Urée, 100 mM NaAc pH 5,00, 500 mM GuHCl) auquel on ajoute 25 mM de DTT frais. Le Bull Cracker est une solution acide, fortement ionique et hautement dénaturante qui permet de solubiliser efficacement les dépôts protéiques et de maintenir c-Myc en solution pour éliminer par centrifugation les complexes de hauts poids moléculaires et l'ADN résiduel. Les tubes sont alors vortexés pour permettre à la solution d'imbiber et décoller le culot, jusqu'à la fragmentation en plus petits morceaux. Le mélange est alors dilué 1:1 avec 15 ml de « Bull CRACKER Part B » (2 M Urée) et soniqué sur glace à puissance minimum pour démanteler la totalité des particules. La sonication ne doit pas être excessive pour ne pas risquer d'endommager la protéine. Lorsque la solution est homogène, elle est centrifugée à 19 500 rpm, (30 000 G); 4°C; 30 min. Au besoin, on équilibre les tubes à centrifugation à l'aide de Bull CRACKER PART B. Lorsque la centrifugation est complétée, le surnageant contenant c-Myc est décanté et les culots peuvent être jetés.

10.3. Purification du surnageant par FPLC

La purification du surnageant est réalisée par chromatographie échangeuse de cations à l'aide de mini-colonnes HiTrapTM SP Sepharose HP de GE Healthcare (17-1152-01). Le système HiTrapTM présente les avantages de fournir une matrice de très haute fidélité, sous un format individuel de 1 ou 5 ml de gel. Les colonnes possèdent des embouts vissables qui permettent de connecter aisément plusieurs colonnes en série pour obtenir un volume de gel sur mesure. Un remplacement des unités vieillottes du sommet par des colonnes neuves à la base permet de toujours disposer d'une colonne de haute fidélité, dans une condition excellente, à un coût raisonnable.

L'élution s'effectue en deux étapes, soit : un lavage en pH faiblement basique et un gradient de force ionique à pH 5,00. L'élution est suivie à l'aide de deux paramètres soit l'absorbance de lumière UV à 280 nm et la conductivité électrique. Pour une purification typique (2,5 litres de culture), le détecteur UV est ajusté à une sensibilité de 0,5 AUFS « *Absorbance Units Full Scale* » et l'enregistreur est réglé à 0,5 V (200% de sensibilité). Les paramètres du fabricant sont utilisés pour la cellule de conductivité et l'enregistreur est réglé à 1,0 V (100% de sensibilité).

10.3.1. Chargement de la colonne

Le Bull Cracker est un tampon optimisé pour permettre une adhésion optimale de c-Myc lors du chargement de la colonne, tout en défavorisant l'adhérence des contaminants de point isoélectrique inférieur à c-Myc ($pI_{c-Myc} \sim 10$). La colonne est formée de 5 unités HiTrap de 5 ml, mises bout à bout pour un volume total de 25 ml de gel. La colonne est préalablement rincée avec du Tampon B (50 mM NaAc, pH 5,00, 5 M NaCl) avant d'être équilibrée par deux volumes de colonne de Tampon A (50 mM NaAc, pH 5,00). La colonne est chargée à raison de 5 ml/min pour un maximum de 75 ml soit l'équivalent de 2,5 litres de culture, puis rééquilibrée avec un volume de Tampon A.

10.3.2. Lavage à l'aide de Tampon U8

Pour éliminer le principal contaminant de c-Myc lorsque élué à pH 5,00, un lavage est réalisé à 2,5 ml/min avec 3 volumes de Tampon U8 (50 mM Tris-HCl pH 7,50, 4 M urée). Le U8 est une solution moyennement dénaturante et de pH légèrement basique qui permet d'éliminer les contaminants ayant un ratio charge/masse comparable à celui de c-Myc à pH 5,00 mais possédant un pI inférieur à celle-ci. L'élévation du pH permet alors d'éluer spécifiquement ces contaminants sans décrocher la protéine de la colonne. La nature dénaturante du U8 est vouée à défavoriser les éventuelles interactions non spécifiques de contaminants avec c-Myc (*e.g.* protéines hydrophobes et/ou acides). Une

fois le lavage à l'aide de U8 terminé, la colonne est rééquilibré avec un volume de Tampon A pour permettre de recharger le gel avant le gradient de force ionique.

10.3.3. Réalisation du gradient de force ionique

Le gradient est réalisé avec les tampons A et B. Comme la concentration en sel du tampon de chargement (Bull CRACKER) correspond à la force ionique d'un mélange d'environ 10% v/v Tampons B/(A+B), un front de 15 ml de 10% Tampon B peut être utilisé pour démarrer le gradient. Le gradient est par la suite réalisé de façon linéaire, de 10% à 35% en 50 ml (0,5%/ml) et à un débit de 2,5 ml/min. L'élution est collectée et fractionnée dans des tubes de polystyrène à raison de 2,5 ml/frc. Le point de chute de c-Myc en solution dans 50 mM acétate de sodium pH 5,00 correspond à 1,5 M NaCl soit environ 30% v/v B/(A+B) et la protéine est éluée avec une pureté supérieure à 90%. Au besoin, les fractions intéressantes peuvent être rassemblées et diluées par deux avec du Tampon A avant d'être rechargées sur la colonne pour reprendre le gradient et obtenir une pureté supérieure à 99%.

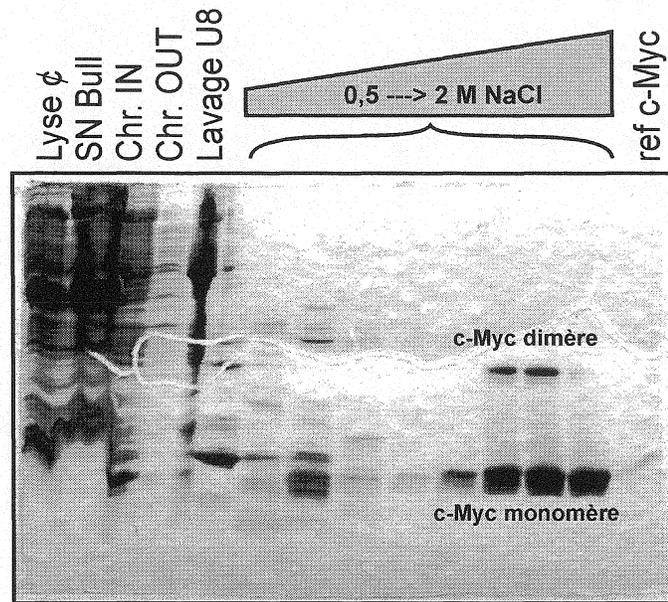


Figure 42 : SDS-PAGE de la purification du b-HLH-LZ de c-Myc

Les échantillons du gradient sont chargés sur gel de polyacrylamide 16.5% sans β -mercaptoéthanol, pour visualiser la proportion de complexe c-Myc'SSc-Myc' dimère. Les pistes réfères à : Lyse ϕ) Lyse cellulaire à l'aide de tampon « Lyse c-Myc », SN Bull) surnageant de centrifugation dans tampons « Bull A + B + DTT », Chr. IN) chargement de la colonne de chromatographie échangeuse de cation, Chr. OUT) élua de la colonne au chargement, Lavage U8) élua de la colonne au lavage à l'aide de tampon « U8 », 0,5 -> 2 M NaCl) fractions du gradient de force ionique par mélanges de tampons « A » et « B », ref c-Myc) référence de migration « protéine c-Myc pure ».

10.4. Dessalage et lyophilisation

Une fois purifiée, la protéine peut être dessalée et lyophilisée. Deux méthodes peuvent être envisagées pour réaliser cette opération, soit le recours à une matrice hydrophobe de type « Silice C8 » ou encore l'utilisation de colonnes HiTrap™ Dessalting de GE Healthcare (17-1408-01). Les essais de membranes à dialyse ont permis d'observer une incompatibilité de c-Myc avec cette technique, conduisant à l'adhésion de la protéine à la membrane et à son agrégation, probablement dû à l'affinité de la protéine pour les charges négatives du complexe cellulosique de la membrane. L'utilisation de matériel à base de verre nu, (*e.g.* colonne de verre non enduit de silanes) est également à proscrire pour les mêmes raisons d'interaction de charge.

10.4.1. Dessalage par Silice C8

Le dessalage à partir de matrice hydrophobe C8 présente l'avantage de permettre un dessalage poussé en peu de temps. En ne requérant qu'une seule étape (*la totalité de la récolte pouvant être liée à la matrice*) la technique permet de concentrer la production dans le volume d'une même élution à l'aide d'un front d'acétonitrile. Le principal problème de cette technique reste le choix de la colonne qui requiert ou bien du verre apprêté ou un matériel plastique résistant à l'acétonitrile. Les mélanges d'eau et d'acétonitrile provoquant la formation de bulles d'air dans la matrice, il faut de plus préalablement dégazer les solutions et utiliser un système à haute pression pour assurer

une élution conviviale. Le recours à un système HPLC « *High Pressure Liquid Chromatography* » utilisant un gradient H₂O 0,05% v/v acide trifluoro-acétique vs acétonitrile 0,025% v/v acide trifluoro-acétique représente la meilleure des options pour cette opération.

10.4.2. Dessalage par Exclusion de taille

Ne disposant pas d'un système HPLC au sein du laboratoire, nous avons eu recours au système HiTrapTM Dessalting qui préconise l'effet de tamis moléculaire pour éluer la protéine par exclusion de taille tout en retardant les ions de sel. La séparation est produite par une différence des vitesses d'élution entre les produits, elle-même fonction de leur capacité à pénétrer ou non les pores des billes de la matrice. La dimension des pores régit la séparation par la différence de liberté de diffusion (d'entropie) qu'elle impose aux petites et aux grosses molécules. Plus une particule est petite, plus elle perçoit un volume important de solvant et peut diffuser, ce faisant plus elle prendra de temps pour traverser la colonne. Cette technique permet des rendements de l'ordre de plus de 95% de récupération de protéine et de 98 à 99% de dessalage.

L'exclusion de taille présente le seul désavantage de limiter le volume de chargement de la colonne relativement au volume de gel. Pour contourner ce problème nous avons préconisé de connecter en série 5 unités HiTrapTM de 5 ml pour obtenir un volume de 25 ml de gel et éviter d'avoir à scinder les récoltes en un nombre dérisoire de chargements.

Le volume de chargement est préférablement en deçà de 7,5 ml et l'injection doit être effectuée dans un régime d'au plus 5 ml/min. Pour assurer la solubilité de c-Myc et permettre l'élimination des contre-ions par lyophilisation, nous avons préconisé d'équilibrer et d'éluer la colonne avec 0,05% v/v acide trifluoro-acétique dans H_2O_{dd} . L'élution est réalisée à 2,5 ml/min et la protéine est récupérée dans des fractions de 2,5 ml dans les 15 premiers millilitres suivant l'injection. La récolte est suivie à l'aide des mêmes paramètres d'absorbance et de conductivité que pour l'élution cationique (voir section 10.3).

10.4.3. Lyophilisation

Les essais de lyophilisation à partir d'élutions sur Silice C8 ayant conduit à un lyophilisat plus spongieux et plus aisément resolubilisable, nous avons systématiquement préconisé de diluer 1:1 avec de l'acétonitrile, les élutions obtenues de dessalage par tamis moléculaire. La récolte est finalement immergée dans l'azote liquide et placée à lyophiliser. Une fois complètement lyophilisée la récolte est pesée puis resolubilisée à l'aide d'un mélange eau/acétonitrile 50% v/v contenant 0,05% v/v acide trifluoro-acétique dans un volume correspondant à 100 μ l/mg. La récolte peut alors être aliquotée dans des tubes eppendorfs à raison de 1 à 10 mg et lyophilisée de nouveau.

11.PRODUCTION DU B-HLH-LZ DE MAX

Les étapes conduisant à l'obtention du domaine b-HLH-LZ de la protéine Max par génie recombinant ayant déjà fait l'objet de travaux antérieurs, seulement les nouvelles observations et principaux développements réalisés durant cette étude seront mentionnés ici. (*Pour plus de détails, voir (NAUD et al., 2003).*)

11.1. Clonage et Expression du b-HLH-LZ de Max

L'ADN encodant les domaines b-HLH-LZ de Max à été sous-cloné dans le vecteur pET-3a et transformé dans la souche *E. coli* pLysS d'Invitrogène. L'induction de l'expression de la souche pLysS étant régulée par l'ajout d'IPTG, la compétition avec le glucose en milieu minimal ne pose pas de problème vu la haute affinité de l'IPTG pour l'opéron *Lac*. La souche pLysS peut donc être utilisée aussi bien pour les expressions en milieu LB qu'en milieu M9. Le milieu de culture servant à la surexpression est enrichi de chloramphénicol à 34 mg/l et d'ampicilline à 50 mg/l pour conserver à la fois le plasmide pLysS et le vecteur pET-3a. Dans l'éventualité d'une toxicité cellulaire causée par la protéine surexprimée, le plasmide pLysS encode le lysozyme du phage T7 qui réduit les fuites de l'élément DE3 en inhibant partiellement la transcription de la polymérase T7 et conséquemment des promoteurs sous le contrôle de cette dernière.

La protéine Max étant soluble, l'expression conduit rapidement à la lyse bactérienne et à la perte du produit d'expression dans le milieu de culture après environ quatre heures d'induction à 30°C. Comme la culture doit être centrifugée hâtivement pour éviter des pertes, nous avons optimisé les rendements en allongeant la phase de croissance jusqu'à une densité optique de 0,9, avant d'induire à l'aide de 14,3 mg/l d'IPTG pour 3 heures. La multiplication cellulaire étant ralentie par l'induction de l'expression, la croissance jusqu'à DO 0,9 permet d'augmenter à son maximum la population bactérienne avant l'induction, tout en évitant la suffocation rencontrée aux densités optiques supérieures à 1,2. Ceci nous permet d'atteindre des rendements de l'ordre de 50 à 75 mg/l de culture.

11.2. Isolement du b-HLH-LZ de Max

La purification du b-HLH-LZ de la protéine Max est beaucoup plus aisée que celle de c-Myc. Le transcrit de gène étant soluble au pH physiologique, sa purification s'effectue à partir du surnageant obtenu de la lyse cellulaire. Chaque culot obtenu à partir d'un litre de culture est d'abord resolubilisé avec 10 ml de tampon « Lyse Max » (50 mM Tris-HCl pH 7,40, 100 mM NaCl, 10 mM MgCl₂, 0,1% v/v Nonidet P40) puis vortexé et combiné aux autres culots dans une même bouteille.

Afin d'augmenter les rendements et d'éviter l'oxydation délétère de la queue cystéine de Max avec d'autres constituants cellulaires, nous avons envisagé d'ajouter du dithiotreitol (DTT) à raison de 5 mM. Cette démarche s'est soldée par un échec dû à la formation d'un co-polymère entre le DTT et le polyethyleneimine (PEI) utilisé pour précipiter l'ADN. Le PEI étant essentiel à la purification de Max, le DTT a été retiré du protocole. Soulignons que ce phénomène posait également problème pour la purification de c-Myc et que nous avons, dans ce cas, préconisé d'éliminer le PEI plutôt que le DTT, et avons simplement prolongé le temps de réaction de la DNaseI.

Une fois resolubilisés, les culots bactériens sont soniqués sur glace à puissance moyenne par impulsions d'environ 30 à 60 secondes, jusqu'à homogénéisation. Par la suite, 300 µl de DNaseI pancréatique bovine (27-0516-01 Amersham) (1mg/ml sln 50% glycérol) sont ajoutés pour chaque litre de culture initialement lysé et incubé 30 minutes à 37°C avec agitation. Une fois le temps de réaction complété, 250 µl d'une solution 5% p/v PEI

sont ajoutés pour chaque litre de culture initial. Le PEI étant un polymère positivement chargé et donc chélateur de charges négatives, son action n'est pas instantanée. Une incubation 15 minutes sur glace avec agitation a été ajoutée au protocole original pour favoriser sa polymérisation avec l'ADN et augmenter son efficacité.

Le mélange est par la suite distribué dans des tubes à centrifugation de type SS34 de manière à occuper un volume maximal, puis les tubes sont centrifugés à 19 500 rpm (30 000G); 30 min; 4°C. Cette procédure assure l'élimination d'un maximum de macromolécules de haut poids moléculaire tout en conservant la totalité de la protéine Max soluble, évitant une surcharge inutile de la matrice de chromatographie.

Pour favoriser une plus grande sélectivité d'adhésion des protéines à la colonne et par le fait même éliminer d'avantage de contaminants, une dilution 1:1 du lysat avec un mélange 20% v/v Tampon B/(A+B) (préalablement réfrigéré) a été ajouté (*Tampon A et B, voir section 10.3.1*). L'acidification apportée par cette dilution engendre la précipitation de nombreux contaminants qui sont aisément éliminés par une centrifugation additionnelle à 19 500 rpm (30 000G); 30 min; 4°C.

11.3. Purification du surnageant par FPLC

La purification du surnageant s'effectue à l'aide d'un assemblage de 5 colonnes HiTrap™ Hp Sepharose SP, préalablement rincé à l'aide de Tampon B et rééquilibré à l'aide de Tampon A. La charge maximum est de 25 ml (environ 2,5 litres de culture) pour assurer une adhérence et une élution convenable. La purification est suivie à l'aide de l'absorbance UV à 280 nm et de la conductivité, mais puisque Max fournit des rendements supérieurs à c-Myc et que la protéine possède deux tyrosines plutôt qu'une seule, l'enregistreur est réglé à 1.0 V (100% de sensibilité) pour les deux paramètres.

Le chargement de la colonne effectué, la colonne est lavée à l'aide d'un front de 25 ml de 10% v/v de Tampon B à raison de 2,5 ml/min. Un gradient linéaire est par la suite réalisé, allant de 10 à 35% v/v B/(A+B), à raison de 2,5 ml/min et l'élution est collectée à l'aide de fraction de 2,5 ml dans des tubes de polystyrène. Le point de chute de Max en solution avec 50 mM acétate de sodium pH 5,00 a été déterminé à 1,25 M NaCl soit environ 25% v/v B/(A+B). La protéine est éluee avec une pureté de plus de 95 %, mais au besoin les fractions intéressantes peuvent être diluées par deux à l'aide de Tampon A et rechargées sur la colonne pour une deuxième élution fournissant une pureté estimée à plus de 99%. La protéine est par la suite dessalée et lyophilisée avant d'être aliquotée selon les mêmes procédures que c-Myc (voir section 10.4).

12.PRODUCTION DES B-HLHS DE C-MYC ET MAX

Pour le bénéfice d'études complémentaires à ce travail, nous avons également procédé au clonage, à l'expression et à la purification des domaines b-HLH de c-Myc et de Max. Tout comme les domaines LZ, les domaines b-HLH des protéines c-Myc et Max représentent des outils moléculaires importants pour l'investigation et la compréhension des comportements associés à la reconnaissance moléculaire de ces deux partenaires de transcription et de transrépression. Bien que le présent ouvrage ne fasse aucune mention de travaux réalisés à partir de ces fragments de protéine, il est important de mentionner que leur disponibilité a été rendue possible par des efforts déployés dans le cadre de ce projet de maîtrise.

12.1. Clonage et expression des b-HLHs de c-Myc et Max

Tout comme pour les domaines b-HLH-LZ, les gènes encodant les domaines b-HLH de c-Myc et Max ont été amplifiés par PCR et clonés dans des vecteurs pET-3a, entre les sites NdeI et BamHI. Les vecteurs ont été respectivement transformés dans les souches C+ pour c-Myc et pLysS pour Max. Leur expression produit des résultats similaires à ceux des domaines b-HLH-LZ entiers, à la différence près que pour Max une concentration deux fois supérieure d'IPTG est nécessaire pour induire la surexpression et ce à une température optimale de 33.5°C pendant 3 heures. Le fragment b-HLH de c-

Myc est pour sa part induit de façon constitutive sous forme de corps d'inclusion tout comme son parent b-HLH-LZ, à 37°C pendant un minimum de 12 heures.

12.2. Purification du b-HLH de Max

La purification du transcrit de Max s'effectue de la même façon que pour le fragment b-HLH-LZ, excluant la dilution 1:1 avec le mélange des Tampon A et B avant le chargement de la colonne et le lavage à l'aide d'un front de 10% de Tampon B. La protéine est éluée en utilisant un gradient de force ionique de 0 à 30% v/v B/(A+B).

12.3. Purification du b-HLH de c-Myc

Le fragment b-HLH de c-Myc a représenté en soi un nouveau défi, puisque l'élimination du domaine LZ entraîne d'une part une augmentation du point isoélectrique de la protéine ($pI \sim 10.81$) mais de plus, la troncation du domaine LZ exalte la tendance agrégative et les capacités interactives de la protéine avec de nouveaux partenaires. Ce faisant la purification exige un protocole différent de chromatographie sur colonne cationique. L'utilisation des conditions déterminées pour le b-HLH-LZ entraîne la protéine dans de multiples interactions avec des partenaires très variés et ne présente donc pas une solution de purification tel quelle. Les conditions permettent toutefois

d'éliminer un certain nombre de partenaires indésirables et c'est pourquoi l'ensemble des étapes de lyse et de précipitation sont conservées telles quelles.

L'élution à partir des conditions d'échange cationique standard s'est révélé un échec, causant l'étalement de la protéine tout au long du gradient ionique. Malgré de nombreux essais, des dénaturants comme l'urée et l'hydrochlorure de guanidine ne parvinrent pas à empêcher les interactions non-spécifiques du b-HLH de c-Myc avec lui-même et ses congénères. La protéine laissée trop longtemps dans ces conditions finit par précipiter au fond des tubes collecteurs, et l'unique façon de resolubiliser les agrégats (de forme fibrillaire) consiste à utiliser un tampon à base de dénaturants et de pH acide $\lll 5,00$. Nous avons donc préconisé d'empêcher l'ionisation des groupements acides au sein des protéines en solution en abaissant drastiquement le pH des tampons en deçà des pKas des acides glutamiques (pI ~ 4.42) et aspartiques (pI ~ 3.9). Pour ce faire, nous avons remplacé l'acétate de sodium à la base des tampons d'élution par du phosphate de manière à obtenir une capacité tampon adaptée pour une plage environnante de pH 2,80. Dans les faits, les produits utilisés furent de l'acide phosphorique H_3PO_4 et du phosphate de sodium monobasique NaH_2PO_4 combinés l'un à l'autre pour obtenir un pH de 2,80 sans influencer la force ionique. Les tampons sont également enrichis de 2 M urée pour favoriser la dénaturation des complexes non-spécifiques formés à partir d'interactions hydrophobes.

Le gradient est réalisé de façon linéaire à l'aide de Tampon A(2.8)(40 mM NaH_2PO_4 , 10 mM H_3PO_4 , 2 M Urée, pH 2,80 ajusté avec 10 N H_3PO_4 et/ou 10 N NaOH) et

B(2.8)(Tampon A(2.8) + 3 M NaCl), et l'élution est effectuée sur colonnes HiTrapTM HP Sepharose SP selon les mêmes paramètres que pour les autres protocoles à l'exception du gradient qui est réalisé de 0 à 100% v/v B/(A+B) en 200 ml (0,5%/ml). Le point de chute du b-HLH de c-Myc dans ces conditions est estimé à 60% v/v B/(A+B) soit environ 1.8 M NaCl. La purification obtenue est d'une qualité estimable à environ 90% par SDS-PAGE, et des ajustements devraient permettre de relever cette barre au même rang que les purifications des autres membres de cette famille de protéine.

BIBLIOGRAPHIE

- ADHIKARY, S., et EILERS, M. (2005) Transcriptional regulation and transformation by Myc proteins. *Nature Reviews Mol Cell Biol.* **6**: 635-645
- ALBERT, T., WELLS, J., FUNK, J.O., PULLNER, A., RASCHKE, E.E., STELZER, G., MEISTERERNST, M., FARNHAM, P.J., et EICK, D. (2001) The chromatin structure of the dual c-myc promoter P1/P2 is regulated by separate elements. *J Biol Chem.* **276**: 20482-20490 .
- ALITALO, K., RAMSAY, G., BISHOP, J.M., PFEIFER, S.O., COLBY, W.W., et LEVINSON, A.D. (1983) Identification of nuclear proteins encoded by viral and cellular myc oncogenes. *Nature.* **306**: 274-277
- AMATI, B., BROOKS, M.W., LEVY, N., LITTLEWOOD, T.D., EVAN, G.I., et LAND, H. (1993) Oncogenic activity of the c-Myc protein requires dimerization with Max. *Cell.* **72**: 233-245
- AMATI, B., FRANK, S.R., DONJERKOVIC, D., et TAUBERT, S. (2001) Function of the c-Myc oncoprotein in chromatin remodeling and transcription. *Biochim Biophys Acta.* **1471**: 135-145

- ANFINSEN, C.B. (1973) Principles that govern the folding of protein chains. *Science*.
181: 223-230
- BANERJEE, A., HU. J., et GOSS. D.J. (2006) Thermodynamics of protein-protein interactions of cMyc, Max, and Mad: effect of polyions on protein dimerization. *Biochemistry*. **45(7)**: 2333-2338
- BENTLEY D.L., BROWN W.L., et GROUDINE M. (1989) Accurate, TATA box-dependent polymerase III transcription from promoters of the c-myc gene in injected *Xenopus* oocytes. *Genes Dev*. **3**: 1179-1189
- BERNARD, O., CORY, S., GERONDAKIS, S., WEBB, E., et ADAMS, J.M. (1983) Sequence of the murine and human cellular myc oncogenes and two modes of myc transcription resulting from chromosome translocation in B lymphoid tumours. *EMBO J*. **2**: 2375-2383
- BEYCHOK, S. (1966) Circular dichroism of biological macromolecules. *Science*.
154(754): 1288-1299
- BLACKWOOD, E., et EISENMAN, R. N. (1991) Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science*. **251**:
1211-1217

- BODESCOT, M., et BRISON, O. (1996) Characterization of new human c-myc mRNA species produced by alternative splicing. *Gene*. **174**: 115-120
- CANTOR, C.R., et SCHIMMEL, P.R. (1980) Biophysical Chemistry, *W.H. Freeman and Company., New York, USA.*
- COLBY, W.W., CHEN, E.Y., SMITH, D.H., et LEVINSON, A.D. (1983) Identification and nucleotide sequence of a human locus homologous to the v-myc oncogene of avian myelocytomatosis virus MC29. *Nature*. **301**: 722-725
- COWIESON, N.P., MILES, A.J., ROBIN, G., FORWOOD, J.K., KOBE, B., MARTIN, J.L., et WALLACE, B.A. (2008) Evaluating protein:protein complex formation using synchrotron radiation circular dichroism spectroscopy. *Proteins*. **70(4)**: 1142-1146
- DESJARDINS, E., et HAY, N. (1993) Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-myc promoters. *Mol Cell Biol*. **13**: 5710-5724
- FIEBER, W., SCHNEIDER, M.L., MATT, T., KRAUTLER, B., KONRAT, R., et BISTER, K. (2001) Structure, function, and dynamics of the dimerization and DNA-binding domain of oncogenic transcription factor v-Myc. *J Mol Biol*. **307**: 1395-1410

- GAGNON, M.C., WILLIAMS, M., DOUCET, A., et BEAUREGARD, M. (2000) Replacement of tyr62 by trp in the designer protein milk bundle-1 results in significant improvement of conformational stability. *FEBS Lett.* **484**: 144-118
- GRANDORI, C., COWLEY, S.M., JAMES, L.P., et EISENMAN, R.N. (2000) The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu Rev Cell Dev Biol.* **16**: 653-699
- GRATZER, W.B., et COWBURN, D.A. (1969) Optical activity of biopolymers. *Nature.* **222(5192)**: 426-431
- GREENFIELD, N., et FASMAN, G.D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry.* **8(10)**: 4108-4116
- HIRVONEN, H., HUKKANEN, V., SALMI, T.T., PELLINIEMI, T.T., et ALITALO, R. (1993) L-myc and N-myc in hematopoietic malignancies. *Leuk Lymphoma.* **11**: 197-205
- HU, J., BANERJEE, A., et GOSS, D.J. (2005) Assembly of b/HLH/z proteins c-Myc, Max, and Mad1 with cognate DNA: importance of protein-protein and protein-DNA interactions. *Biochemistry.* **44(35)**: 11855-11863

HURLIN, P.J., QUÉVA, C., KOSKINEN, P.J., STEINGRÍMSSON, E., AYER, D.E., COPELAND, N.G., JENKINS, N.A., et EISENMAN, R.N. (1995) Mad3 and Mad4: novel Max-interacting transcriptional repressors that suppress c-myc dependent transformation and are expressed during neural and epidermal differentiation. *EMBO J.* **14(22)**: 5646-5659

JOHNSON, W.C. JR. (1988) Secondary structure of proteins through circular dichroism spectroscopy. *Annu Rev Biophys Biophys Chem.* **17**: 145-166

KAGAYA, S., KITANAKA, C., NOGUCHI, K., MOCHIZUKI, T., SUGIYAMA, A., ASAI, A., YASUHARA, N., EGUCHI, Y., TSUJIMOTO, Y., et KUCHINO, Y. (1997) A functional role for death proteases in s-Myc- and c-Myc-mediated apoptosis. *Mol Cell Biol.* **17**: 6736-6745

KELLY, S.M., JESS, T.J., et PRICE, N.C. (2005) How to study proteins by circular dichroism. *Biochim Biophys Acta.* **1751(2)**: 119-139

KLEIN, E., KLEIN, G., NADKARNI, J.S., NADKARNI, J.J., WIGZELL, H., et CLIFFORD, P. (1968) Surface IgM-kappa specificity on a Burkitt lymphoma cell in vivo and in derived culture lines. *Cancer Res.* **28**: 1300-1310

KLEMPNAUER, K.H., GONDA, T.J., et BISHOP, J.M. (1982) Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell*. **31**: 453-463

KOHLER, J.J., METALLO, S.J., SCHNEIDER, T.L., et SCHEPARTZ, A. (1999) DNA specificity enhanced by sequential binding of protein monomers. *Proc Natl Acad Sci USA*. **96(21)**: 11735-11739

LAVIGNE, P., CRUMP, M.P., GAGNÉ, S.M., HODGES, R.S., KAY, C.M., et SYKES, B.D. (1998) Insights into the mechanism of heterodimerization from the ¹H-NMR solution structure of the c-Myc-Max heterodimeric Leucine Zipper. *J Mol Biol*. **281**: 165-181

LAVIGNE, P., KONDEJEWSKI, L.H., HOUSTON JR., M.E., SÖNNICHSEN, F.D., LIX, F.D., SYKES, B.D. HODGES, R.S., et KAY, C.M. (1995) Preferential heterodimeric parallel coiled-coil formation by synthetic Max and c-Myc Leucine zippers: A description of putative electrostatic interactions responsible for the specificity of heterodimerization. *J Mol Biol*. **254**: 505-520

LAVIGNE, P., SÖNNICHSEN F.D., KAY C.M., et HODGES, R.S. (1996) Interhelical salt bridges, coiled-coil stability, and specificity of dimerization. *Science*. **271**: 1136-1138

LI, Z., CALCAR, S.V., QU, C., CAVENEE, W.K., ZHANG, M.Q., et REN, B. (2003)

A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells.

Proc Natl Acad Sci. **100**: 8164-8169

MAO, D., WATSON, J., YAN, P., BARSYTE-LOVEJOY, D., KHOSRAVI, F.,

WONG, W., FARNHAM, P., HUANG, T. et PENN, L. (2003) Analysis of Myc

Bound Loci Identified by CpG Island Arrays Shows that Max Is Essential for Myc-

Dependent Repression. *Curr Biol.* **13**: 882-886

MARIADASON, J.M., COURTNEY, N.M., L'ITALIEN, K.E., ZHUANG, M.,

SMARTT, H.J.M., HEERDT, B.G., YANG, W., CORNER, G.A., WILSON, A.J.,

ARANGO, D., et AUGENLICHT, L.H. (2005) Gene expression profiling of

intestinal epithelial cell maturation along the crypt-villus axis. *Gastroenterology.*

128: 1081-1088

MCCMAHON, S.B., WOOD, M.A., et COLE, M.D. (2000) The essential cofactor

TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. *Mol Cell Biol.* **20**:

556-562

MEULIA, T., KRUMM, A., SPENCER, C., et GROUDINE, M. (1992) Sequences in the

human c-myc P2 promoter affect the elongation and premature termination of

transcripts initiated from the upstream P1 promoter. *Mol Cell Biol.* **12**: 4590-4600

MILEV, S., GORFE, A.A., KARSHIKOFF, A., CLUBB, R.T., BOSSHARD, H.R., et JELESAROV, I. (2003) Energetics of sequence-specific protein-DNA association: conformational stability of the DNA binding domain of integrase Tn916 and its cognate DNA duplex. *Biochemistry*. **42**: 3492-3502

MILEV, S., GORFE, A.A., KARSHIKOFF, A., CLUBB, R.T., BOSSHARD, H.R., et JELESAROV, I. (2003) Energetics of sequence-specific protein-DNA association: binding of integrase Tn916 to its target DNA. *Biochemistry*. **42**: 3481-3491

MILLER, J.C., HOLMES, M.C., WANG, J., GUSCHIN, D.Y., LEE, Y.L., RUPNIEWSKI, I., BEAUSEJOUR, C.M., WAITE, A.J., WANG, N.S., KIM, K.A., GREGORY, P.D., PABO, C.O., et REBAR, E.J. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*. **25(7)**: 778-785

MONTAGNE, M., NAUD, J.F., MCDUFF, F.O., et LAVIGNE, P. (2005) Towards the elucidation of the structural determinants responsible for the molecular recognition between Mad1 and Max. *Biochemistry*. **44**: 12860-12869

MUHLE-GOLL, C., NILGES, M., et PASTORE, A. (1995) The leucine zippers of the HLH-LZ proteins Max and c-Myc preferentially form heterodimers. *Biochemistry*. **34**: 13554-13564

- MYERS, J.K., PACE, C.N., et SCHOLTZ, J.M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4(10)**: 2138-2148
- NAIR, S.K., et BURLEY, S.K. (2003) X-ray structures of Myc-Max and Mad-Max recognizing DNA: Molecular bases of regulation by proto-oncogenic transcription factors. *Cell.* **112**: 193-205
- NAUD, J.F., GAGNON, F., WELLINGER, R., CHABOT, B., et LAVIGNE, P. (2003) Improving the Stability of the LZ domain of Max increases the stability of its B-HLH-LZ:E-Box complex. *J Mol Biol.* **326**: 1577-1595
- NAUD, J.F., MCDUFF, F.O., SAUVÉ, S., MONTAGNE, M., WEBB, B.A., SMITH, S.P., CHABOT, B., et LAVIGNE, P. (2005) Structural and thermodynamical characterization of the complete p21 gene product of Max. *Biochemistry.* **44**: 12746-12758
- NIKIFOROV, M.A., CHANDRIANI, S., PARK, J., KOTENKO, I., MATHEOS, D., JOHNSON, A., MCMAHON, S.B., et COLE, M.D. (2002) TRRAP-dependent and TRRAP-independent transcriptional activation by Myc family oncoproteins. *Mol Cell Biol.* **22**: 5054-5063

- OAKLEY, M.T., et HIRST, J.D. (2006) Charge-transfer transitions in protein circular dichroism calculations. *J Am Chem Soc.* **128(38)**: 12414-12415
- OAKLEY, M.T., BULHELLER, B.M., et HIRST, J.D. (2006) First-principles calculations of protein circular dichroism in the far-ultraviolet and beyond. *Chirality.* **18(5)**: 340-347
- PEUKERT, K., STALLER, P., SCHNEIDER, A., CARMICHAEL, G., HANEL, F., et EILERS, M. (1997) An alternative pathway for gene regulation by Myc. *EMBO J.* **16**: 5672–5686
- PRITSKER, M., JONES, P., BLUMENTHAL, R., et SHAI, Y. (1998) A synthetic all D-amino acid peptide corresponding to the N-terminal sequence of HIV-1 gp41 recognizes the wild-type fusion peptide in the membrane and inhibits HIV-1 envelope glycoprotein-mediated cell fusion. *Proc Natl Acad Sci USA.* **95(13)**: 7287-7292
- PRIVALOV, P.L. (1996) Intermediate states in protein folding. *J Mol Biol.* **258**: 707-725
- PULLNER, A., MAUTNER, J., ALBERT, T., et EICK, D. (1996) Nucleosomal structure of active and inactive c-myc genes. *J Biol Chem.* **271**: 31452-31457

- RAMSAY, G.M., et HAYMAN, M.J. (1982) Isolation and biochemical characterization of partially transformation-defective mutants of avian myelocytomatosis virus strain MC29: localization of the mutation to the myc domain of the 110,000-dalton gag-myc polypeptide. *J Virol.* **41**: 745-753
- REINHARDT, F., et SOEDER, H. (1997) Atlas des mathématiques, *La Pochothèque*, Librairie Générale Française. : 316-317
- ROTTMANN, S., et LUSCHER, B. (2006) The Mad side of the Max network: antagonizing the function of Myc and more. *Curr Top Microbiol Immunol.* **302**: 63-122
- SAKAMURO, D., et PRENDERGAST, G.C. (1999) New Myc-interacting proteins: a second Myc network emerges. *Oncogene.* **18**: 2942-2954
- SAUVÉ, S., NAUD, J.F., et LAVIGNE, P. (2007) The Mechanism of Discrimination between Cognate and Non-Specific DNA by Dimeric b/HLH/LZ Transcription Factors. *J Mol Biol.* **365**: 1163–1175
- SAUVÉ, S., TREMBLAY, L., et LAVIGNE, P. (2004) The NMR Solution Structure of a Mutant of the Max b/HLH/LZ Free of DNA: Insights into the Specific and Reversible DNA Binding Mechanism of Dimeric Transcription Factors. *J Mol Biol.* **342**: 813-832

STONE, J., DE LANGE, T., RAMSAY, G., JAKOBOVITS, E., BISHOP, J.M.,
VARMUS, H., et LEE, W. (1987) Definition of regions in human c-myc that are
involved in transformation and nuclear localization. *Mol Cell Biol.* **7**: 1697-1709

SUGIYAMA, A., KUME, A., NEMOTO, K., LEE, S.Y., ASAMI, Y., NEMOTO, F.,
NISHIMURA, S., et KUCHINO, Y. (1989) Isolation and characterization of s-
myc, a member of the rat myc gene family. *Proc Natl Acad Sci USA.* **86**: 9144-
9148

SZCZEPEK, M., BRONDANI, V., BUCHEL, J., SERRANO, L., SEGAL, D.J., et
CATHOMEN, T. (2007) Structure-based redesign of the dimerization interface
reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol.* **25(7)**: 786-793

VELAZQUEZ-CAMPOY, A., LEAVITT, S.A., et FREIRE, E. (2004) Characterization
of protein-protein interactions by isothermal titration calorimetry. *Methods Mol
Biol.* **261**: 35-54

VENNSTROM, B., SHEINESS, D., ZABIELSKI, J., et BISHOP, J.M. (1982) Isolation
and characterization of c-myc, a cellular homolog of the oncogene (v-myc) of
avian myelocytomatosis virus strain 29. *J Virol.* **42**: 773-779

VOLLHARDT, K.P.C., et SCHORE, N.E. (1995) *Traité de chimie organique*, 2^e édition,

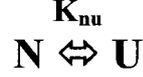
De Boeck-Wesmael S.A., Bruxelles. : 136-160, 325-329, 387-392

VUILLARD, L., BRAUN-BRETON, C., et RABILLOUD, T. (1995) Non-detergent sulphobetaines: a new class of mild solubilization agents for protein purification. *Biochem J.* **305**: 337-343

WATT, R., STANTON, L.W., MARCU, K.B., GALLO, R.C., CROCE, C.M., et ROVERA, G. (1983) Nucleotide sequence of cloned cDNA of human c-Myc oncogene. *Nature.* **303**: 725–728

ANNEXE I

Équations de Gibbs-Helmholtz pour un système monomérique à 2 états



$$K_{nu} = \frac{[\mathbf{U}]}{[\mathbf{N}]} \quad \text{si: } [\mathbf{N}] = [\mathbf{T}] - [\mathbf{U}]$$

$$K_{nu} = \frac{[\mathbf{U}]}{[\mathbf{T}] - [\mathbf{U}]} \quad \rightarrow \quad K_{nu} [\mathbf{T}] - K_{nu} [\mathbf{U}] = [\mathbf{U}]$$

$$K_{nu} [\mathbf{T}] = [\mathbf{U}] + K_{nu} [\mathbf{U}] \quad \rightarrow \quad K_{nu} [\mathbf{T}] = [\mathbf{U}] (1 + K_{nu})$$

$$[\mathbf{U}] = \frac{K_{nu} [\mathbf{T}]}{(1 + K_{nu})}$$

$$P_U = \frac{[\mathbf{U}]}{[\mathbf{T}]} \quad \rightarrow \quad P_U(T) = \frac{K_{nu}(T) [\mathbf{T}]}{(1 + K_{nu}(T)) [\mathbf{T}]} = \frac{K_{nu}(T)}{(1 + K_{nu}(T))}$$

$$K_{nu}(T) = e^{\left(\frac{-\Delta G_{nu}(T)}{RT} \right)}$$

$$\Delta G = \Delta H - T\Delta S + \Delta Cp$$

$$\Delta G_{nu}(T) = \Delta H_{nu}(T_{nu}) - \frac{T\Delta H_{nu}(T_{nu})}{T_{nu}} + \Delta Cp_{nu} \left(T - T_{nu} - T \ln \left(\frac{T}{T_{nu}} \right) \right)$$

$$\Delta G_{nu}(T) = \Delta H_{nu}(T_{nu}) \left(1 - \frac{T}{T_{nu}} \right) + \Delta Cp_{nu} \left(T - T_{nu} - T \ln \left(\frac{T}{T_{nu}} \right) \right)$$

ANNEXE II

Dérivation de P_U à partir de la fonction de partition d'un système homodimérique à 2 états.

$$Q = 2[N] + [U]$$

$$K_D = \frac{[U][U]}{[N]} = \frac{[U]^2}{[N]} \quad \text{et} \quad P_U = \frac{[U]}{[T]} \quad \text{et} \quad [T] = [U] + 2[N]$$

$$P_U = \frac{[U]}{([U] + 2[N])}$$

$$[U] = \sqrt{K_D [N]}$$

$$P_U = \frac{\sqrt{K_D [N]}}{(\sqrt{K_D [N]} + 2[N])}$$

$$P_U = \frac{\sqrt{K_D} \sqrt{[N]}}{(\sqrt{K_D} \sqrt{[N]} + 2[N])}$$

$$P_U = \frac{\sqrt{K_D} \sqrt{[N]}}{\sqrt{[N]}(\sqrt{K_D} + 2\sqrt{[N]})}$$

$$P_U = \frac{\sqrt{K_D}}{(\sqrt{K_D} + 2\sqrt{[N]})}$$

Dérivation de P_N à partir de la fonction de partition d'un système homodimérique à 2 états.

$$Q = 2[N] + [U]$$

$$K_D = \frac{[U][U]}{[N]} = \frac{[U]^2}{[N]} \quad \text{et} \quad P_N = \frac{[N]}{[T]} \quad \text{et} \quad [T] = [U] + 2[N]$$

$$P_N = \frac{[N]}{([U] + 2[N])}$$

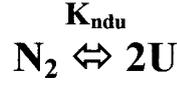
$$[N] = \frac{[U]^2}{K_D}$$

$$P_N = \frac{\frac{[U]^2}{K_D}}{\left([U] + 2 \frac{[U]^2}{K_D}\right)}$$

$$P_N = \frac{[U] \left(\frac{[U]}{K_D}\right)}{[U] \left(1 + 2 \frac{[U]}{K_D}\right)}$$

$$P_N = \frac{\frac{[U]}{K_D}}{\left(1 + 2 \frac{[U]}{K_D}\right)}$$

Équations de Gibbs-Helmholtz pour un système homodimérique à 2 états



$$K_{ndu} = \frac{[\text{U}][\text{U}]}{[\text{N}_2]} \quad \text{si: } [\text{N}_2] = \frac{[\text{T}] - [\text{U}]}{2}$$

$$K_{ndu} = \frac{[\text{U}]^2}{\left(\frac{[\text{T}] - [\text{U}]}{2}\right)} \quad \rightarrow \quad K_{ndu} [\text{T}] - K_{ndu} [\text{U}] = 2[\text{U}]^2$$

$$0 = 2[\text{U}]^2 + K_{ndu} [\text{U}] - K_{ndu} [\text{T}]$$

$$[\text{U}] = \frac{-K_{ndu} + \sqrt{K_{ndu}^2 - 4 \cdot 2 \cdot -K_{ndu} [\text{T}]}}{4}$$

$$P_U = \frac{[\text{U}]}{[\text{T}]} \quad \rightarrow \quad P_U(T) = \frac{-K_{ndu}(T_{ndu}) + \sqrt{K_{ndu}(T_{ndu})^2 + 8K_{ndu}(T_{ndu})[\text{T}]}}{4[\text{T}]}$$

$$K_{ndu}(T) = e^{\left(\frac{-\Delta G_{ndu}(T)}{RT}\right)}$$

$$\Delta G = \Delta H - T\Delta S - RT \ln K + \Delta Cp$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - \frac{T\Delta H_{ndu}(T_{ndu})}{T_{ndu}} - RT \ln K_{ndu}(T_{ndu}) + \Delta Cp_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln K_{ndu}(T_{ndu}) \right) + \Delta Cp_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

$$\ln K_{ndu}(T_{ndu}) \approx \frac{\ln[N]^{n-1} \cdot n^{n/2} \cdot \prod_{i=1}^k (m_i)^{m_i}}{(\sqrt{n} + 1)^{n-1}}$$

Où : $[N]$ = le nombre possible de complexe Natif (*Associés*)

Et : n = le nombre total de mères dans le complexe
 m_i = le nombre de copies de chaque différente mère

e.g. $A_n \Leftrightarrow mD$

$A_2 \Leftrightarrow 2A$	$n = 2$	$m_1 = 2$	
$AB \Leftrightarrow 1A + 1B$	$n = 2$	$m_1 = 1$	$m_2 = 1$
$AB_2 \Leftrightarrow 1A + 2B$	$n = 3$	$m_1 = 1$	$m_2 = 2$

$$\ln K_{ndu}(T_{ndu})_{HOMO} \approx \frac{\ln[N]^1 \cdot 2^1 \cdot 2^2}{(\sqrt{2} + 1)^1} = \ln[N] \cdot \frac{8}{2,4142}$$

$$\ln K_{ndu}(T_{ndu})_{HOMO} \approx \ln[N] \cdot 3,3137$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln([N] \cdot 3,3137) \right) + \Delta C p_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

$$\text{si : } [N] = \frac{[T]}{2}$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln \left(\frac{3,3137[T]}{2} \right) \right) + \Delta C p_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln(1,6569[T]) \right) + \Delta C p_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

ANNEXE III

Dérivation de P_U à partir de la fonction de partition d'un système hétérodimérique à 2 états.

$$Q = [N1N2] + [U1] + [U2]$$

$$K_D = \frac{[U1][U2]}{[N1N2]} \quad \text{et} \quad P_{U1} = \frac{[U1]}{[T1]} \quad \text{et} \quad [T1] + [T2] = [U1] + [U2] + 2[N1N2]$$

$$P_{U1} = \frac{[U1]}{([U1] + [U2] + 2[N1N2] - [T2])}$$

$$[U1] = \frac{K_D [N1N2]}{[U2]}$$

$$P_{U1} = \frac{\frac{K_D [N1N2]}{[U2]}}{\left(\frac{K_D [N1N2]}{[U2]} + [U2] + 2[N1N2] - [T2] \right)}$$

Équations de Gibbs-Helmholtz pour un système hétérodimérique à 2 états



$$K_{ndu} = \frac{[U^1][U^2]}{[N^1 N^2]} \quad \text{si : } [N^1 N^2] = [T^1] - [U^1]$$

$$K_{ndu} = \frac{[U^1][U^2]}{[T^1] - [U^1]} \quad \rightarrow \quad K_{ndu} [T^1] - K_{ndu} [U^1] = [U^1][U^2]$$

$$\text{si : } [U^2] = [U^1] \quad \rightarrow \quad K_{ndu} [T^1] - K_{ndu} [U^1] = [U^1]^2$$

À concentrations égales de partenaires, le terme $[T^1]$ équivaut ici à 100% de $[N]$

$$0 = [U^1]^2 + K_{ndu} [U^1] - K_{ndu} [T^1]$$

$$[U^1] = \frac{-K_{ndu} \pm \sqrt{K_{ndu}^2 + 4K_{ndu} [T^1]}}{2}$$

$$P_{U^1} = \frac{[U^1]}{[T]} \quad \rightarrow \quad P_{U^1}(T) = \frac{-K_{ndu}(T_{ndu}) + \sqrt{K_{ndu}(T_{ndu})^2 + 4K_{ndu}(T_{ndu})[T^1]}}{2[T^1]}$$

$$K_{ndu}(T) = e^{\left(\frac{-\Delta G_{ndu}(T)}{RT}\right)}$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln K_{ndu}(T_{ndu}) \right) + \Delta C p_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

$$\ln K_{ndu}(T_{ndu}) \approx \frac{\ln[N]^{n-1} \cdot n^{n/2} \cdot \prod_{i=1}^k (m_i)^{m_i}}{(\sqrt{n} + 1)^{n-1}}$$

$$AB \Leftrightarrow 1A + 1B \quad n = 2 \quad m_1 = 1 \quad m_2 = 1$$

$$\ln K_{ndu}(T_{ndu})_{HETERO} \approx \frac{\ln[N]^1 \cdot 2^1 \cdot (1^1 \cdot 1^1)}{(\sqrt{2} + 1)^1} = \ln[N] \cdot \frac{2}{2,4142}$$

$$\ln K_{ndu}(T_{ndu})_{HETERO} \approx \ln[N] \cdot 0,8284$$

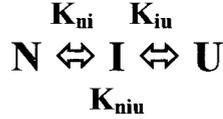
$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln([N] \cdot 0,8284) \right) + \Delta C p_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

$$\text{si: } [N] = \frac{[T]}{2}$$

$$\Delta G_{ndu}(T) = \Delta H_{ndu}(T_{ndu}) - T \left(\frac{\Delta H_{ndu}(T_{ndu})}{T_{ndu}} + R \ln(0,4142[T]) \right) + \Delta C p_{ndu} \left(T - T_{ndu} - T \ln \left(\frac{T}{T_{ndu}} \right) \right)$$

ANNEXE IV

Équations de Gibbs-Helmholtz pour un système monomérique à 3 états.



$$K_{ni} = \frac{[\mathbf{I}]}{[\mathbf{N}]} \quad \& \quad K_{iu} = \frac{[\mathbf{U}]}{[\mathbf{I}]}$$

$$K_{niu} = \frac{[\mathbf{U}]}{[\mathbf{N}]} = K_{ni} \cdot K_{iu}$$

$$Q = 1 + \sum_{i=1}^n K_i = \frac{[\mathbf{N}] + [\mathbf{I}] + [\mathbf{U}]}{[\mathbf{N}]} = 1 + \frac{[\mathbf{I}]}{[\mathbf{N}]} + \frac{[\mathbf{U}]}{[\mathbf{N}]} = 1 + K_{ni} + K_{ni} K_{iu}$$

$$P_N = \frac{[\mathbf{N}]}{[\mathbf{T}]} = \frac{1}{Q} \quad \rightarrow \quad P_N(T) = \frac{1}{(1 + K_{ni}(T) + K_{ni}(T) \cdot K_{iu}(T))}$$

$$P_I = \frac{[\mathbf{I}]}{[\mathbf{T}]} = \frac{K_{ni}}{Q} \quad \rightarrow \quad P_I(T) = \frac{K_{ni}(T)}{(1 + K_{ni}(T) + K_{ni}(T) \cdot K_{iu}(T))}$$

$$K_{ni}(T) = e^{\left(\frac{-\Delta G_{ni}(T)}{RT}\right)}$$

$$\Delta G_{ni}(T) = \Delta H_{ni}(T_{ni}) - \frac{T \Delta H_{ni}(T_{ni})}{T_{ni}} + \Delta C p_{ni} \left(T - T_{ni} - T \ln \left(\frac{T}{T_{ni}} \right) \right)$$

$$\Delta G_{ni}(T) = \Delta H_{ni}(T_{ni}) \left(1 - \frac{T}{T_{ni}} \right) + \Delta C p_{ni} \left(T - T_{ni} - T \ln \left(\frac{T}{T_{ni}} \right) \right)$$

$$P_U = \frac{[U]}{[T]} = \frac{K_{ni} \bullet K_{iu}}{Q} \rightarrow P_U(T) = \frac{K_{ni}(T) \bullet K_{iu}(T)}{(1 + K_{ni}(T) + K_{ni}(T) \bullet K_{iu}(T))}$$

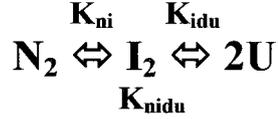
$$K_{iu}(T) = e^{\left(\frac{-\Delta G_{iu}(T)}{RT}\right)}$$

$$\Delta G_{iu}(T) = \Delta H_{iu}(T_{iu}) - \frac{T\Delta H_{iu}(T_{iu})}{T_{iu}} + \Delta Cp_{iu} \left(T - T_{iu} - T \ln \left(\frac{T}{T_{iu}} \right) \right)$$

$$\Delta G_{iu}(T) = \Delta H_{iu}(T_{iu}) \left(1 - \frac{T}{T_{iu}} \right) + \Delta Cp_{iu} \left(T - T_{iu} - T \ln \left(\frac{T}{T_{iu}} \right) \right)$$

ANNEXE V

Équations de Gibbs-Helmholtz pour un système homodimérique à 3 états



$$K_{ni} = \frac{[I_2]}{[N_2]} \quad \& \quad K_{idu} = \frac{[U][U]}{[I_2]} = \frac{[U]^2}{[I_2]}$$

$$K_{nidu} = \frac{[U][U]}{[N]} = \frac{[U]^2}{[N]} = K_{ni} \cdot K_{idu}$$

[T₁] = concentration totale de monomère

$$K_{ni} \cdot K_{idu} = \frac{[U]^2}{[N_2]} \quad \text{si : } [N_2] = \frac{([T_1] - [U] - 2[I_2])}{2}$$

$$K_{ni} \cdot K_{idu} = \frac{2[U]^2}{([T_1] - [U] - 2[I_2])} \quad \text{si : } [I_2] = \frac{[U]^2}{K_{idu}}$$

$$K_{ni} \cdot K_{idu} \cdot [T_1] - K_{ni} \cdot K_{idu} \cdot [U] - 2K_{ni} [U]^2 = 2[U]^2$$

$$0 = (2K_{ni} + 2)[U]^2 + K_{nidu} \cdot [U] - K_{nidu} \cdot [T_1]$$

$$0 = \frac{-K_{nidu} + \sqrt{(K_{nidu}^2 - 4 \cdot (2K_{ni} + 2) \cdot -K_{nidu} \cdot [T_1])}}{2 \cdot (2K_{ni} + 2)}$$

$$\text{si : } P_U = \frac{[U]}{[T_1]} \quad \rightarrow \quad P_U = \frac{-K_{nidu} + \sqrt{(K_{nidu}^2 + 8K_{ni} \cdot K_{nidu} \cdot [T_1] + 8K_{nidu} \cdot [T_1])}}{(4K_{ni} + 4) \cdot [T_1]}$$

$$K_{mi}(T) = e^{\left(\frac{-\Delta G_{mi}(T)}{RT}\right)}$$

$$\Delta G_{mi}(T) = \Delta H_{mi}(T_{mi}) - \frac{T\Delta H_{mi}(T_{mi})}{T_{mi}} + \Delta Cp_{mi} \left(T - T_{mi} - T \ln \left(\frac{T}{T_{mi}} \right) \right)$$

$$\Delta G_{mi}(T) = \Delta H_{mi}(T_{mi}) \left(1 - \frac{T}{T_{mi}} \right) + \Delta Cp_{mi} \left(T - T_{mi} - T \ln \left(\frac{T}{T_{mi}} \right) \right)$$

$$K_{idu}(T) = e^{\left(\frac{-\Delta G_{idu}(T)}{RT}\right)}$$

$$\Delta G_{idu}(T) = \Delta H_{idu}(T_{idu}) - T \left(\frac{\Delta H_{idu}(T_{idu})}{T_{idu}} + R \ln K_{idu}(T_{idu}) \right) + \Delta Cp_{idu} \left(T - T_{idu} - T \ln \left(\frac{T}{T_{idu}} \right) \right)$$

$$\ln K_{idu}(T_{idu}) \approx \frac{\ln[I]^{n-1} \cdot n^{n/2} \cdot \prod_{i=1}^k (m_i)^{m_i}}{(\sqrt{n} + 1)^{n-1}}$$

$$A_2 \Leftrightarrow 2A \quad n = 2 \quad m_1 = 2$$

$$\ln K_{idu}(T_{idu})_{HOMO} \approx \frac{\ln[I]^1 \cdot 2^1 \cdot 2^2}{(\sqrt{2} + 1)^1} = \ln[I] \cdot \frac{8}{2,4142}$$

$$\ln K_{idu}(T_{idu})_{HOMO} \approx \ln[I] \cdot 3,3137$$

$$\Delta G_{idu}(T) = \Delta H_{idu}(T_{idu}) - T \left(\frac{\Delta H_{idu}(T_{idu})}{T_{idu}} - R \ln(1,6569[T]) \right) + \Delta Cp_{idu} \left(T - T_{idu} - T \ln \left(\frac{T}{T_{idu}} \right) \right)$$

$$K_{midu}(T) = K_{mi}(T) \bullet K_{idu}(T)$$

$$\text{si: } \frac{[T_1] - [U]}{2} = [N_2] + [I_2] \rightarrow [N_2] = \frac{[T_1] - [U]}{2} - [I_2]$$

$$K_{mi} = \frac{[I_2]}{[N_2]} \rightarrow K_{mi} = \frac{[I_2]}{\left(\frac{[T_1] - [U]}{2} - [I_2]\right)}$$

$$K_{mi} \left(\frac{[T_1] - [U]}{2}\right) - K_{mi} \bullet [I_2] = [I_2]$$

$$K_{mi} \left(\frac{[T_1] - [U]}{2}\right) = [I_2] + K_{mi} \bullet [I_2]$$

$$[I_2] = \frac{K_{mi} \left(\frac{[T_1] - [U]}{2}\right)}{(1 + K_{mi})}$$

$$\text{si: } P_I = \frac{2[I_2]}{[T_1]} \rightarrow P_I = \frac{K_{mi}([T_1] - [U])}{(1 + K_{mi})[T_1]} = \frac{K_{mi}([T_1] - P_U \bullet [T_1])}{(1 + K_{mi})[T_1]} = \frac{K_{mi} \bullet [T_1](1 - P_U)}{(1 + K_{mi})[T_1]}$$

$$P_I = \frac{K_{mi}(1 - P_U)}{(1 + K_{mi})}$$

$$\text{si: } [N_2] = \frac{([T_1] - 2[I_2] - [U])}{2}$$

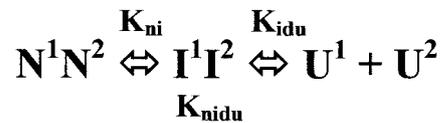
$$\text{si: } P_N = \frac{2[N_2]}{[T_1]} \rightarrow P_N = \frac{([T_1] - 2[I_2] - [U])}{[T_1]}$$

$$P_N = \frac{\left([T_1] - \frac{2P_I[T_1]}{2} - P_U[T_1]\right)}{[T_1]} = 1 - P_I - P_U$$

ANNEXE VI

Équations de Gibbs-Helmholtz pour un système hétérodimérique à 3 états

Il faut ici distinguer entre $K_{\text{réel}}$ et K_{apparent} . Le $K_{\text{réel}}$ représente l'affinité pure de c-Myc pour Max indépendamment de l'interférence occasionnée par les affinités homodimériques. Il doit être extrait d'analyses faites sur des données épurées de leurs interférences, soit par le biais d'une méthode de reconstruction de spectres à partir des expériences fondamentales. Le K_{apparent} est celui qui peut être déterminé à partir des présentes équations, mais il représente une sous-estimation de l'affinité réelle des partenaires l'un pour l'autre.



$$K_{ni} = \frac{[I^1 I^2]}{[N^1 N^2]} \quad \& \quad K_{idu} = \frac{[U^1][U^2]}{[I^1 I^2]}$$

$$K_{nidu} = \frac{[U^1][U^2]}{[N^1 N^2]} = K_{ni} \cdot K_{idu}$$

si : $[U^1] = [U^2]$

...ce qui est vrai seulement dans la mesure où il n'y a aucune interaction de type homodimérique et que les concentrations initiales des partenaires sont les mêmes.

Alors :

$$[U^1] + [U^2] = [U]$$

$$[U^1][U^2] = [U^1]^2 = [U^2]^2 = ([U]/2)^2$$

$[T_1]$ = concentration totale en monomère #1 et #2

$$\text{si : } [N^1 N^2] = \frac{([T_1] - [U^1] - [U^2] - 2[I^1 I^2])}{2} = \frac{([T_1] - [U] - 2[I^1 I^2])}{2}$$

$$K_{ni} \cdot K_{idu} = \frac{\left(\frac{[U]}{2}\right)^2}{[N^1 N^2]} \rightarrow K_{ni} \cdot K_{idu} = \frac{2\sqrt{2}[U]^2}{([T_1] - [U] - 2[I^1 I^2])}$$

$$\text{si : } [I^1 I^2] = \frac{\sqrt{2}[U]^2}{K_{idu}}$$

$$K_{ni} \cdot K_{idu} \cdot [T_1] - K_{ni} \cdot K_{idu} \cdot [U] - 2K_{ni} [U]^2 = 2\sqrt{2}[U]^2$$

$$0 = (2\sqrt{2}K_{ni} + 2\sqrt{2})[U]^2 + K_{midu} \cdot [U] - K_{midu} \cdot [T_1]$$

$$[U] = \frac{-K_{midu} + \sqrt{(K_{midu}^2 - 4 \cdot (2\sqrt{2}K_{ni} + 2\sqrt{2})) \cdot -K_{midu} \cdot [T_1]}}{2 \cdot (2\sqrt{2}K_{ni} + 2\sqrt{2})}$$

$$\text{si : } P_U = \frac{[U]}{[T_1]} \rightarrow P_U = \frac{-K_{midu} + \sqrt{(K_{midu}^2 + 8\sqrt{2}K_{ni} \cdot K_{midu} \cdot [T_1] + 8\sqrt{2}K_{midu} \cdot [T_1])}}{(4\sqrt{2}K_{ni} + 4\sqrt{2}) \cdot [T_1]}$$

$$P_{U^1} = P_{U^2} = \frac{P_U}{2}$$

$$K_m(T) = e^{\left(\frac{-\Delta G_m(T)}{RT}\right)}$$

$$\Delta G_m(T) = \Delta H_m(T_m) - \frac{T\Delta H_m(T_m)}{T_m} + \Delta Cp_m \left(T - T_m - T \ln\left(\frac{T}{T_m}\right) \right)$$

$$\Delta G_m(T) = \Delta H_m(T_m) \left(1 - \frac{T}{T_m} \right) + \Delta Cp_m \left(T - T_m - T \ln\left(\frac{T}{T_m}\right) \right)$$

$$K_{idu}(T) = e^{\left(\frac{-\Delta G_{idu}(T)}{RT}\right)}$$

$$\Delta G_{idu}(T) = \Delta H_{idu}(T_{idu}) - T \left(\frac{\Delta H_{idu}(T_{idu})}{T_{idu}} - R \ln K_{idu}(T_{idu}) \right) + \Delta Cp_{idu} \left(T - T_{idu} - T \ln\left(\frac{T}{T_{idu}}\right) \right)$$

$$\ln K_{idu}(T_{idu}) \approx \frac{\ln[I]^{n-1} \cdot n^{n/2} \cdot \prod_{i=1}^k (m_i)^{m_i}}{(\sqrt{n} + 1)^{n-1}}$$

Où : $[I]$ = le nombre maximum possible de complexe Intermédiaire (*Associés*)

$$AB \Leftrightarrow 1A + 1B \quad n = 2 \quad m_1 = 1 \quad m_2 = 1$$

$$\ln K_{idu}(T_{idu})_{HETERO} \approx \frac{\ln[I]^1 \cdot 2^1 \cdot (1^1 \cdot 1^1)}{(\sqrt{2} + 1)^1} = \ln[I] \cdot \frac{2}{2,4142}$$

$$\ln K_{idu}(T_{idu})_{HETERO} \approx \ln[I] \cdot 0,8284$$

$$\Delta G_{idu}(T) = \Delta H_{idu}(T_{idu}) - T \left(\frac{\Delta H_{idu}(T_{idu})}{T_{idu}} - R \ln([I] \cdot 0,8284) \right) + \Delta Cp_{idu} \left(T - T_{idu} - T \ln\left(\frac{T}{T_{idu}}\right) \right)$$

$$\Delta G_{idu}(T) = \Delta H_{idu}(T_{idu}) - T \left(\frac{\Delta H_{idu}(T_{idu})}{T_{idu}} - R \ln(0,4142[T]) \right) + \Delta Cp_{idu} \left(T - T_{idu} - T \ln\left(\frac{T}{T_{idu}}\right) \right)$$

$$K_{midu}(T) = K_{ni}(T) \cdot K_{idu}(T)$$

$$\text{si: } \frac{[T_1] - [U]}{2} = [N_2] + [I_2] \rightarrow [N_2] = \frac{[T_1] - [U]}{2} - [I_2]$$

$$K_{ni} = \frac{[I_2]}{[N_2]} \rightarrow K_{ni} = \frac{[I_2]}{\left(\frac{[T_1] - [U]}{2} - [I_2]\right)}$$

$$K_{ni} \left(\frac{[T_1] - [U]}{2} \right) - K_{ni} \cdot [I_2] = [I_2]$$

$$K_{ni} \left(\frac{[T_1] - [U]}{2} \right) = [I_2] + K_{ni} \cdot [I_2]$$

$$[I_2] = \frac{K_{ni} \left(\frac{[T_1] - [U]}{2} \right)}{(1 + K_{ni})}$$

$$\text{si: } P_I = \frac{2[I_2]}{[T_1]} \rightarrow P_I = \frac{K_{ni} \left(\frac{[T_1] - [U]}{[T_1]} \right)}{(1 + K_{ni})} = \frac{K_{ni} \cdot [T_1] (1 - P_U)}{(1 + K_{ni}) [T_1]} = \frac{K_{ni} (1 - P_U)}{(1 + K_{ni})}$$

$$\text{si: } [N_2] = \frac{([T_1] - 2[I_2] - [U])}{2}$$

$$\text{si: } P_N = \frac{2[N_2]}{[T_1]} \rightarrow P_N = \frac{([T_1] - 2[I_2] - [U])}{[T_1]}$$



**AUTORISATION D'INTÉGRATION
D'UN ARTICLE ÉCRIT EN COLLABORATION
À UN MÉMOIRE OU UNE THÈSE**

Je (ou nous), soussignée(s), soussigné(s), co-auteur(e)s, co-auteur(s) de l'article intitulé : Reversible
Dimerization Between Human C-myc and Max b-HLH-LZ
Domain

reconnais (reconnaissons) que ledit article sera inclus comme partie constituante du mémoire de la thèse

de l'étudiant(e) (nom) : François-Olivier McDuff

Inscrit(e) au programme de Pharmacologie

de la Faculté de Médecine de l'Université de Sherbrooke

En foi de quoi, j'ai (nous avons) signé cet engagement en un nombre suffisant d'exemplaires*

Signature Date : X 09/03/09

Nom Jean-François Naud

Coordonnées Biozentrum, Physiol. Chemie II
Universität Würzburg, 97074 Würzburg
Allemagne

Signature Date : X 01/12/08

Nom Réjean Labelle

Coordonnées Dép. Médecine Nucléaire et Radiobiologie
U. de S.

Signature Date :

Nom Simon Sauvé

Coordonnées

Signature _____ Date : _____

Nom _____

Coordonnées _____

* Un exemplaire pour l'étudiante, l'étudiant, un exemplaire pour chaque personne signataire et un exemplaire pour le Service des bibliothèques (à remettre avec le mémoire ou la thèse au moment du dépôt final).

**AUTORISATION D'INTÉGRATION
D'UN ARTICLE ÉCRIT EN COLLABORATION
À UN MÉMOIRE OU UNE THÈSE**

Je (ou nous), soussignée(s), soussigné(s), co-auteur(s), co-auteur(s) de l'article intitulé :

Reversible Dimerization Between Human c-Myc and Max
b-ABL-LZ Domains

reconnais (reconnaissons) que ledit article sera inclus comme partie constituante du mémoire de la thèse

de l'étudiant(e) (nom) : François-Olivier MeDuff

inscrit(e) au programme de Pharmacologie

de la Faculté de Médecine de l'Université de Sherbrooke

En foi de quoi, j'ai (nous avons) signé cet engagement en un nombre suffi

Signature _____ Date : 13/09/2007

Nom MARTIN MONTAGNE

Coordonnées Pharmacologie, Médecine, CHUS

Signature _____ Date : 13/09/2007

Nom KYRÈRE LANGRIS

Coordonnées Pharmacologie FMSSUS

Signature _____ Date : _____

Nom _____

Coordonnées _____

Signature _____ Date : _____

Nom _____

Coordonnées _____

* Un exemplaire pour l'étudiante, l'étudiant, un exemplaire pour chaque personne signataire et un exemplaire pour le Service des bibliothèques (à remettre avec le mémoire ou la thèse au moment du dépôt final).