

Université de Montréal

**Caractérisation structurale et biophysique de l'impact de
l'acétylation de SUMO1 sur son interaction dépendante de
la phosphorylation avec PML**

par Christina Gagnon

Département de Biochimie et Médecine Moléculaire, Université de Montréal
Faculté de Médecine

Mémoire présenté
en vue de l'obtention du grade de Maîtrise
en Biochimie
option Générale

Juillet 2017

© Christina Gagnon, 2017

Résumé

La formation et la dynamique des corps nucléaires contenant PML (PML-NBs) est extensivement régulée par un réseau complexe de modifications post-traductionnelles, incluant la SUMOylation des protéines résidentes des NBs, la phosphorylation des motifs interagissant avec SUMO (SIMs) et l'acétylation des protéines de la famille SUMO. Des études récentes suggèrent que l'acétylation des protéines SUMO contrôle la formation des PML-NBs, mais le rôle précis de cette modification demeure énigmatique. Usant des techniques de calorimétrie par titration isotherme (ITC) et de cristallographie par diffraction des rayons X nous avons investigué l'impact de l'acétylation de SUMO1 sur son interaction avec les phospho-SIMs de PML et Daxx. La neutralisation de la charge positive des résidus K37, K39, K45, ou K46 par acétylation ou par remplacement avec un résidu glutamine diminue de manière significative l'affinité de liaison envers les variants phospho-mimétiques de PML-SIM telle que mesurée par ITC, mais la perte d'interaction est plus importante avec les variants K39 et K46 qu'avec les variants K37 et K45. Les structures cristallographiques des phospho-SIMs de PML et Daxx en complexe avec le variant SUMO1 K46Ac démontrent des changements dramatiques au niveau des interactions à l'interface de liaison, mais ces changements semblent distincts pour chaque SIM. Les autres variants acétylés en complexe avec le phospho-SIM de PML démontrent une réorganisation plus globale d'effet de charge à l'interface de liaison. En résumé, nos résultats apportent une description au niveau atomique sur la manière dont l'acétylation de SUMO1 peut moduler son interaction avec les différents phospho-SIMs, ce qui indique qu'il y a une plasticité considérable impliquée dans les interactions SUMO/SIM. Cette plasticité pourrait être altérée par divers types de modifications post-traductionnelles afin de permettre la régulation de fonctions cellulaires distinctes.

Mots-clés : SUMO, PML, SUMOylation, cristallographie par diffraction des rayons X, calorimétrie par titration isotherme (ITC), corps nucléaires contenant PML (PML-NBs), motif interagissant avec SUMO (SIM), suppresseur de tumeur, phospho-sérines

Abstract

PML-Nuclear Bodies (PML-NBs) formation and dynamics is regulated by an extensive network of post-translational modifications (PTMs), which includes SUMOylation of PML-NB-resident proteins, phosphorylation of SIMs (SUMO-Interacting Motifs) and acetylation of SUMO proteins. Recent studies suggest that acetylation of SUMO proteins regulates PML-NB function, but the precise role of acetylation is still not clear. Using Isothermal Titration Calorimetry (ITC) and X-ray crystallography studies, we have investigated the impact of SUMO1 acetylation on its interaction with the phospho-SIMs of PML and Daxx. Interestingly, neutralizing the positive charges of K37, K39, K45 or K46 by either acetylation or glutamine substitution significantly decreases the binding affinity towards phospho-mimetic PML-SIMs as measured by ITC, but the decreases are more dramatic with the K39 and K46 variants than with the K37 and K45 variants. In the crystal structures of the phospho-SIMs of PML and Daxx in complex with K46Ac, there are dramatic disruptions of interactions at the binding interfaces, but the changes that occur at the interface appear to be distinct for the SIM involved. The other acetylated variants in complex with the phospho-SIM of PML displayed a more global effect of charge re-organization at the binding interface. Together, our results provide an atomic level description of how SUMO1 acetylation can modulate its interaction with different phospho-SIMs, indicating that there is considerable plasticity in SUMO/SIM interactions and this plasticity could be altered by different types of PTMs in order to regulate distinct cellular functions.

Keywords: SUMO, PML, SUMOylation, X-ray crystallography, Isothermal Titration Calorimetry (ITC), PML-Nuclear Bodies (PML-NBs), SUMO-Interacting Motif (SIM), tumour suppressor, phosphoserines

Table des matières

Résumé	i
Abstract.....	ii
Table des matières	iii
Liste des tableaux	v
Liste des figures.....	vi
Liste des sigles.....	viii
Liste des abréviations	xi
Remerciements	xiii
1. Introduction	1
1.1. Les modifications post-traductionnelles des protéines.....	2
1.1.1. SUMOylation	2
1.1.2. Phosphorylation.....	13
1.1.3. Acétylation	16
1.2. Cancer et suppresseurs de tumeur	20
1.2.1. La protéine <i>ProMyelocytic Leukemia</i> (PML).....	20
1.3. But du projet.....	26
1.3.1. Problématique et hypothèses	26
1.3.2. Objectifs de recherche	27
1.4. Approche expérimentale.....	27
1.4.1. Production de protéines modifiées	27
1.4.2. Calorimétrie par titration isotherme	31
1.4.3 Cristallographie par diffraction des rayons X	34
2. Structural and biophysical characterization of the impact that SUMO1 acetylation has on its phosphorylation-dependent interaction with PML.....	37
Detailed authors contribution	39
Summary.....	40

Introduction	41
Results	44
Acetyl-mimetic variants of SUMO1 have lower affinities for phospho-mimetic PML-SIMs	44
Acetylated Δ N-SUMO1 variants behave similarly to K-to-Q variants in ITC studies	45
Acetylation of SUMO1 does not alter binding to the hydrophobic core PML-SIM	46
Crystallization trials and structure of the Δ N-SUMO1-K46Ac:PML-SIM-4PO ₄ complex	47
Structure of the Δ N-SUMO1-K39Q:PML-SIM-4PO ₄ complex	48
Structures of the Δ N-SUMO1-K37Ac:PML-SIM-4PO ₄ and the Δ N-SUMO1-K45Ac:PML-SIM-4PO ₄ complex	48
Structure of the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO ₄ complex	50
Discussion.....	52
Experimental procedures	55
Recombinant peptides and proteins constructs.....	55
Peptides and proteins expression and purification	55
<i>In vitro</i> phosphorylation of SIM peptides	57
ITC experiments	57
Crystallization and data collection	57
Structure determination and refinement	58
Authors contribution.....	59
Acknowledgements	59
Funding.....	59
Conflict of interest.....	59
Tables	60
Figures	66
Discussion générale et conclusion.....	80
Bibliographie	85
Supplementary figures.....	xiv

Liste des tableaux

Table I. ITC results for ΔN-SUMO1 K-to-Q variants binding to either PML-SIM-4SD or PML-SIM-4SE.	60
Table II. ITC results for ΔN-SUMO1 K-to-Ac variants binding to either PML-SIM-4SD or PML-SIM-4SE.	60
Table III. Crystal data collection and refinement statistics for acetylated SUMO variants in complex with PML-SIM.	61
Table IV. Crystal data collection and refinement statistics for acetylated SUMO variants in complex with PML-SIM-4PO₄.	62
Table V. Summary of distances between phosphoserines (pS) of PML-SIM-4PO₄ and key residues of ΔN-SUMO1, ΔN-SUMO1-K46Ac and ΔN-SUMO1-K39Q.	63
Table VI. Summary of distances between phosphoserines (pS) of PML-SIM-4PO₄ and key residues of ΔN-SUMO1, ΔN-SUMO1-K37Ac and ΔN-SUMO1-K45Ac.	63
Table VII. Data collection and refinement statistics for ΔN-SUMO1-K46Ac in complex with mDaxx-SIM-2PO₄.	64
Table VIII. Summary of distances between phosphoserines (pS) of Daxx-SIM and key residues of ΔN-SUMO1 and ΔN-SUMO1-K46Ac.	65

Liste des figures

Figure 1. Comparaison de la structure tridimensionnelle des protéines de la famille SUMO.....	3
Figure 2. Alignement de séquences des quatre paralogues SUMO chez l'humain (SUMO1-4_HUMAN) et des homologues de SUMO1 chez la levure <i>S. cerevisiae</i> (SMT3_YEAST) et chez le ver <i>C. elegans</i> (SUMO_CAEEL).....	4
Figure 3. Comparaison du repliement de ubiquitine et SUMO1.....	5
Figure 4. Mécanisme général de la SUMOylation.....	7
Figure 5. Comparaison des orientations possibles des interactions SUMO/SIM. (A) Exemple de l'interaction parallèle entre SUMO1 (gris) et RanBP2 (cyan) [PDB ID 1Z5S]. (B) Exemple de l'interaction antiparallèle entre SUMO1 (gris) et un SIM dérivé de PIASX (magenta) [PDB ID 2ASQ].	12
Figure 6. Mécanisme général de la phosphorylation <i>in vivo</i>.....	14
Figure 7. Mécanisme général de l'acétylation <i>in vivo</i>.	17
Figure 8. Lysines acétylées et implication dans les interactions SUMO/SIM.....	18
Figure 9. Représentation schématique des exons de la protéine encodée par le gène <i>PML</i> et des différentes isoformes générées par épissage alternatif.....	21
Figure 10. Schématisation des multiples fonctions des PML-NBs.....	23
Figure 11. Représentation schématique de la forme générale de la protéine PML montrant les principaux domaines et modifications post-traductionnelles.	24
Figure 12. Représentation schématique d'un modèle de biogénèse des PML-NBs.....	25
Figure 13. Production d'une protéine acétylée en bactérie utilisant l'expansion du code génétique.....	30
Figure 14. Représentation schématique de l'intérieur d'un calorimètre (gauche) et d'une expérience de titration typique (droite, haut) accompagnée de son évaluation (droite, bas).	33
Figure 15. Schématisation des principales étapes de la résolution de structures macromoléculaires utilisant la cristallographie par diffraction de rayons X.....	36

Figure 16. Representation of important post-translational modification sites on SUMO1 and PML-SIM.	66
Figure 17. SUMO1 K-to-Q variants dramatically decrease binding to PML-SIM peptides.	67
Figure 18. Acetylated SUMO1 variants bind to PML-SIM peptides in a similar manner as K-to-Q variants.	68
Figure 19. Acetylation of SUMO1 does not affect interactions with the hydrophobic core of the PML-SIM.	69
Figure 20. Crystal structure of the ΔN-SUMO1-K46Ac:PML-SIM-4PO₄ complex.	70
Figure 21. Comparison of ΔN-SUMO1:PML-SIM-4PO₄ (PDB entry 4WJN) and ΔN-SUMO1-K46Ac:PML-SIM-4PO₄ complexes.	71
Figure 22. Crystal structure of the ΔN-SUMO1-K39Q:PML-SIM-4PO₄ complex.	72
Figure 23. Comparison of ΔN-SUMO1:PML-SIM-4PO₄ (PDB entry 4WJN) and ΔN-SUMO1-K39Q:PML-SIM-4PO₄ complexes.	73
Figure 24. Crystal structure of ΔN-SUMO1-K37Ac:PML-SIM-4PO₄ complexes.	74
Figure 25. Comparison of ΔN-SUMO1:PML-SIM-4PO₄ (PDB entry 4WJN) and ΔN-SUMO1-K37Ac:PML-SIM-4PO₄ complexes.	75
Figure 26. Crystal structures of ΔN-SUMO1-K45Ac:PML-SIM-4PO₄ complexes.	76
Figure 27. Comparison of ΔN-SUMO1:PML-SIM-4PO₄ (PDB entry 4WJN) and ΔN-SUMO1-K45Ac:PML-SIM-4PO₄ complexes.	77
Figure 28. Crystal structure of ΔN-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complexes.	78
Figure 29. Comparison of ΔN-SUMO1:mDaxx-SIM-2PO₄ (PDB entry 4WJP) and ΔN-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complexes.	79

Liste des sigles

AcK : *N*ε-acetyl-lysine

APL : Acute Promyelocytic Leukemia

As₂O₃ : arsenic trioxide

BRD : BRomoDomain

CCD : Charge-Coupled Device

CCP4 : Collaborative Computational Project Number 4

CK2 : Casein Kinase 2

CLS : Canadian Light Source

DTT : DiThioTreitol

EDTA : Ethylene Diamine Tetra-Acetic

ERK : Extracellular signal-Regulated Kinase

ERK1/2 : Extracellular-Regulated Kinases 1/2

GSH : Reduced (SH) Glutathion (G)

GST : Glutathion-S-Transferase

HAT : Histone AceTyltransferase

HCl : chlorhydric acid

HDAC : Histone DeACetyltransferase

HIPK2 : Homeodomain-Interacting Protein Kinase 2

HPLC: High-Pressure Liquid Chromatography

IPTG : IsoPropyl-β-D-ThioGalactopyranoside

ITC : Isothermal Titration Calorimetry

KAT : Lysine (K) AceTyltransferase

KDAC : Lysine (K) DeACetyltransferase

LB : Luria Broth

MacCHESS : Macromolecular diffraction facility at Cornell High Energy Synchrotron Source

Na₂HPO₄ : sodium hydrogenophosphate

NaCl : sodium chloride

NaH₂PO₄ : sodium dihydrogenophosphate

NAM : NicotinAMide

NB : Nuclear Bodies
NDSM : Negatively charged-Dependent SUMOylation Motif
NLS : Nuclear Localization Signal
O.D.₆₀₀ : Optical Density at 600nm
PC : PolyComb protein
PDB : Protein Data Bank
PDSM : Phosphorylation-Dependent SUMOylation Motif
PEG : PolyEthylene Glycol
PIAS : Protein Inhibitor of Activated STAT
PINIT : Pro-Ile-Asn-Ile-Thr
PLK1 : Polo-Like Kinase 1
PML : ProMyelocytic Leukemia protein
Ran : Ras-related nuclear protein
RanBP2 : Ran Binding Protein 2
RAR : Retinoic Acid Receptor
RBCC : RING domain, B-box, Coiled-Coil
RLD : RING-finger-Like zing-binding Domain
rpm : revolution per minute
SAE : SUMO-Activating Enzyme
SAP : N-terminal Scaffold attachments factor-A/B, Acinus and PIAS domain
SENP : SENtrin-specific Protease
SIM : SUMO-Interacting Motif
Siz1 : SAP and miz-finger domain
SLD : SUMO-Like Domain
SUMO : Small Ubiquitin-like MOdifier
TEV : Tobacco Etch Virus
TFA : TriFluoroAcetic acid
TOPORS : TOP1 binding arginine/serine-rich protein
TRIM : TRIPartite Motif
UBC9 : UBiquitin-like Conjugating enzyme 9

w/v : weight/volume

Liste des abréviations

°C : degré Celcius

3D : tridimensionnel(le)

C. elegans : *Caenorhabditis elegans*

E. coli : *Escherichia coli*

etc. : et cætera

h : heure

kDa : kiloDalton

M : molaire

min. : minute

ml : millilitre

mM : millimolaire

S. cerevisiae : *Saccharomyces cerevisiae*

Se-Met : séléno-méthionine

Tris : Tris(hydroxyméthyl)aminométhane

ug : microgramme

uM : micromolaire

Pour M. et P.

Remerciements

Je tiens à remercier sincèrement tous les membres des divers laboratoires du Département de Biochimie qui ont participé de près ou de loin à la qualité de ma formation. Plus particulièrement, la rédaction de ce document aurait été impossible sans la généreuse implication de mon superviseur Jim Omichinski ainsi que les précieux et très appréciés commentaires éditoriaux de mes collègues Xavier Mascle et Soumiya Amellah. Merci également à tous mes ami(e)s et membres de famille qui ont fait preuve d'une compréhension et d'une coopérativité incomparable; ceux-ci m'ayant permis de conserver ma ligne directrice et mon moral. Finalement, un grand merci aux professeurs John M. Pascal et Damien d'Amours qui ont pris le temps d'effectuer la lecture de cet ouvrage et d'y apporter des remarques constructives afin de grandement en améliorer la qualité.

1. Introduction

1.1. Les modifications post-traductionnelles des protéines

Les modifications post-traductionnelles sont définies comme l'ajout covalent de groupements chimiques variés (phosphate, acétyle, sulfate, méthyle, etc.) à des acides aminés distincts constituant la séquence de protéines cibles. Ces modifications surviennent une fois la synthèse de la protéine cible complétée par le ribosome, d'où leur appellation. Les réactions chimiques impliquées dans ces modifications sont facilitées par des enzymes spécifiques. Ainsi, le contrôle de l'activité catalytique des enzymes et des réseaux d'interaction protéiques représente un mécanisme important de régulation de multiples voies de signalisation cellulaire (1, 2). Ces fonctions sont accomplies via la reconnaissance de ces modifications par des domaines d'interaction spécifiques (3). À titre d'exemple, la phosphorylation d'acides aminés contenant un groupement hydroxyle dans leur chaîne latérale (sérine, thréonine, tyrosine) est utilisée extensivement *in vivo* comme un moyen de moduler l'activité catalytique de nombreuses enzymes (kinases, phosphatases, etc.) (4). Alternativement, les modifications post-traductionnelles incluent également l'addition des petites protéines sur des résidus lysines cibles de substrats protéiques. Parmi ces processus, le mieux caractérisé est l'ubiquitinylation, au cours duquel la protéine ubiquitine est conjuguée sur la cible via une cascade enzymatique impliquant plusieurs enzymes. Des protéines structurellement similaires à l'ubiquitine telles que SUMO (*Small Ubiquitin-like MOdifier*) et NEDD8 (*Neural precursor cell Expressed Developmentally Down-regulated protein 8*), peuvent être employées dans des cascades enzymatiques semblables à celle de l'ubiquitinylation. Ces réactions enzymatiques sont respectivement qualifiées de SUMOylation et NEDDylation. Toutes ces protéines ont une répartition cellulaire particulière et possèdent des fonctions qui leur sont propres. En particulier, la SUMOylation est un processus survenant dans le noyau des cellules eucaryotes et l'intérêt porté à cette modification provient de son implication dans la régulation de processus nucléaires fondamentaux (5).

1.1.1. SUMOylation

1.1.1.1. Les protéines de la famille SUMO

On entend par SUMOylation la conjugaison covalente d'une ou de plusieurs protéine(s) de la famille SUMO sur des résidus lysines de substrats protéiques (6). Chez

l'humain (*Homo sapiens*), la famille des protéines SUMO comprend quatre paralogues (SUMO1-4) d'environ 100 acides aminés (poids moléculaire ~12kDa). Ceux-ci présentent une structure tridimensionnelle similaire (**Figure 1**) malgré une variabilité importante de leur séquence primaire (**Figure 2**). Les protéines SUMO2 et SUMO3 présentent un haut degré d'identité de séquences. On les regroupe d'ailleurs sous l'appellation SUMO2/3 puisque leur distinction par un anticorps est impossible (7). Par opposition, elles ne partagent que 40% d'identité de séquence avec SUMO1 (8). Les variations au sein des séquences primaires de ces paralogues semblent être associées à leur utilisation intracellulaire et à leurs fonctions respectives (9).

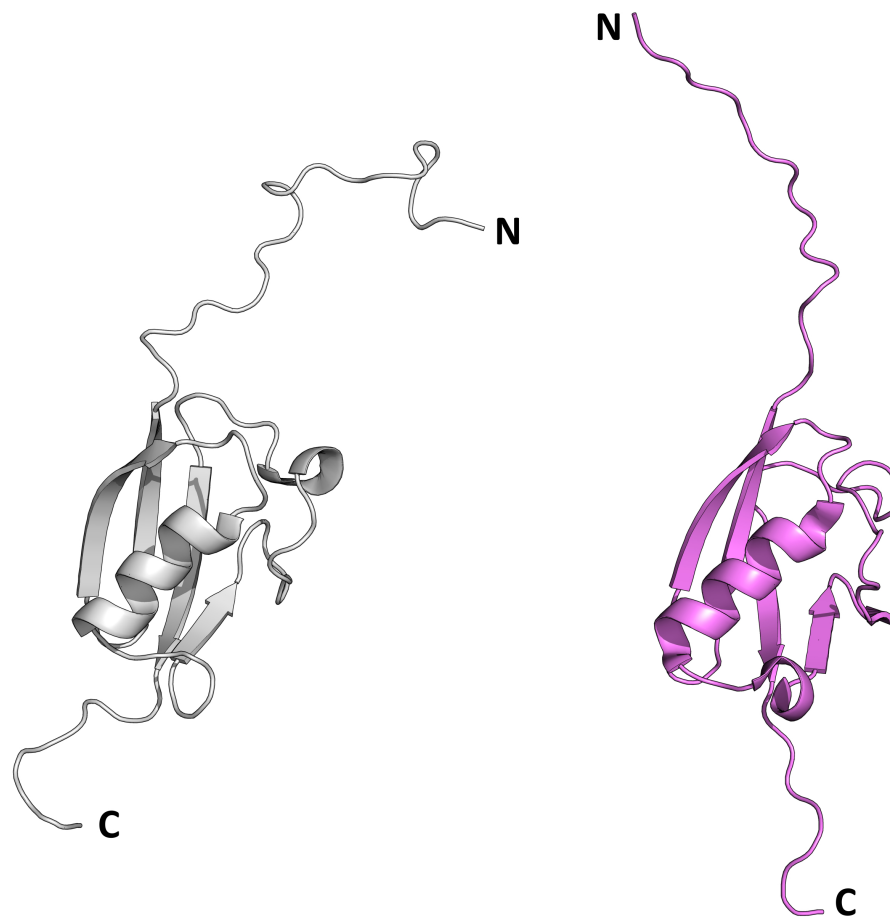


Figure 1. Comparaison de la structure tridimensionnelle des protéines de la famille SUMO. Malgré une identité de séquence de seulement 40% entre les paralogues SUMO1 (gris, PDB ID [1A5R](#)) et SUMO2/3 (magenta, PDB ID [2N9E](#)), les deux protéines possèdent un repliement similaire comprenant une hélice α et quatre brins β antiparallèles (représentation *cartoon* du logiciel PyMOL). Une région non structurée étendue est retrouvée en N-terminal (N) des deux protéines, alors que la conjugaison au substrat s'effectue en C-terminal (C).

Le paralogue SUMO4 partage une plus grande identité de séquence avec SUMO2/3 (76%) qu'avec SUMO1 (40%). SUMO1-2/3 sont exprimées de façon ubiquitaire chez l'humain, alors que SUMO4 est exprimée dans certains organes seulement (la rate, les reins, le pancréas et les ganglions) (10). Chez d'autres Eucaryotes tels que la levure *Saccharomyces cerevisiae* et le ver *Caenorhabditis elegans*, on retrouve une protéine homologue de SUMO1 : SMT3 chez la levure (11) et SMO-1 chez le ver (**Figure 2**). Des études sur divers arthropodes ont démontré l'existence de protéines analogues à SUMO1 et SUMO2/3 chez les insectes basaux tels que *Blattella germanica*, alors que le paralogue correspondant à SUMO1 humain aurait été perdu chez *Drosophila melanogaster* au cours de l'évolution, ne laissant que SUMO3 chez cette espèce de mouche (12).

```

SUMO1_HUMAN MS----DQEAKPSTEDLGDKKEGEYIKLKVIGQDSSEIHFKVMTTHLKKLKESYCQRQG 56
SUMO2_HUMAN MA----DE--KPK--EGVKTENNDHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCERQG 52
SUMO3_HUMAN MS----EE--KPK--EGVKTE--NDHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCERQG 51
SUMO4_HUMAN MA----NE--KPT--EEVKTENNNHINLKVAGQDGSVVQFKIKRQTPLSKLMKAYCEPRG 52
SMT3_YEAST  MSDSEVNQEAKPEVKPE--VKPETHINLKVS-DGSSEIFFKIKKTTPLRRLMEAFKRQG 57
SUMO_CAEEL  MA-----DDAAQAGDNAEYIKIKVVGQDSNEVHFRVKYGTSMALKKSYADRTG 49
* :                               .      *:::*  :... : *::*  * : :*  :...  *

SUMO1_HUMAN VPMNSLRFLFEGQRIADNHTPKELGMEEDVIEVYQEQTGGHSTV----- 101
SUMO2_HUMAN LSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGGVY----- 95
SUMO3_HUMAN LSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGGVPESLAGHSF 103
SUMO4_HUMAN LSMKQIRFRFDGQPISGTDKPAQLEMEDEDTIDVFQQOPTGGVY----- 95
SMT3_YEAST  KEMDSLRFlyDGIrIQADQTPEDLDMEDNDIIEAHREQIGGATY----- 101
SUMO_CAEEL  VAVNSLRFLFDGRRINDDDTPKTLMEDEDDVIEVYQEQLGGF----- 91
: .::* : * * ..* * *::* *:::  **

```

Figure 2. Aligement de séquences des quatre paralogues SUMO chez l'humain (SUMO1-4_HUMAN) et des homologues de SUMO1 chez la levure *S. cerevisiae* (SMT3_YEAST) et chez le ver *C. elegans* (SUMO_CAEEL). Les résidus conservés (*), la conservation de la polarité d'un résidu (:), et les résidus conservés dans au moins la moitié des séquences (.) sont représentés. Les lysines acétylées de SUMO1 et leur équivalent dans les autres protéines sont surlignés en gris. Les lysines (K) 37 et 39 de SUMO1 sont conservées chez la levure et dans les autres paralogues humains. La lysine 45 de SUMO1 n'est pas conservée dans les autres protéines. La lysine 46 est conservée à travers les paralogues humains et chez le ver, mais remplacée par une arginine (R) chez la levure.

Le corps globulaire des protéines SUMO est très similaire à celui retrouvé chez la protéine ubiquitine (**Figure 3**) malgré que celles-ci ne partagent que 18% d'identité de séquence et possèdent des fonctions distinctes (13). Les protéines SUMO se caractérisent par le repliement $\beta\beta\alpha\beta\beta\alpha\beta$ typique des protéines de la famille de l'ubiquitine. De plus, les formes matures de SUMO et de l'ubiquitine présentent un motif diglycine (GG) en C-terminal (C).

Tout comme l'ubiquitine, des chaînes de poly-SUMOs ont été identifiées *in vivo*. Cependant, cette capacité à la polymérisation semble exclusive à SUMO2/3, puisque SUMO1 ne contient pas dans sa séquence le site consensus interne qui permet l'assemblage de chaînes poly-SUMOs. Néanmoins, des chaînes mixtes peuvent être observées où SUMO1 se retrouve à leur extrémité (14).

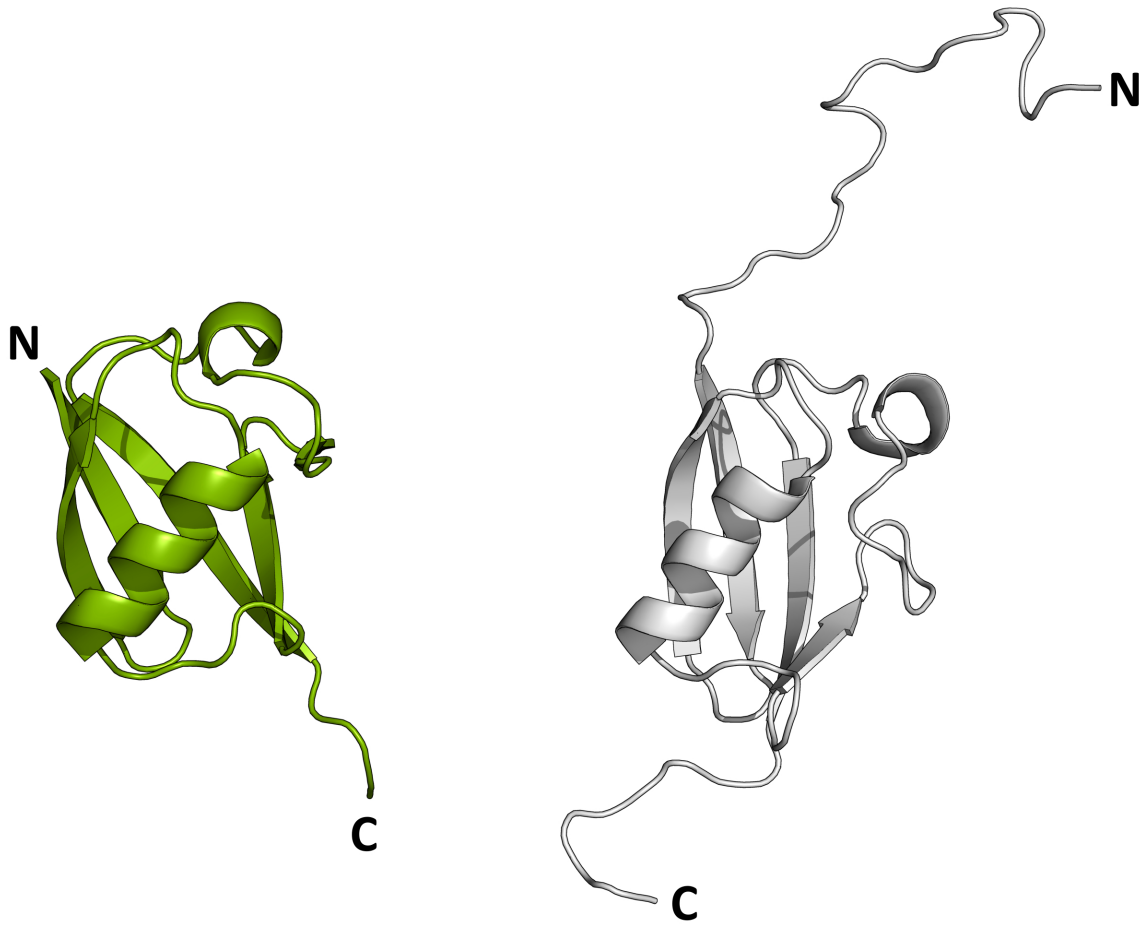


Figure 3. Comparaison du repliement de ubiquitine et SUMO1. Ubiquitine (vert, PDB ID [1UBQ](#)) et SUMO1 (gris, PDB ID [1A5R](#)) possèdent une structure tridimensionnelle semblable lorsque les molécules sont représentées en tant qu'éléments de structure secondaire : deux séries de feuillets β antiparallèles et une hélice α séparés par des boucles. La différence structurale majeure entre les deux protéines est la présence d'une boucle étendue en N-terminal (N) chez SUMO1 qui est absente chez l'ubiquitine.

1.1.1.2. La voie de SUMOylation

1.1.1.2.1. Cascade enzymatique

Susnommées, les protéines SUMO sont hautement semblables à l'ubiquitine d'un point de vue structural malgré une faible identité de séquence. Cette ressemblance s'étend d'autant plus au niveau biochimique, puisque les réactions de SUMOylation et d'ubiquitinylation présentent des cascades enzymatiques analogues impliquant des enzymes E1, E2 et E3 (**Figure 4**). Durant la SUMOylation, les protéines SUMO sont liées de manière covalente aux groupements latéraux ϵ -amine de résidus lysines cibles sur leurs substrats. Les protéines SUMO sont initialement traduites sous forme de précurseurs qui sont ensuite maturés par des protéases spécifiques (*sentrin-specific proteases* ; SENPs) qui clivent l'extrémité C-terminale des protéines SUMO afin de révéler le motif diglycine. Ce dernier est ensuite adénylé à son extrémité carboxy-terminale par l'enzyme dimérique E1 afin de générer la forme activée de SUMO (5). Une fois la maturation complétée, un résidu cystéine conservé de l'enzyme E1 (complexe *SUMO-Activating Enzymes 1/2* ; SAE1-SAE2) procède à une attaque nucléophile sur le groupement adénylé de SUMO mature, ceci menant à la formation d'un lien thioester. Ensuite, SUMO est transféré à l'enzyme de conjugaison E2 par un mécanisme similaire formant un deuxième intermédiaire thioester. À ce niveau, deux cas de figure sont possibles : l'enzyme E2 peut directement transférer SUMO sur la lysine du substrat, ou bien une enzyme E3 ligase peut intervenir. Les E3 ligases peuvent faciliter le processus de conjugaison en stabilisant le complexe entre l'intermédiaire thioester E2~SUMO et le substrat, ou en augmentant la probabilité de conjugaison en stimulant la décharge de SUMO par E2 sur la protéine cible. Chez l'humain, la cascade de SUMOylation dépend d'une unique enzyme E2 nommée UBC9 (*UBiquitin-Conjugating enzyme 9*) et de quelques E3 ligases, telles que la protéine RanBP2 (*Ran Binding Protein 2*) ou les membres de la famille PIAS (*Protein Inhibitor of Activated STAT*) (15-17). A contrario, la cascade d'ubiquitinylation possède plusieurs E2 et des centaines d'enzymes E3 caractérisées (18). Tout comme l'ubiquitinylation, la SUMOylation est un processus réversible où le retrait de la modification est catalysé par des isopeptidases spécifiques à chaque protéine. Le retrait de l'ubiquitine est effectué par les USPs/DUBs (*Ubiquitin Specific Proteases/DeUBiquitinating enzymes*), tandis que les

protéines SUMO sont déconjuguées de leurs substrats par les SENPs, les mêmes enzymes responsables de la maturation des précurseurs SUMO (19, 20).

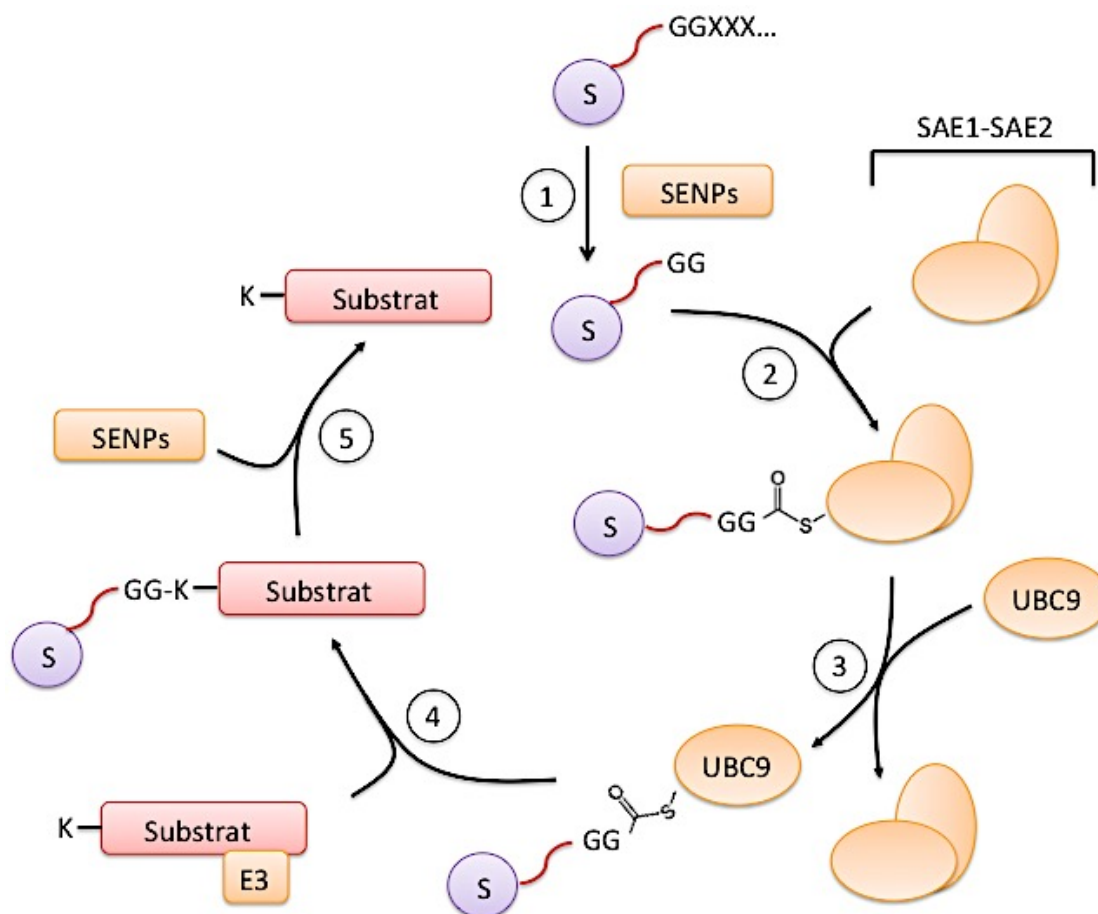


Figure 4. Mécanisme général de la SUMOylation. Le précurseur de SUMO est maturé par les SENPs (1) afin de révéler le motif diglycine (GG), puis activé et lié par l'enzyme E1, SAE1-SAE2 (2). SUMO est ensuite transféré à l'enzyme E2, UBC9 (3). Le transfert final à la lysine (K) cible du substrat peut ensuite se faire avec (4) ou sans l'intervention d'une E3 ligase par divers mécanismes et interactions protéiques. Les SENPs sont aussi responsables de la déconjugaison de SUMO du substrat (5). *Figure adaptée de : (21).*

1.1.1.2.2. Motifs consensus de conjugaison

La majorité des protéines ciblées pour conjugaison aux protéines SUMO partagent un motif consensus de conjugaison qui est reconnu par l'enzyme E2 de la machinerie de SUMOylation. Ceci contraste avec l'ubiquitinylation qui semble être un processus moins spécifique puisqu'aucun motif consensus de conjugaison clair n'a été caractérisé pour cette modification ; il semble plutôt que toute lysine accessible puisse potentiellement être

conjuguée à l'ubiquitine. La SUMOylation affecte principalement les résidus lysine (**K**) de substrats au sein de la séquence canonique ψ -**K**-X-D/E, où ψ représente un acide aminé hydrophobe, X peut être n'importe quel acide aminé et D/E est un acide aminé chargé négativement (22). Les résidus de ce motif interagissent directement avec UBC9 E2 ligase, ce qui stabilise l'interaction entre le substrat et l'enzyme (23). Des évidences structurales démontrent que, lors de ce contact intermoléculaire, la lysine ciblée se retrouve dans une poche hydrophobe du site actif de UBC9. Les résidus adjacents, quant à eux, forment des interactions polaires (liaisons hydrogènes et électrostatiques) avec des résidus en surface de UBC9. Ces interactions semblent faciliter la reconnaissance du substrat en participant au bon positionnement de la lysine cible tout en diminuant son pKa afin d'augmenter le taux de catalyse (24, 25). Pour que cette interaction puisse avoir lieu, le motif consensus de SUMOylation doit être dans une conformation étendue (26). En raison de ce critère stérique, il en découle que les motifs consensus se retrouvent de façon prédominante dans les régions flexibles des protéines cibles, telles que les longues boucles et les régions intrinsèquement désordonnées (5).

Au motif canonique s'ajoute deux autres types de séquences consensus sujettes à la SUMOylation : les motifs dépendants de la phosphorylation (*Phosphorylation-Dependant SUMOylation Motifs* ; PDSMs) et les motifs dépendants de résidus négativement chargés (*Negatively charged-Dependant SUMOylation Motifs* ; NDSMs). Ces séquences non-canoniques sont des formes plus étendues de la séquence canonique qui dépendent de la présence d'un résidu phosphorylé (dans le cas des PDSMs) ou d'un résidu acide (dans le cas des NDSMs) en aval de la lysine acceptrice. La forme générale d'une séquence de type PDSM est ψ -**K**-X-D/E-X-X-S-P, alors que celle d'un NDSM est ψ -**K**-X-D/E-X-X-S-D/E. La présence d'une proline adjacente au site de phosphorylation dans un PDSM implique que ce type de motif est retrouvé dans des substrats de kinases dirigées par la proline (27). La charge négative du groupement phosphate ou du résidu acide composant ces motifs augmente le taux de conjugaison de SUMO. Ceci s'explique par le fait que la charge négative additionnelle peut interagir avec des résidus positifs à la surface de UBC9 (K65 en particulier), ce qui étend la région d'interaction effective entre le substrat et l'enzyme E2, favorisant ainsi une interaction plus stable (28).

1.1.1.2.3. Fonctions biologiques

La principale différence entre SUMO et l'ubiquitine est le sort de la protéine cible suite à sa modification. L'ubiquitinylation engendre typiquement la dégradation des protéines cibles par le protéasome, bien que des fonctions alternatives lui sont désormais attribuées à travers les différents compartiments cellulaires (29). En contraste, on retrouve principalement les protéines SUMO dans le noyau des cellules où elles y effectuent des fonctions diverses : transport cytosol-noyau (30, 31) ; régulation de la stabilité génomique et de la transcription (32-34), du cycle cellulaire (35), et de l'apoptose (36) ; altération de la stabilité et des fonctions protéiques (37).

Plus spécifiquement, la modification par SUMO2/3 semble être associée à la réponse au stress (38), alors que la modification de substrats par SUMO4 semble être liée plus particulièrement au stress métabolique (39). De plus, la répartition cellulaire des différents paralogues est étroitement liée au cycle cellulaire : lors de la mitose, SUMO1 se situe au niveau du fuseau mitotique, alors que SUMO2/3 est associée avec les chromosomes et les centromères où elle se retrouve conjuguée à la topoisomérase II (40). SUMO4 étant le paralogue le plus récemment identifié, très peu de données concernant ses fonctions spécifiques sont actuellement disponibles. Toutefois, il a été rapporté que SUMO4 modifie la protéine I κ B α et serait associée au diabète de type I (41).

1.1.1.3. Modes de régulation

1.1.1.3.1. Activité, localisation et spécificité des enzymes E3 ligases

Il existe trois familles de SUMO E3 ligases : Siz/PIAS (42), RanBP2 (43), et *PolyComb 2* (PC2) (44). Les protéines de la famille PIAS (et leurs homologues Siz dans la levure) possèdent une structure comprenant des motifs et domaines particuliers qui permettent d'expliquer leur mécanisme d'action : le domaine SAP (*Scaffolds Attachment factor-A/B, acinus and PIAS*) et le motif PINIT (*Proline Isoleucine Asparagine Isoleucine Threonine*) permettent à ces protéines d'agir au niveau nucléaire, le RLD (*RING-finger-Like zinc-binding Domain*) est impliqué dans le processus de SUMOylation et permet l'interaction avec d'autres protéines, et le SIM (*SUMO-Interacting Motif*) est responsable de la reconnaissance et de l'interaction avec les protéines de la famille SUMO (45). Dans la levure, la localisation

cellulaire de la protéine Siz1 est dépendante du cycle cellulaire, ce qui module son accessibilité à ses substrats durant les divers évènements cellulaires (46).

Aussi, des modifications post-traductionnelles peuvent altérer l'activité des enzymes E3 ligase. Un exemple de régulation positive est l'interaction entre HIPK2 et PC2. HIPK2 est une protéine ciblée pour SUMOylation qui phosphoryle PC2 afin d'augmenter son activité envers ce même substrat (47). Il a aussi été démontré que certaines E3 ligases sont sujettes à de l'automodification par SUMO. La SUMOylation des enzymes E3 peut engendrer une augmentation de leur activité enzymatique et/ou une délocalisation cellulaire qui permettrait d'agir sur leurs substrats (48). Cette large collection de modifications post-traductionnelles dévoile l'idée d'une convergence entre les voies de SUMOylation et d'ubiquitinylation. Il semble exister une boucle de rétrocontrôle entre ces voies biochimiques dans laquelle l'ubiquitinylation des enzymes SUMO E3 ligase et la SUMOylation des enzymes de l'ubiquitinylation permet de moduler leur activité enzymatique respective (49).

En ce qui a trait à la spécificité, il a été déterminé qu'un même substrat peut être conjugué à SUMO par plusieurs différentes E3. Cette redondance semble indiquer que ce type d'interaction dépend d'une association entre E2 et E3, plutôt que de la reconnaissance du substrat (50). Cependant, certaines E3 ligases démontrent un pairage spécifique avec leur(s) substrat(s) et elles sont par conséquent nécessaires pour leur conjugaison à SUMO (51, 52). Une plus grande spécificité ainsi qu'une stimulation de la conjugaison sont tout spécialement nécessaires lorsque la lysine acceptrice ne fait pas partie d'une séquence consensus canonique (53). Il faut également considérer la spécificité du point de vue des différents paralogues de SUMO. En effet, certaines E3 ligases conjuguent un paralogue particulier à des substrats donnés. Cette spécificité peut être rationalisée par la stabilité relative d'un substrat face aux SENPs lorsque conjugué par SUMO1 ou SUMO2/3. Par exemple, la protéine RanGAP1 (le premier substrat modifié par SUMO1 ayant été identifié (30, 54)) peut être conjuguée à SUMO1 ou SUMO2/3 dans des essais *in vitro*, mais on ne retrouve que la forme conjuguée à SUMO1 *in vivo*. Cette dernière est protégée dans la cellule par la formation d'un complexe stable comprenant les enzymes UBC9 E2 ligase et RanBP2 E3 ligase (55). Il n'est pas clair si ce mode de régulation est généralisé dans la cellule, mais les exemples décrits suggèrent que le

contrôle de l'activité des SUMO E3 ligases peut être utilisé pour dicter la stabilité des substrats.

1.1.1.3.2. Motifs interagissant avec SUMO

Les protéines de la famille SUMO, sous forme libre ou conjuguées à leurs substrats, ont la capacité d'interagir de manière non-covalente avec d'autres protéines contenant des motifs particuliers étant typiquement non-structurés. Ces motifs interagissant avec SUMO sont appelés SIMs (56). Les séquences des SIMs consistent en de courts fragments de résidus hydrophobes, (V/I)X(V/I)(V/I) ou (V/I)(V/I)X(V/I) (57), qui sont parfois flanqués de régions riches en acides aminés chargés négativement (D/E) (58-60) ou phosphorylables (S/T) (61). Suite à son interaction avec SUMO, le SIM adopte une structure secondaire de type brin β . Cette interaction est stabilisée à travers des contacts proximaux entre les résidus hydrophobes du SIM et une poche hydrophobe de SUMO (60, 62). L'interaction SUMO/SIM peut être de type parallèle ou antiparallèle dépendamment du positionnement relatif du SIM par rapport au feuillet β impliqué de SUMO. L'orientation (parallèle versus antiparallèle) de l'interaction semble être dictée par l'emplacement des résidus acides/phosphorylables adjacents au SIM (**Figure 5**). À l'interface de l'interaction, ces résidus polaires peuvent interagir via des ponts salins ou des liaisons hydrogène avec des résidus conservés et positivement chargés des protéines SUMO (K37/39/45/46 chez SUMO1, K33/35/42 chez SUMO2/3) (63).

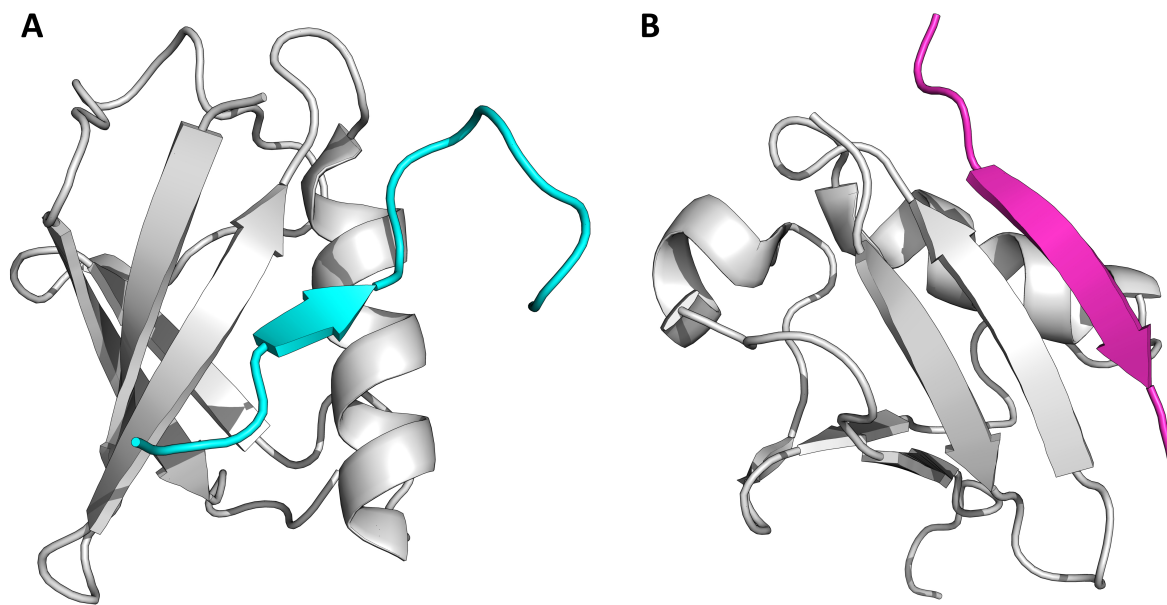


Figure 5. Comparaison des orientations possibles des interactions SUMO/SIM. (A) Exemple de l'interaction parallèle entre SUMO1 (gris) et RanBP2 (cyan) [PDB ID [1Z5S](#)]. (B) Exemple de l'interaction antiparallèle entre SUMO1 (gris) et un SIM dérivé de PIASX (magenta) [PDB ID [2ASQ](#)].

On retrouve des SIMs chez certaines protéines ciblées par la SUMOylation, mais aussi chez des protéines interagissant avec SUMO, des ubiquitine ligases guidées par SUMO, et dans plusieurs enzymes de la voie de SUMOylation (56, 64). Lorsqu'ils font partie des substrats ou des partenaires d'interaction avec SUMO, les SIMs peuvent avoir un impact sur l'activité ou la localisation à travers la modification de substrats par SUMO (65). Bien que la plupart des SIMs ne montrent aucune spécificité d'interaction pour les différentes protéines SUMO, certains présentent toutefois une spécificité pour un paralogue en particulier (66, 67). Les ubiquitine ligases guidées par SUMO utilisent leur(s) SIM(s) pour interagir avec des protéines SUMOylées ou contenant des *SUMO-Like Domains* (SLDs). Ces enzymes possèdent souvent plusieurs SIMs dans leur séquence, suggérant qu'elles pourraient interagir avec plusieurs molécules SUMO d'une chaîne (68). Alternativement, les SLDs permettent des interactions avec les composantes de la machinerie de SUMOylation telles que l'enzyme UBC9 E2 ligase et la sous-unité SAE2 du complexe SUMO E1 ligase (69). Les SIMs de ces enzymes leur permettent d'effectuer la reconnaissance de leurs substrats SUMOylés afin de procéder à leur ubiquitylation (70, 71). Dans les composantes de la machinerie de

SUMOylation, des SIMs sont présents dans les enzymes E1 et E3. Dans le cas des E1 ligases, deux SIMs ont été identifiés dans la partie C-terminale de la sous-unité SAE2 du complexe. La mécanistique de cette interaction SUMO/SIM est similaire à d'autres (72), mais sa fonction exacte demeure énigmatique puisque SAE2 conserve sa fonction en cellules humaines et de levures lorsque son C-terminus est trunqué (73). Chez les E3 ligases, plusieurs SIMs ont été identifiés dans la protéine RanBP2 (74). Il n'est toujours pas clair si tous ces motifs sont requis pour que RanBP2 puisse accomplir ses fonctions, mais il a tout de même été établi que la délétion du SIM entre les résidus 2629-2635 de RanBP2 humain affecte son activité SUMO E3 ligase. Ceci suggère du moins que son interaction non-covalente avec SUMO1 est importante pour la reconnaissance de l'intermédiaire E2~SUMO lors de la conjugaison (62). La caractérisation structurale du complexe SUMO1/UCB9/RanBP2 établie que le SIM de RanBP2 se lie à SUMO1 dans une orientation parallèle (**Figure 5A**) et qu'un regroupement de résidus acides en amont du cœur hydrophobe du SIM joue un rôle clé dans la détermination de l'orientation de l'interaction. Dans la famille des E3 ligases Siz/PIAS, un SIM est retrouvé en C-terminal (61). Cette famille de ligases possède également une série de résidus acides et/ou phosphorylables adjacents au SIM, mais contrairement à RanBP2, ceux-ci sont plutôt situés en aval de la partie hydrophobe. Ici aussi des évidences structurales suggèrent que les résidus chargés sont impliqués dans l'orientation de l'interaction SUMO1/PIAS1 qui est de type antiparallèle (**Figure 5B**). Récemment, il a été démontré que la phosphorylation du SIM de PIAS1 est cruciale à la formation du complexe tertiaire non-covalent entre SUMO1/PIAS1/UBC9 qui agit comme répresseur de la transcription (75).

1.1.2. Phosphorylation

1.1.2.1. Mécanisme d'action général

La phosphorylation est une modification post-traductionnelle réversible catalysée *in vivo* par une famille d'enzymes nommées kinases. Le groupement phosphate, provenant de l'ATP, est ajouté de manière covalente aux acides aminés contenant un hydroxyle dans leur chaîne latérale (S, T, Y). La réaction inverse, c'est-à-dire l'hydrolyse du phosphate, est effectuée par des enzymes nommées phosphatases (**Figure 6**). Certaines enzymes possèdent une grande spécificité pour leur substrat, alors que d'autres agissent d'une façon moins

discriminatoire. La phosphorylation est une modification post-traductionnelle majeure et conservée au cours de l'évolution comme en témoigne le nombre important de kinases et de phosphatases exprimés (76, 77) ainsi que la grande quantité de substrats protéiques (78). En effet, les organismes exploitent la phosphorylation/déphosphorylation pour réguler des réactions et processus cellulaires fondamentaux (79, 80).

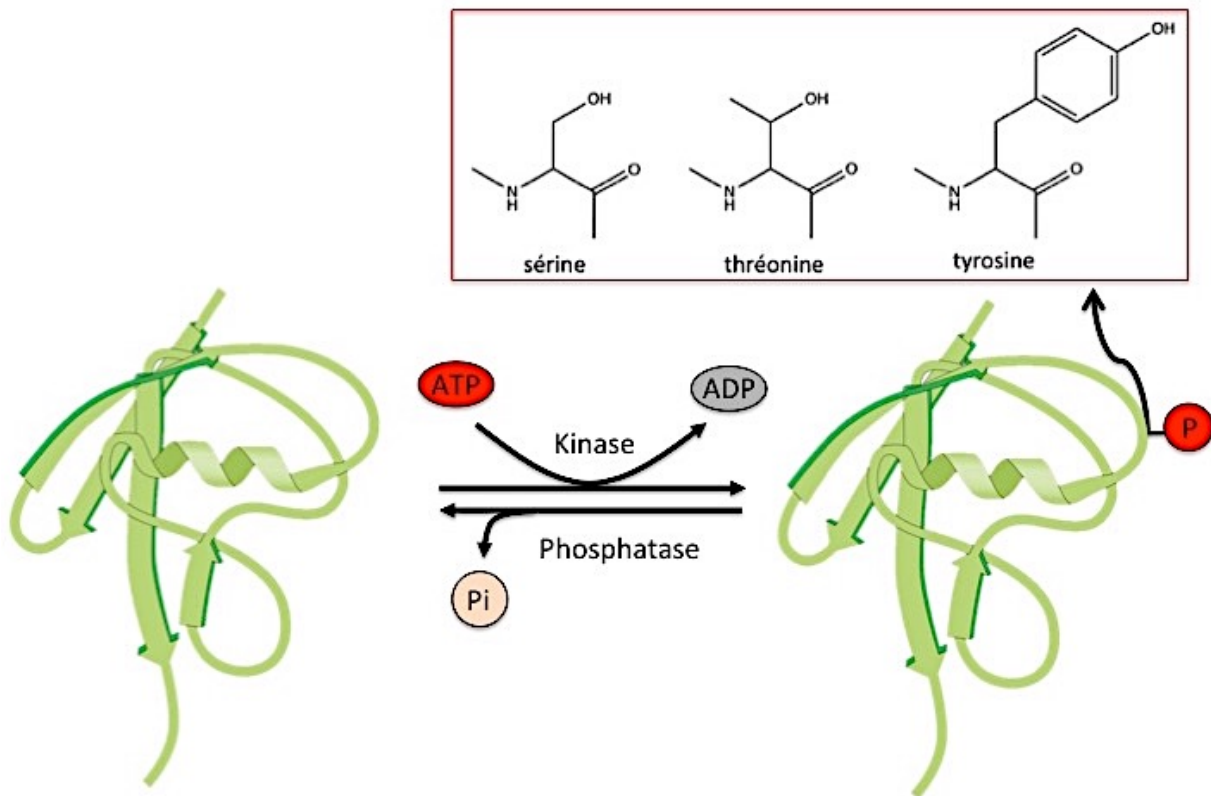


Figure 6. Mécanisme général de la phosphorylation *in vivo*. Le substrat (protéine verte représentant ubiquitine en tant qu'exemple) est phosphorylé (P) par une kinase sur un ou plusieurs résidu(s) sérine, thréonine et/ou tyrosine. Le groupement phosphate provient d'une molécule d'ATP qui est ensuite relâchée sous forme d'ADP. La déphosphorylation est catalysée par une phosphatase, ce qui génère une molécule de phosphate inorganique (Pi).

1.1.2.2. Cas spécifique de l'enzyme CK2

La famille des enzymes homologues Caséine Kinase (CK) tombe dans la catégorie des kinases possédant une grande variété de substrats. Un membre de cette famille chez les Eucaryotes est Caséine Kinase 2 (CK2). CK2 est un tétramère composé de deux sous-unités catalytiques (nommées α et α' chez *Homo sapiens*) et deux sous-unités régulatrices (chez

l'humain on retrouve deux copies d'une même sous-unité régulatrice β). Chez les organismes eucaryotes, CK2 est constitutivement active et distribuée de manière ubiquitaire dans les différents types et compartiments cellulaires. Cette kinase est un acteur clé dans plusieurs processus cellulaires et ceci peut s'expliquer par sa grande diversité de substrats et sa localisation étendue dans la cellule. Des expériences biochimiques et des analyses de séquences ont permis la catégorisation de CK2 en tant que sérine/thréonine kinase (81). L'enzyme reconnaît le motif consensus S/T-X-X-D/E, où S/T est le résidu ciblé par la phosphorylation, X est n'importe quel acide aminé, et D/E est un acide aminé chargé négativement (82). Parmi les substrats connus de CK2 sont retrouvés plusieurs protéines contenant des SIMs, telles que le suppresseur de tumeur PML et les E3 ligases de la famille PIAS. Ces SIMs sont souvent phosphorylés sur plusieurs résidus sérine en raison de la présence de motifs consensus pour CK2 en C-terminal de la partie hydrophobe. De ce fait, la régulation de l'activité de CK2 peut être exploitée comme un moyen de réguler les interactions SUMO/SIM (61).

1.1.2.3. Utilisation de la SUMOylation et des interactions SUMO/SIM comme mécanisme régulateur

Il a été montré que la phosphorylation régule à la fois la SUMOylation, en modulant la conjugaison de SUMO aux substrats, et les interactions non-covalentes entre les protéines SUMO et les phospho-SIMs. La protéine PML est un exemple de régulation positive du substrat en plus d'être une protéine particulièrement intéressante de part sa dualité fonctionnelle : PML est SUMOylée et interagit de manière non-covalente avec les protéines SUMO via son phospho-SIM. Suite à un stress génotoxique, PML peut être phosphorylée par ERK1/2 ou HIPK2 ce qui mène à sa SUMOylation subséquente. Dans ce cas de figure, la phosphorylation survient sur un résidu ne faisant pas partie du SIM. HIPK2 phosphoryle les sérines 8, 36 et 38 de PML (83) ce qui promeut la SUMOylation de PML aux lysines 65, 160 et 490 (84). L'explication moléculaire appuyant ces observations est que la phosphorylation de ces résidus mène à un changement de conformation dans PML, ce qui expose les sites de modification par SUMO offrant un accès à la machinerie de SUMOylation. Également, la phosphorylation du SIM par l'enzyme CK2 active résulte en une affinité de liaison supérieure entre PML-SIM et SUMO1 (61, 75, 85). Cette interaction SUMO/SIM se caractérise par des

liaisons électrostatiques entre les phospho-sérines de PML-SIM et des lysines chargées positivement à la surface de SUMO1 (85).

1.1.3. Acétylation

1.1.3.1. Voies enzymatiques

Chimiquement, l'acétylation est définie comme l'addition d'un groupement acétyle sur une molécule organique (**Figure 7**). Les protéines peuvent être acétylées par voies non-enzymatiques *in vitro*, mais, *in vivo*, elles sont modifiées par des enzymes communément appelées acétyltransférases. Les *N*-acétyltransférases catalysent le transfert d'un groupement acétyle issu du métabolite acétyl-CoA sur les amines terminales de résidus lysines cibles. Il y a de nombreuses acétyltransférases regroupées en superfamilles enzymatiques selon leurs caractéristiques structurales et leur localisation cellulaire (86). Deux exemples de superfamilles sont les histones acétyltransférases (HATs) et les lysines acétyltransférases (KATs). La conversion d'une lysine chargée positivement en *N*-acétyl-lysine neutre peut altérer la structure de protéines cibles et avoir des effets sur leurs interactions avec d'autres biomolécules. Les enzymes responsables de la réaction inverse, c'est-à-dire de l'hydrolyse du groupement acétyle des protéines cibles, font partie d'un petit groupe d'enzymes nommées déacétylases. Tout comme les acétyltransférases, les déacétylases sont regroupées en superfamilles selon leur structure et séquence. Notamment, les histones déacétylases (HDACs) et les lysines déacétylases (KDACs) sont les antagonistes en terme de fonction des HATs et KATs, respectivement. Deux différentes classes de HDACs évolutivement conservées ont été décrites (HDAC1 et HDAC2), avec quelques sous-classes, mais leur nombre est beaucoup plus limité que celui des HATs (87).

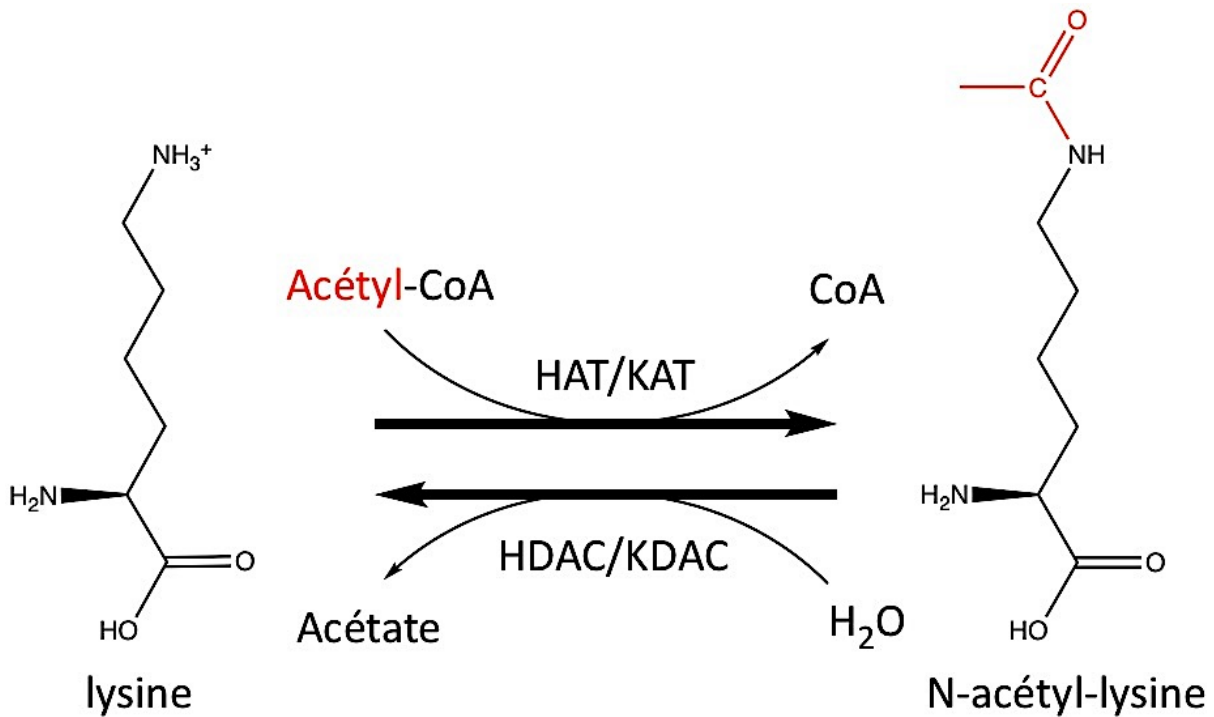


Figure 7. Mécanisme général de l'acétylation *in vivo*. Un groupement acétyle provenant du métabolite **acétyl-CoA** est ajouté de manière covalente à la lysine cible via l'activité acétyltransférase des HATs/KATs. L'hydrolyse (H_2O) de l'acétyl-lysine est catalysée par les HDACs/KDACs et le groupement acétyle est libéré sous forme d'acétate. À pH physiologique, la lysine est chargée positivement, alors que l'acétylation engendre la neutralisation de la basicité de la chaîne latérale. Figure générée avec *ChemDraw Professional 16.0*.

1.1.3.2. Acétylation des protéines de la famille SUMO

En plus d'être une modification post-traductionnelle à proprement parler, les protéines SUMO peuvent elles-mêmes être modifiées post-traductionnellement (concept de « *modified modifier* ») (88). Des analyses de spectrométrie de masse ont déterminé que des lysines spécifiques des protéines SUMO1 et SUMO2/3 sont ciblées pour acétylation *in vivo* (89, 90). Un groupe a démontré que l'enzyme CBP/p300 est en mesure d'acétyler SUMO1 *in vitro* aux lysines 37, 39, 45, 46, et 48 (**Figure 8A**). Ces résultats ont ensuite été corroborés *in vivo* par des études de spectrométrie de masse. L'acétylation du SUMO1 survient majoritairement aux résidus K37 et K39, seuls ou en combinaison, mais l'acétylation des résidus K45, K46, et K48 a également pu être détectée (63, 89). Il semble donc que l'acétylation des résidus K37 et K39 puissent jouer un rôle important dans la régulation des interactions SUMO/SIM puisque ces résidus sont conservés à travers les différents paralogues de SUMO ainsi que chez la protéine

de levure SMT3. Cependant, chez le ver, seul K39 est conservée alors que K37 est remplacée par une arginine. La lysine 45 n'est conservée chez aucun paralogue de SUMO, alors que K46 est conservée à travers les paralogues humains et chez le ver, mais remplacée par une arginine dans la levure (**Figure 2**). De plus, puisque l'acétylase CBP/p300 est en mesure de modifier les lysines identifiées dans les essais *in vitro*, il semble plausible que cette enzyme soit responsable de l'acétylation des protéines SUMO *in vivo*. Il a aussi été montré que l'acétylation de SUMO est contrôlée par les HDACs (63). Les protéines SUMO subissent l'acétylation lorsqu'elles sont conjuguées à leur substrat, ce qui confirme l'importance fonctionnelle de cette modification. Il a aussi été observé que l'acétylation des protéines SUMO affecte leurs interactions avec certains SIMs (**Figure 8B**), ce qui occasionne des conséquences observables sur certains processus cellulaires (dynamique des PML-NBs, apoptose et transcription médiés par p53).

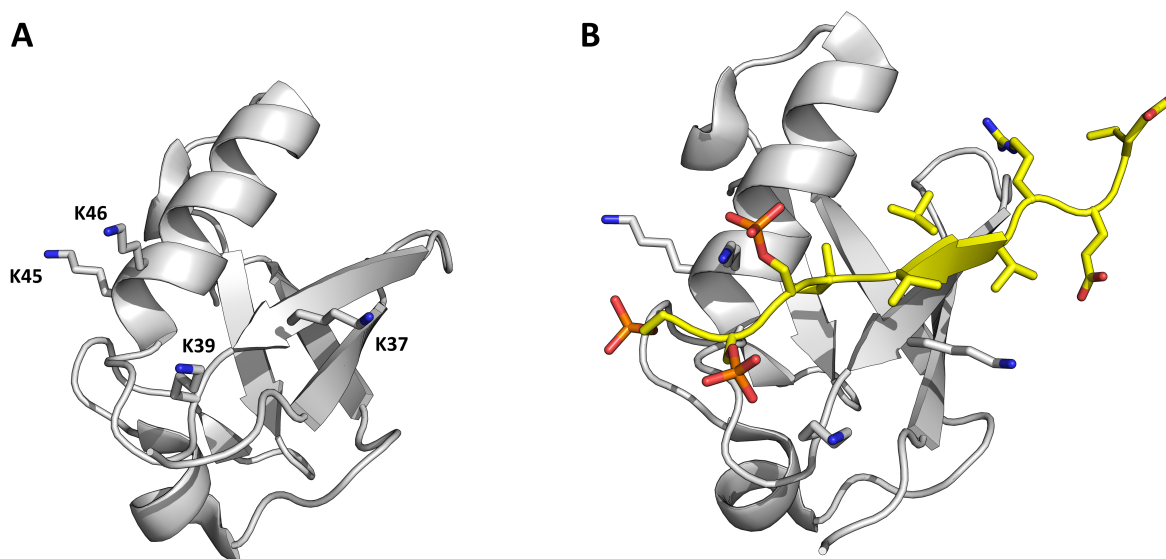


Figure 8. Lysines acétylées et implication dans les interactions SUMO/SIM. (A) Identification des lysines de SUMO1 (gris, PDB ID [4WJN](#)) pouvant être modifiées par acétylation et étant localisées à l'interface d'interaction avec les SIMs. **(B)** Exemple du complexe entre SUMO1 (gris) et le SIM de PML sous forme phosphorylée (jaune, PDB ID [4WJN](#)).

1.1.3.3. Modulation potentielle des interactions SUMO/SIM

Quelques-unes de ces modifications ont été étudiées dans un contexte biologique en partant de l'hypothèse que l'acétylation des protéines SUMO interfère avec les interactions SUMO/SIM via la neutralisation de la charge positive de résidus lysines en surface de SUMO. La lysine K39 de SUMO1 et son équivalent K35 chez SUMO2 sont impliquées dans des interactions électrostatiques avec plusieurs phospho-SIMs. Puisqu'il a été démontré que K39 est modifiée par acétylation *in vivo*, il est possible que la neutralisation de la charge positive de cette lysine puisse avoir un impact sur les interactions SUMO/SIM (61, 66, 91). De plus, il a été démontré que les charges positives des lysines 37 de SUMO1 et 33 de SUMO2 sont requises dans leur interaction avec certains SIMs tels que ceux appartenant à la famille des protéines PIAS (61, 92). Il a ainsi été suggéré que cette diminution d'affinité entre les protéines SUMO et les séquences SIMs de diverses protéines est seulement observée si le résidu lysine acétylé sur SUMO est impliqué dans des ponts salins avec des résidus négativement chargés du SIM. Structuralement, les lysines 39 et 46 de SUMO1 sont impliquées dans des interactions ioniques directes avec les phospho-sérines du SIM de PML, ce qui suggère un rôle prédominant de ces deux résidus dans le contexte de l'acétylation (85). De plus, des études moléculaires plus approfondies ont déterminé que l'acétylation des lysines de SUMO interfère dans les interactions non-covalentes entre SUMO et les SIMs de PML, de PIAS1, et de Daxx. Par conséquent, l'acétylation de SUMO modifie la dynamique de structures comme les PML-NBs et altère la répression transcriptionnelle médiée par SUMO (63). Également, il a été démontré que le bromodomaine de CBP/p300 reconnaît spécifiquement SUMO2 acétylée en position K33. Il est ainsi suggéré que l'acétylation des protéines SUMO affecte dans un premier temps leur interaction avec les SIMs et permettrait dans un deuxième temps la liaison à des protéines contenant des bromodomains, ou encore l'accès aux enzymes de désSUMOylation (63).

1.2. Cancer et suppresseurs de tumeur

1.2.1. La protéine *ProMyelocytic Leukemia* (PML)

1.2.1.1. Fondement génétique et implication dans les maladies

Chez l'humain, la protéine PML est encodée par un gène situé sur le chromosome 15 qui appartient à la famille *TRi partite Motif* (TRIM). Les membres de la famille TRIM possèdent plusieurs caractéristiques structurales distinctes collectivement connues sous l'appellation de motif RBCC : trois domaines liant le zinc (domaine RING, B-box de type 1 et 2) et une région superhélice (*coiled-coil*, CC). La protéine PML est présente sous diverses isoformes (PML-I à PML-VII) produites par épissage alternatif de l'exon en 3' du transcrit du gène *PML* (**Figure 9**). Le motif RBCC est retrouvé chez toutes les isoformes de PML et est essentiel à la formation de larges complexes macromoléculaires via sa capacité d'homomultimérisation (93). Plusieurs membres de la famille TRIM/RBCC sont des ubiquitine ligases qui peuvent s'auto-assembler afin de former des sous-structures cellulaires (94). De plus, la présence d'un signal de localisation nucléaire (*Nuclear Localization Signal*, NLS) chez toutes les isoformes (à l'exception de PML-VII) explique la prédominance nucléaire de ces protéines. On pense que les variations de séquences en C-terminal des isoformes PML génèrent peut-être différentes interfaces d'interaction, ce qui influencerait leur spécificité de fonction. L'isoforme de la protéine la plus étudiée est PML-IV en raison de son implication dans la sénescence, la différenciation et l'apoptose cellulaire (95-97).

Une condition pathologique nommée leucémie promyélocytaire aigue (*Acute Promyelocytic Leukemia*, APL) est causée par une translocation chromosomique du gène *PML* avec le gène codant pour le récepteur de l'acide rétinoïque α (*Retinoic Acid Receptor alpha*, RAR α). La protéine de fusion qui en résulte, PML/RAR α , perturbe la formation de multimères contenant PML dans le noyau (98). Il est intéressant de noter que l'acide rétinoïque ainsi que le trioxyde d'arsenic (As_2O_3) promeuvent la dégradation de la protéine de fusion PML/RAR α dans des conditions cellulaires. Ainsi, ces composés sont utilisés comme agents thérapeutiques pour les patients présentant cette maladie (99).

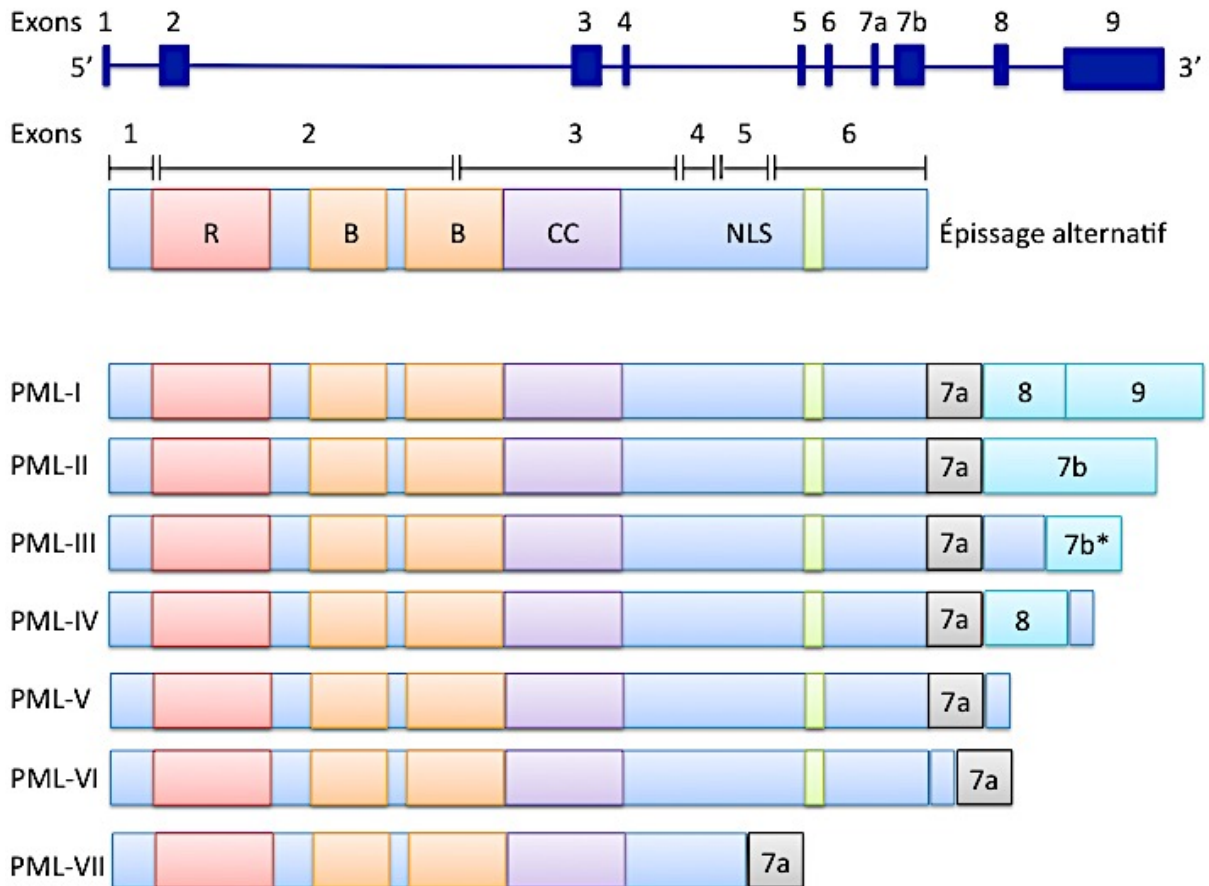


Figure 9. Représentation schématique des exons de la protéine encodée par le gène *PML* et des différentes isoformes générées par épissage alternatif. Le gène *PML* humain contient 9 exons qui peuvent être épissés différemment afin de générer sept isoformes distinctes. La plupart des isoformes de la protéine PML se retrouve principalement dans le noyau en raison de la présence d'un NLS dans l'exon 6. Cependant, un épissage des exons 4-6 a été reporté, ce qui rend possible l'existence de formes cytoplasmiques de la protéine. Toutes les isoformes contiennent par contre les exons 1-3 où se retrouve le motif TRIM/RBCC. *Figure adaptée de : (65).*

1.2.1.2. Corps nucléaires contenant PML

La protéine PML a initialement été étudiée pour ses fonctions dans la suppression tumorale, mais son implication dans la régulation de processus tels que le métabolisme et la survie cellulaire a également été caractérisé (100). PML exerce ses diverses fonctions notamment via l'organisation de superstructures nucléaires portant le nom de corps nucléaires contenant PML (*PML-Nuclear Bodies*, PML-NBs) (101). Ces structures sont retrouvées dans l'espace entre les chromosomes et sont conjuguées solidement à la matrice nucléaire (102) où

elles s'associent de manière non aléatoire avec des régions riches en gènes et transcriptionnellement actives du génome (103). La quantité et la taille de ces structures peuvent être altérées via le contrôle du niveau de transcription du gène *PML* (101, 104) en réponse à des stress cellulaires tels qu'une infection virale (105), un choc thermique, une exposition à des métaux lourds, un stress oxydatif, ou encore une transformation cellulaire (106). Les PML-NBs sont des structures hétérogènes et dynamiques où transitent divers facteurs impliqués dans la régulation des processus biologiques (**Figure 10**). Un point commun entre PML et les protéines associées aux PML-NBs est qu'elles sont souvent des cibles de la SUMOylation et/ou qu'elles contiennent des SIMs dans leur séquence (65). Ainsi, les PML-NBs sont impliqués dans une grande variété de processus cellulaires où leurs rôles pourraient être d'agir en tant que 1) structure nucléaire séquestrant et relâchant des protéines, 2) surface catalytique pour les modifications post-traductionnelles des protéines qui résident ou transitent à l'intérieur des corps nucléaires, et/ou 3) site actif de transcription ou de régulation de la chromatine (106, 107).

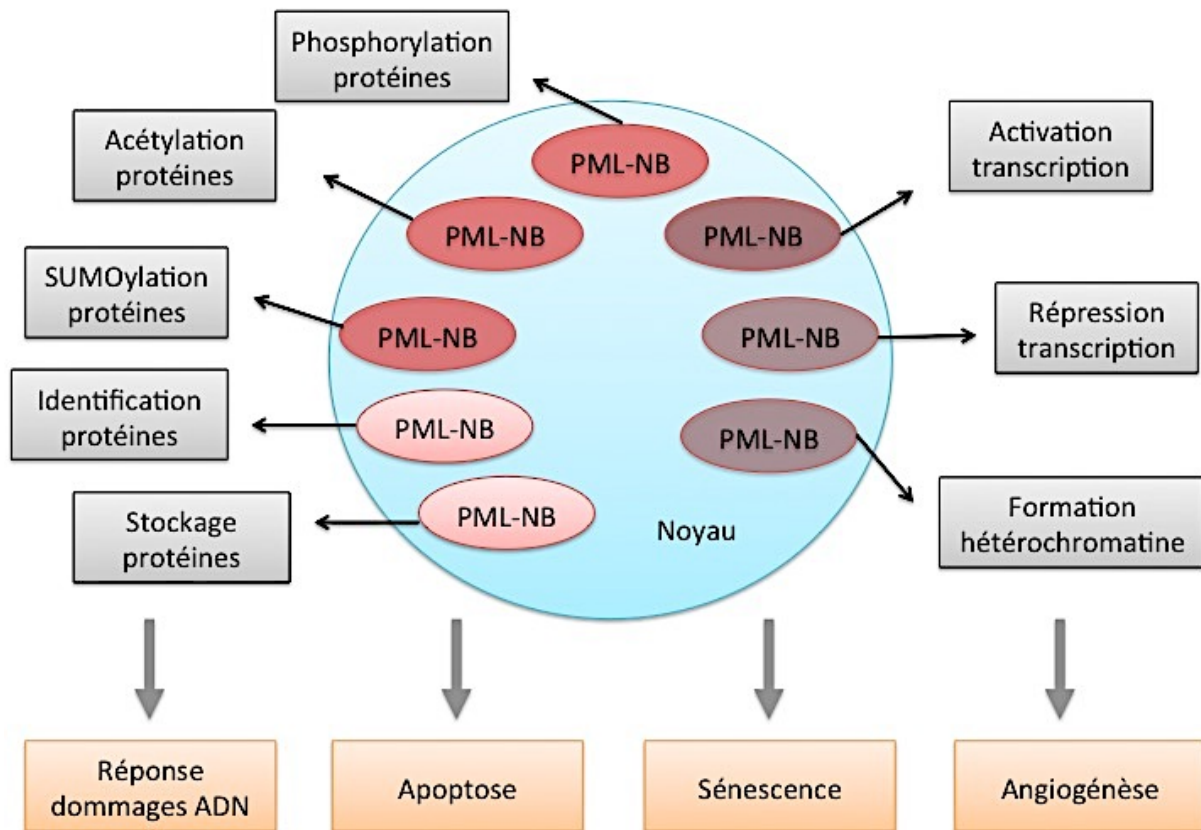


Figure 10. Schématisation des multiples fonctions des PML-NBs. Les fonctions cellulaires attribuées aux PML-NBs (rectangles oranges) sont multiples : réponse aux dommages à l'ADN, contrôle de l'apoptose, de la sénescence et de l'angiogénèse. Ces structures parviennent à réguler toutes ces fonctions par divers moyens pouvant être classifiés en trois catégories : identification et stockage de protéines ; modifications post-traductionnelles de protéines ; et régulation des activités nucléaires (activité transcriptionnelle de gènes, organisation de la chromatine, etc.). Les PML-NBs (cercles rouges) sont représentés en trois différentes teintes afin de souligner leur potentiel hétérogénéité de localisation dans le noyau (cercle bleu) et de fonctions (rectangles noirs). *Figure adaptée de : (65).*

1.2.1.3. Régulation et modifications post-traductionnelles

1.2.1.3.1. SUMOylation

PML est modifiée par SUMO sur au moins trois lysines : K65 dans le domaine RING, K160 dans le B-box de type 1, et K490 dans le NLS. SUMO est ajouté de manière covalente à PML via l'interaction directe entre PML et l'enzyme UBC9 E2 SUMO ligase. Non seulement PML est modifiée par les trois paralogues de SUMO, mais elle co-localise également avec

eux, suggérant que PML possède un domaine capable d'interagir de manière non-covalente avec les protéines de la famille SUMO (108). La nature de ce domaine est désormais connue : il s'agit du SIM (**Figure 11**) retrouvé dans toutes les isoformes de la protéine à l'exception de PML-VII (**Figure 9**).

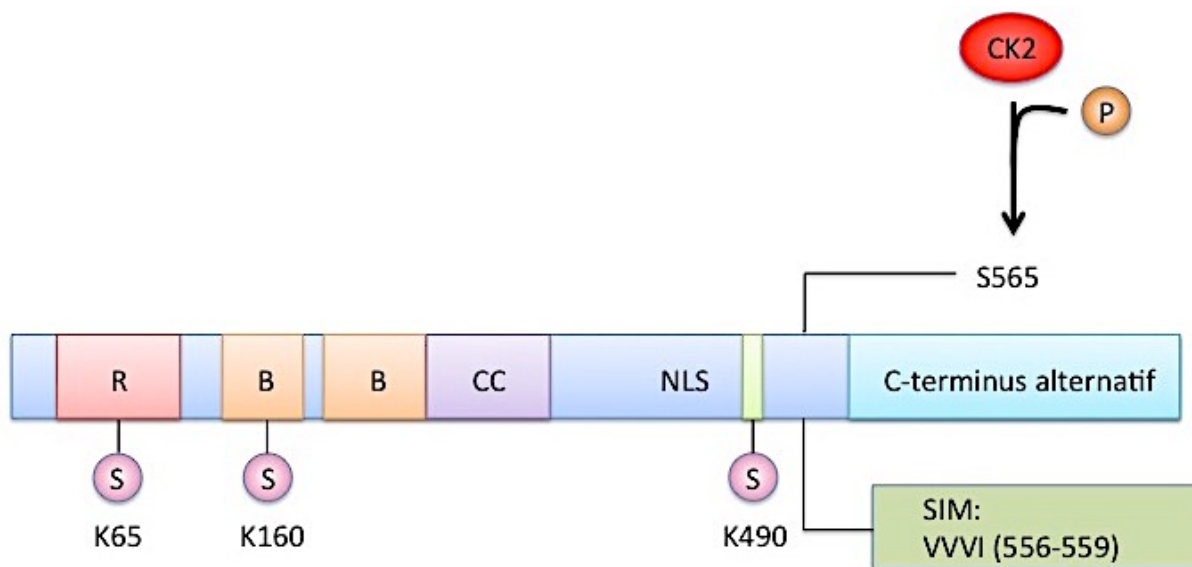


Figure 11. Représentation schématique de la forme générale de la protéine PML montrant les principaux domaines et modifications post-traductionnelles. L'enzyme CK2 phosphoryle PML à la position S565 en aval du SIM (résidus 556-559, séquence VVVI). Les trois lysines (K65, K160, K490) ciblées par SUMOylation chez PML sont également indiquées (sphères roses). Ces modifications post-traductionnelles sont communes à la plupart des isoformes de PML, puisqu'elles surviennent dans des régions conservées. R, RING finger ; B, B-box ; CC, Coiled-Coil domain ; NLS, Nuclear Localization Signal ; S, SUMO ; P, Phosphate. *Figure adaptée de : (65).*

Suite à ces découvertes, il a été démontré que la présence de sites de SUMOylation et d'un domaine SIM sur PML joue un rôle crucial dans le processus de nucléation des corps nucléaires, ainsi que dans le recrutement de partenaires d'interaction qui sont généralement SUMOylés et/ou qui possèdent des SIMs dans leur séquence (109). Il ressort de ces évidences que les interactions SUMO/SIM possèdent une importance clé dans la régulation des diverses fonctions des PML-NBs (**Figure 12**).

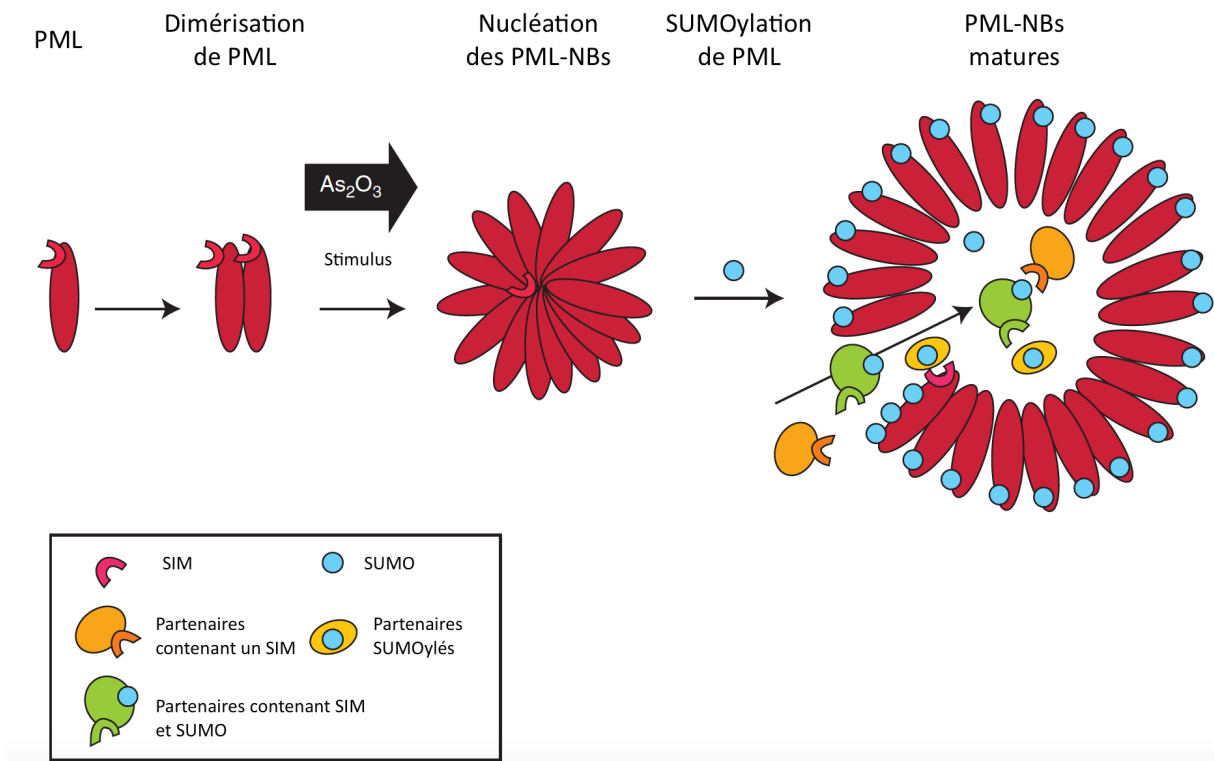


Figure 12. Représentation schématique d'un modèle de biogénèse des PML-NBs. En premier lieu, la protéine PML dimérise. S'ensuit une multimérisation des protéines qui engendre la nucléation des PML-NBs. Cette nucléation survient typiquement suite à un stress oxydatif (ici il s'agit de l'agent oxyde d'arsenic, As_2O_3). Il est proposé que la SUMOylation de PML mène à une organisation sphérique des corps nucléaires. Des partenaires protéiques autres sous forme SUMOylée et/ou possédant des SIMs sont ensuite recrutés dans ces structures via des interactions SUMO/SIM directes avec PML et/ou toutes autres facteurs co-localisant dans les corps nucléaires. *Figure adaptée de : (109).*

1.2.1.3.2. Phosphorylation

En plus d'être SUMOylée sur plusieurs lysines, PML est également la cible de kinases qui sont activées suite à des dommages à l'ADN ou lors de conditions de stress cellulaire. Il semble que la phosphorylation soit utilisée comme moyen de régulation de la stabilité et des fonctions apoptotiques de PML (65, 110) à travers le recrutement et/ou le relâchement de facteurs multiples qui transitent par les NBs (83, 111). Notamment, la phosphorylation de plusieurs sérines et thréonines de PML par des kinases régulées par des signaux extracellulaires (*Extracellular signal-Regulated Kinases*; ERKs) est préalable à la SUMOylation de PML (112). Également, la kinase proto-oncogénique CK2 est responsable de

la phosphorylation des quatre sérines adjacentes au SIM de PML (**Figure 11**). Ainsi, trois des résidus sérine (S560/561/565) font partie de séquences consensus de reconnaissance par CK2 et sont directement ciblés pour modification par la kinase. La modification de la quatrième sérine (S562) est potentiellement dépendante de l'état de phosphorylation de S565. Puisque ces modifications surviennent en aval du domaine SIM de PML, il est suggéré que l'apport de charges négatives par les groupements phosphates ajoutés pourrait réguler l'interaction de PML avec des résidus lysines conservées chez les protéines de la famille SUMO (85). De plus, la phosphorylation de PML par CK2 à la sérine 565 stimulerait l'ubiquitinylation subséquente de la protéine PML (110). La communication entre ces diverses modifications post-traductionnelles apporte une complexité supplémentaire aux processus de régulation cellulaires dépendant de PML. Ainsi, en plus d'être particulièrement pertinente dans un contexte pathologique comme le cancer, où CK2 tend à être hyperactive, la phosphorylation de PML par CK2 influence les propriétés de ses multiples substrats (*ex.* : stabilité de PML).

1.3. But du projet

1.3.1. Problématique et hypothèses

La protéine PML est un suppresseur de tumeur impliqué dans la régulation d'une grande variété de processus biologiques incluant la transcription, le contrôle du cycle cellulaire et l'intégrité génomique (113-115). Plusieurs des fonctions de PML sont attribuables à sa capacité de recrutement d'autres facteurs à l'intérieur de structures sous-nucléaires nommées PML-NBs. De nombreuses protéines ciblées par la SUMOylation transitent dans ces structures en réponse à divers stimuli (*ex.* : stress oncogénique, dommage par radiation, infection virale) (109). PML joue un rôle central dans la régulation de ce transit via l'établissement d'interactions non-covalentes entre son SIM et les protéines SUMOylées (60, 116). Le SIM de PML se compose de quatre résidus hydrophobes (VVVI) suivis d'une séquence riche en résidus acides et sérine (S560, S561, S562 et S565) peuvent être phosphorylées par la kinase CK2 (110, 117, 118). Il a été démontré que l'altération de l'état de phosphorylation du SIM de PML régule son affinité de liaison pour SUMO1 (61). En outre, il a été proposé que cette interaction SUMO/SIM pourrait être régulée via l'acétylation de résidus lysine conservés des protéines SUMO (63). De plus, une combinaison d'études

fonctionnelles, biochimiques et structurales au sein de notre laboratoire a permis la caractérisation de l'interaction entre SUMO1 et le SIM de PML au niveau atomique (85). Ces travaux ont établi que l'interaction entre SUMO1 et PML est grandement favorisée lorsque les résidus sérine du SIM de PML sont phosphorylés. Les données structurales de SUMO1 en complexe avec le SIM de PML dans sa forme non-phosphorylée ou tétra-phosphorylée indiquent que deux des quatre sérines du SIM (S560 et S561) établissent des interactions ioniques directes avec des lysines (K46, K39) de SUMO1 qui sont conservées chez les autres paralogues. En considérant toutes ces données, nous formulons l'hypothèse que l'acétylation des lysines conservées K39 et K46 de SUMO1 altère l'interaction avec le SIM de PML lorsque ce dernier est sous forme phosphorylée.

1.3.2. Objectifs de recherche

- Caractériser le rôle de l'acétylation des protéines de la famille SUMO sur leur affinité de liaison au SIM de PML.
- Déterminer l'effet de la phosphorylation et de l'acétylation sur la structure des complexes PML-SIM lié aux protéines de la famille SUMO.

1.4. Approche expérimentale

1.4.1. Production de protéines modifiées

1.4.1.1. Peptides phosphorylés

Puisque les protéines sont typiquement phosphorylées par des kinases spécifiques *in vivo*, il est possible d'exposer des peptides recombinants purifiés, tels que les SIMs, à ces mêmes kinases dans des essais *in vitro*. Plusieurs phospho-SIMs, incluant celui de la protéine PML, sont des cibles de l'enzyme CK2 (61, 91). Cependant, puisque le SIM de PML contient quatre sérines ciblées pour phosphorylation par CK2 (les sérines 560, 561 et 565 sont des sites consensus de modification, alors que la sérine 562 devient un site potentiel suite à la phosphorylation de S565), la complétion de cette réaction est relativement difficile à obtenir *in vitro*. Ceci rend particulièrement problématique la génération d'une quantité importante de peptide phosphorylé homogène requis pour les expériences thermodynamiques. Afin de

contourner ce problème, des variants des résidus sérine considérés comme imitant la charge négative associées à la phosphorylation peuvent être employés. Pour se faire, le codon sérine est muté dans la séquence d'ADN du peptide d'intérêt en codon correspondant à l'acide aspartique (Asp, **D**) ou l'acide glutamique (Glu, **E**). Ainsi, les résidus sérine en question sont substitués par les deux acides aminés naturels chargés négativement à pH physiologique, ce qui mime partiellement l'effet de charge apporté par la phosphorylation.

1.4.1.2. Protéines acétylées

Il est maintenant clairement établi que de nombreuses modifications post-traductionnelles ciblent les protéines de la plupart des organismes vivants et altèrent ainsi leurs fonctions biochimiques. Le potentiel de pouvoir diriger l'incorporation d'acides aminés modifiés post-traductionnellement dans des protéines et organismes variés durant leur synthèse *in vivo* permet d'étudier plus précisément les conséquences de ces modifications sans avoir recours à des essais enzymatiques *in vitro*. Différentes approches sont désormais disponibles pour investiguer l'impact de ces modifications sur les propriétés de protéines. L'une d'elles, l'expansion du code génétique, permet l'incorporation site-spécifique d'acides aminés non-naturels dans la séquence de protéines recombinantes. Un système aisément manipulable dans ce contexte est celui des archées du genre *Methanosarcina* : l'aminoacyl-ARNt synthétase est nommé PyIRS et encodé par le gène PyIS, alors que l'ARNt est nommé ARNt^{Pyl}_{CUA} et est encodé par le gène PyIT (119). Ainsi, des paires orthogonales aminoacyl-ARNt synthétase/ARNt_{CUA} reconnaissent un codon spécifique *amber* qui, suite à la transcription du gène d'intérêt, est adéquatement positionné dans la séquence de l'ARN messenger codant pour la protéine recombinante. Le codon *amber* est en fait le codon le moins exploité dans le code génétique des souches bactériennes utilisées pour la production de protéines recombinantes.

On considère les acides aminés non-naturels comme étant ceux qui ne sont pas naturellement incorporés lors de la traduction protéique effectuée par le ribosome. On compte dans cette catégorie les acides aminés naturels qui subissent des modifications post-traductionnelles tels que les lysines acétylées (acétyl-lysines), ou les sérines phosphorylées (phospho-sérines). Entre autres, l'acétyl-lysine a été génétiquement encodée dans des histones

recombinantes exprimées dans la bactérie *E. coli* en utilisant une paire acétyl-lysine ARNt synthétase/ARNt^{PyI}_{CUA} (120). Cette stratégie peut être adaptée pour la modification d'autres protéines en bactérie, mais également pour l'expression dans d'autres organismes (121).

Cette stratégie a donc été utilisée pour l'expression des protéines SUMO acétylées. Le codon *amber* stop UAG est incorporé à la place du codon de la lysine d'intérêt au sein de la séquence d'ADN des protéines SUMO. Le codon stop sera reconnu par la machinerie de traduction synthétique comme étant le signal d'incorporation de l'acétyl-lysine. Les enzymes nécessaires à cette incorporation sont fournies à la bactérie par le biais d'ADN plasmidique. Sur un des plasmides, identifié pCDF PyIT, se trouvent le gène codant pour la protéine d'intérêt (*ex.* : Histone H3, SUMO1, etc.), un gène de résistance à l'antibiotique streptomycine (SpectR), la séquence d'ADN correspondant à l'ARNt^{PyI}_{CUA} (PyIT) et le gène LacI permettant l'induction de l'expression de la protéine recombinante. Sur le second, identifié pBK-AcKRS3, se trouvent le gène codant pour l'acétyl-lysine ARNt synthétase (AcKRS) et le gène de résistance à l'antibiotique kanamycine (KanR). Les bactéries transformées et positivement sélectionnées sont inoculées en milieu de croissance liquide où de l'acétyl-lysine (AcK) et de la nicotinamide (NAM) sont ajoutés à la culture. La nicotinamide a pour rôle d'inhiber les déacétylases susceptibles d'hydrolyser le résidu lysine acétylé. Finalement, l'induction de l'expression permet la synthèse complète de la protéine SUMO contenant l'acétyl-lysine à la position définie (**Figure 13**).

Alternativement, il est possible d'imiter cette modification en remplaçant la lysine d'intérêt pour un acide aminé naturel dont les propriétés physico-chimiques sont reconnues pour être comparables à celles d'une lysine acétylée. La glutamine est l'acide aminé naturel ayant la structure la plus semblable à l'acétyl-lysine et son potentiel d'imitation de l'acétylation a d'ailleurs été démontré. Pour se faire, le codon lysine d'intérêt est substitué par un codon glutamine dans la séquence de l'ADN codant pour la protéine recombinante. Cette stratégie est particulièrement utile dans le contexte de l'étude des modifications post-traductionnelles *in vivo* lorsque le système de production de protéines modifiées n'est pas compatible ou adapté (63, 122).

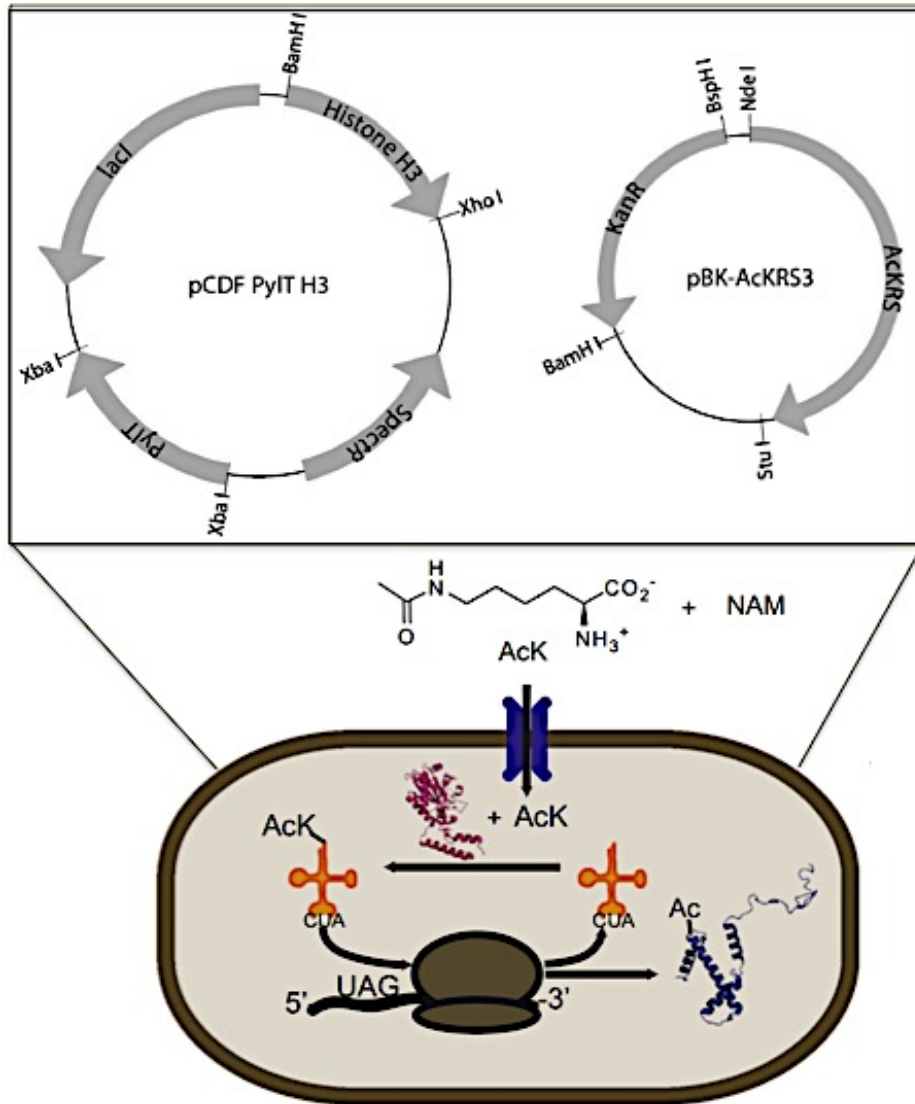


Figure 13. Production d'une protéine acétylée en bactérie utilisant l'expansion du code génétique. Les plasmides pCDF PyIT (exemple codant pour l'Histone H3, la résistance à la streptomycine, StrepR, l'aminocyl-ARNt transférase, PyIT, et l'inhibiteur de l'opéron lactose, lacI) et pBK-AcKRS3 (codant pour la résistance à la kanamycine, KanR, et l'ARN de transfert, AcKRS) sont transformés dans la bactérie d'expression (cercle brun) et ensuite les cellules sont positivement sélectionnées. Lors de la production de la protéine, de l'acétyl-lysine (AcK) et de la nicotinamide (NAM) sont ajoutés au milieu préalablement à l'induction de la protéine d'intérêt et diffusent à travers la membrane bactérienne (canal bleu). L'acétyl-lysine est lié de manière covalente à l'acétyl-lysine-ARNt synthétase (molécule rose) et cette enzyme catalyse ensuite le transfert de l'acétyl-lysine sur l'ARNt_{CUA} (molécule orange). L'ARNt chargé reconnaît le codon *amber* stop UAG sur l'ARNm de la protéine induite. Ainsi, l'acétyl-lysine est spécifiquement incorporée dans la chaîne peptidique au cours de la traduction menant à la production d'une protéine acétylée (molécule bleue). *Figure adaptée de : (120).*

1.4.2. Calorimétrie par titration isotherme

1.4.2.1. Généralités

La calorimétrie par titrage isotherme (*Isothermal Titration Calorimetry* ; ITC) est un moyen quantitatif utilisé pour évaluer les paramètres thermodynamiques (la constante d'association, K_a , la stœchiométrie, n , l'enthalpie de liaison, ΔH , sont calculés, alors que l'entropie, ΔS , est déterminée à partir de K_a) d'une interaction entre un ligand et un partenaire de liaison (Éq. 1). On mesure directement l'équilibre de l'interaction en déterminant le changement de température (chaleur émise ou relâchée) induit lors de l'association d'un ligand avec son partenaire de liaison (123). L'étude des interactions protéine-protéine a également une importance dans le repliement des protéines. On considère désormais que le repliement protéique est favorisé au niveau entropique par l'effet hydrophobe (enfouissement des résidus hydrophobes dans la partie centrale de la protéine). Expérimentalement, la détection d'une diminution de la capacité de chaleur corrèle proportionnellement avec l'exposition de la surface hydrophobe (124). Cependant, l'importance de l'effet hydrophobe dans les interactions protéine-protéine n'est toujours pas clair, puisque des molécules de solvant sont souvent observées à l'interface d'interaction (125).

$$\Delta G = \Delta H - T\Delta S = -RT\ln K_a = \frac{-RT\ln}{K_d} \quad \text{Éq. 1}$$

L'avantage principal de l'ITC est que le signal observé correspond directement à la chaleur générée ou absorbée lors de la formation du complexe, ce qui contraste grandement avec d'autres méthodes où les composantes protéiques doivent être modifiées par une étiquette fluorescente (ex. : polarisation de fluorescence) ou être immobilisées sur une plaque (ex. : résonance plasmon de surface, SPR). Ceci peut aussi être un inconvénient puisqu'il faut effectivement qu'une enthalpie de réaction soit détectée afin de permettre l'étude d'une interaction par ITC. Également, la constante d'association de complexes exhibant une très forte affinité (ex. : liaison anticorps-antigène) ne peut être évaluée précisément, mais l'enthalpie de liaison peut tout de même être déterminée. Une expérience d'ITC se conduit

assez aisément, mais des précautions sont de mise tout particulièrement lors de la manipulation des échantillons (*ex.*: détermination exacte de la concentration des macromolécules). Lorsque combiné à des informations structurales (pouvant provenir par exemple de données cristallographiques ou de résonance magnétique nucléaire, RMN), les données d'ITC permettent de bien appréhender les propriétés thermodynamiques régissant une interaction macromoléculaire. Ce type de données devient tout spécialement utile pour la conception de composés pharmacologiques et l'élucidation des facteurs déterminant la spécificité de ces interactions (123).

1.4.2.2. Aspect de l'appareil

À l'intérieur d'un instrument d'ITC se trouvent deux cellules (pour la référence et l'échantillon) composées d'un matériel conducteur de chaleur hautement efficace qui sont enrobées par des boucliers adiabatiques (**Figure 14, gauche**). Les différences de température entre la cellule de référence et la cellule de l'échantillon, ou entre les cellules et les boucliers adiabatiques, sont détectées par des circuits thermophiles/thermocouples sensibles. Une température constante entre toutes les composantes est maintenue grâce à des éléments thermorégulateurs. Lors de l'expérience, la solution de macromolécules est placée dans la cellule de l'échantillon alors qu'un standard (eau ou tampon) est placé dans la cellule de référence. Le titrant (typiquement une solution d'une autre biomolécule) est placé dans la seringue, dont l'extrémité aplatie est insérée dans la cellule de l'échantillon. Le titrant est périodiquement injecté dans la cellule dont le contenu est continuellement mélangé. Afin d'établir le signal de base, une puissance (flux énergétique) constante est appliquée sur la cellule de référence et ce signal guide le circuit de rétroaction menant à l'activation de l'élément chauffant sur la cellule de référence. La mesure directe effectuée par l'appareil correspond à l'ajustement de puissance requis pour le maintien d'une température équivalente entre les deux cellules (échantillon *versus* référence) en fonction du temps (**Figure 14, droite, haut**). La thermodynamique de la réaction de titration sera détectée par l'appareil s'il y a un dégagement (réaction exothermique, $\Delta G < 0$) ou une absorption (réaction endothermique, $\Delta G > 0$) de chaleur. Cette différence de chaleur est proportionnelle à la fraction de ligand lié : l'ampleur de cette différence diminue au fil de la réaction puisque la macromolécule de la cellule devient saturée par le ligand. Le circuit de rétroaction ajuste son alimentation en

puissance (désactivation lors de dégagement de chaleur et augmentation lors de l'absorption de chaleur) afin de maintenir la température dans la cellule de l'échantillon. L'enthalpie peut être évaluée mathématiquement et une courbe de l'énergie en fonction du ratio molaire est ainsi générée (**Figure 14, droite, bas**). La valeur du ratio molaire au point d'inflexion de la courbe correspond à la stœchiométrie de la réaction et la valeur de la pente correspond à la constante d'affinité (123).

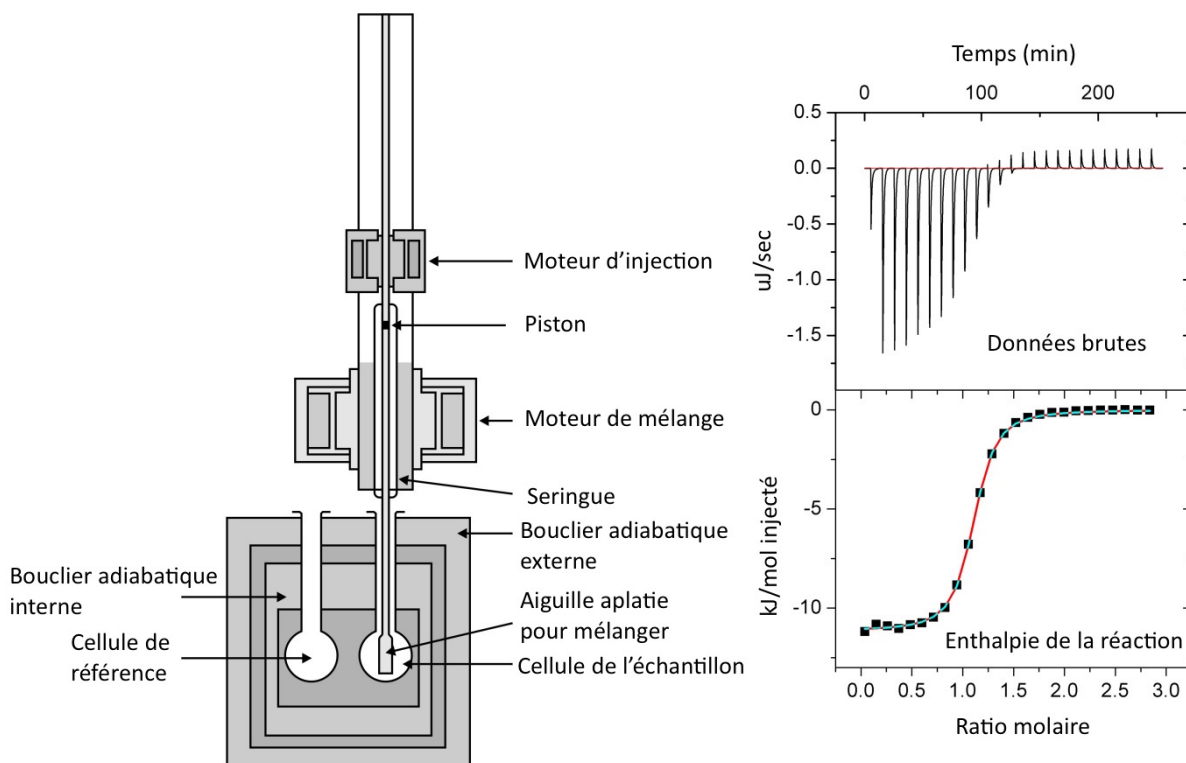


Figure 14. Représentation schématique de l'intérieur d'un calorimètre (gauche) et d'une expérience de titration typique (droite, haut) accompagnée de son évaluation (droite, bas). L'appareil calcule une différence d'enthalpie entre la réaction survenant dans la cellule de l'échantillon *versus* la valeur de référence. L'évaluation de l'enthalpie de la réaction génère une courbe à partir de laquelle la stœchiométrie (valeur du ratio molaire au point d'inflexion de la courbe) et l'affinité de liaison (pente de la courbe) peuvent être estimées. *Figure adaptée de :*

[HTTP://ELTE.PROMPT.HU/SITES/DEFAULT/FILES/TANANYAGOK/INTRODUCTIONTOPRACTICALBIOCHEMISTRY/CH08S07.HTML](http://elte.prompt.hu/sites/default/files/tananyagok/introductiontopracticalbiochemistry/ch08s07.html)

1.4.3 Cristallographie par diffraction des rayons X

1.4.3.1. Généralités

La cristallographie de macromolécules permet la détermination des structures tridimensionnelles (3D) des molécules biologiques. L'élucidation de la structure 3D de biomolécules (protéines et acides nucléiques) fournit des informations importantes sur leurs fonctions, leurs interactions, ainsi que sur l'implication de ces molécules dans des voies mécanistiques au sein d'organismes vivants variés. D'un point de vue fondamental, les structures cristallographiques approvisionnent nos connaissances de base sur la biologie, les mécanismes biochimiques et les pathologies en fournissant des données essentielles à la compréhension des mécanismes moléculaires (126).

Les cristaux de macromolécules se distinguent des autres types de cristaux par leur haut contenu en solvant et les faibles forces intermoléculaires supportant la maille cristalline. D'ailleurs, le succès de l'expérience de diffraction dépend de l'ordre cristallin. La structure 3D finale représente une moyenne dans le temps et l'espace de toutes les macromolécules composant le cristal. L'empaquetage du cristal biaise d'une certaine façon l'expérience, puisqu'une mince fraction du comportement dynamique des molécules sera échantillonnée lors de l'expérience où seul le désordre statique sera révélé (126). La grande majorité des structures résolues dernièrement ont été acquises à partir de cristaux maintenus à des cryotempératures loin des standards physiologiques. De basses températures sont utilisées afin de réduire les dommages induits par la radiation lors de l'acquisition des données de diffraction. Néanmoins, il a été montré que le positionnement du squelette carbone alpha n'est pas influencé par la température. Aussi, les cryotempératures diminuent l'agitation thermique dans le cristal, ce qui engendre une stabilisation des molécules d'eau du solvant et une meilleure définition des différentes conformations des chaînes latérales (127).

1.4.3.2. Pipeline

La résolution de structures de biomolécules s'effectue typiquement à travers un pipeline où l'étape limitante est l'obtention de cristaux diffractant de bonne qualité (**Figure 15**). Tout d'abord, une quantité suffisante de matériel pur et homogène (mono-dispersion, état

oligomérique simple) doit être produite très souvent sous forme recombinante provenant d'un extrait cellulaire (bactérie, levure, cellules humaines, etc.). La détermination des conditions de cristallisation s'effectue à l'aide de criblage de centaines de conditions. Les conditions chimiques (tampon, température, pH, concentration protéique, présence/absence de cations divalents, détergents, précipitants, additifs) sont variées dans le but d'obtenir un mélange optimal permettant la cristallisation. Les conditions positives sont sélectionnées et optimisées davantage dans le but d'obtenir des cristaux à fort potentiel diffractant. Le criblage est effectué de manière semi-automatisée, alors que l'optimisation des conditions de cristallisation est effectuée de manière empirique manuelle. La présence de cryoprotectant est nécessaire lors de la préparation du cristal pour la collecte de données effectuées à des cryotempératures. Le cristal est maintenu dans une boucle par tension superficielle, exposé au cryoprotectant, puis congelé dans l'azote liquide. Ce montage est ensuite utilisé pour la collecte des patrons de diffraction avec une source de rayons X (synchrotron) et un détecteur (CCD, Pilatus, etc.). Pour minimiser le bruit de fond et maximiser le ratio signal-sur-bruit, la taille du rayon incident est ajustée à celle du cristal. Les données de diffraction sont ensuite analysées par méthodes informatiques en utilisant des logiciels tels que ceux de la suite *Collaborative Computational Project Number 4 (CCP4)* (128) et *PHENIX* (129). Le processus comprend la réduction des données à travers le phasage et les calculs de la carte de densité électronique, l'interprétation de la carte (construction du modèle), le raffinement de la structure (complétion) et le dépôt dans la *Protein Data Bank (PDB)*. Les phases sont obtenues par remplacement moléculaire dans *Phaser* (130), par dérivatization séléno-méthionine (Se-Met) (131), ou par la détection de très faibles signaux anomaux. Suite à la génération d'un modèle initial, la structure doit être raffinée afin que le modèle atomique représente de manière optimale la densité expérimentale. La construction du modèle peut être effectuée manuellement avec l'aide de programmes tels que *Coot* (132) qui aide le processus itératif de raffinement et de reconstruction de modèle. Finalement, des outils sont disponibles pour la validation de la géométrie et de la qualité des structures avant soumission des coordonnées atomiques finales dans la PDB (133).

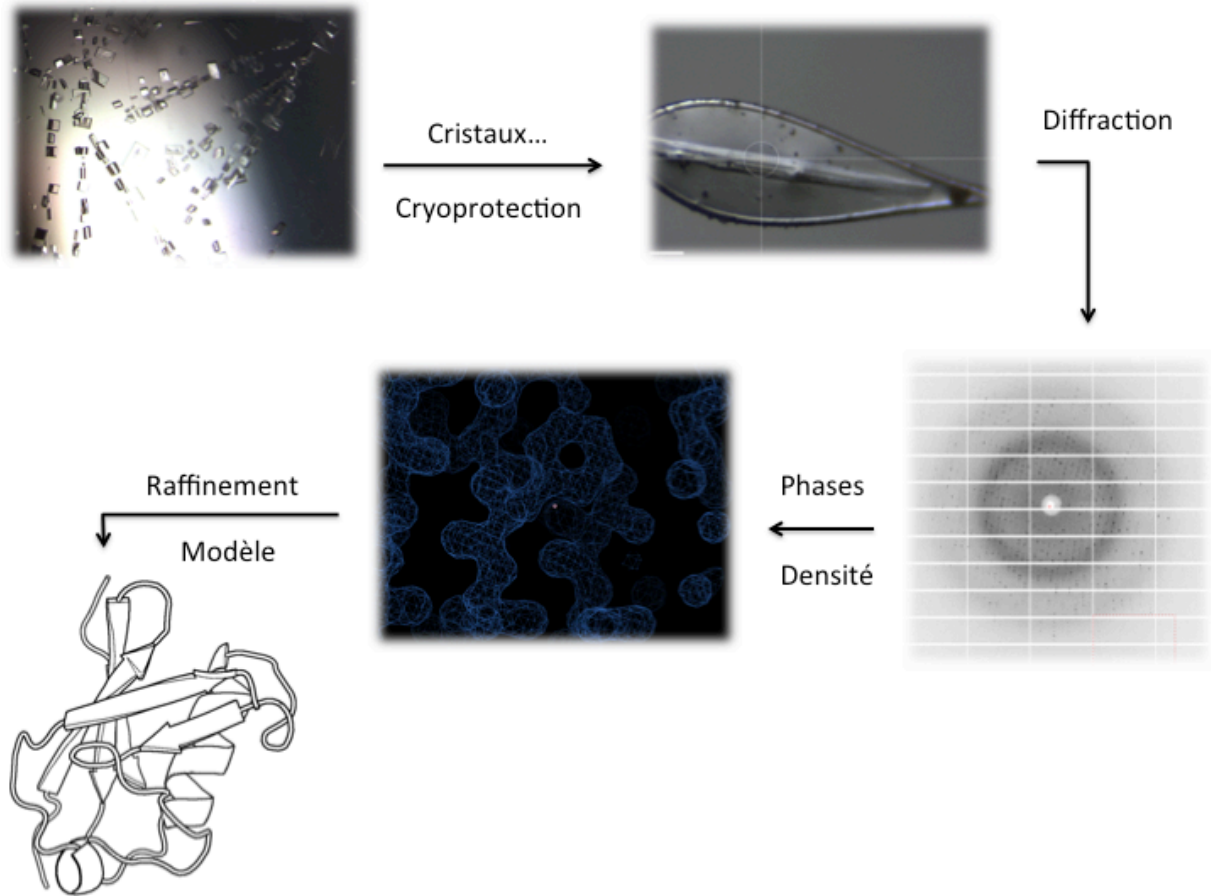


Figure 15. Schématisation des principales étapes de la résolution de structures macromoléculaires utilisant la cristallographie par diffraction de rayons X. Les protéines sont cristallisées dans différentes conditions. Les cristaux sont ensuite prélevés à l'aide d'une boucle de nylon, puis cryo-protégés dans l'azote liquide. Le cristal est exposé aux rayons X et les patrons de diffraction sont collectés par un détecteur. Les diffractions sont traitées informatiquement et les phases sont récupérées, ce qui permet la génération de la densité électronique des atomes formant le cristal. Un modèle initial est déterminé à partir des données expérimentales puis raffiné itérativement jusqu'à l'obtention d'un modèle final représentant la structure à déposer.

**2. Structural and biophysical characterization of the
impact that SUMO1 acetylation has on its
phosphorylation-dependent interaction with PML**

Structural and biophysical characterization of the impact that SUMO1 acetylation has on its phosphorylation-dependent interaction with PML

Christina Gagnon¹, Mathieu Lussier-Price¹, Haytham Wahba¹, Xavier H. Masclé¹, James G. Omichinski^{1}*

¹ Département de Biochimie et Médecine Moléculaire, Université de Montréal, Montréal, QC, Canada

***Correspondence:**

James G. Omichinski, Département de Biochimie et Médecine Moléculaire, Université de Montréal C.P. 6128 Succursale Centre-Ville, Montréal, QC H3C 3J7 Canada.

Email: jg.omichinski@umontreal.ca

Phone: 514-343-7341

Fax: 514-343-2210

Running Title: Structure of the Phospho-SIM of PML Bound to Acetylated SUMO1 Variants

Detailed authors contribution

Christina Gagnon performed cloning, protein expression and purification, sample preparation for each experiment entitled, protein crystallization, X-ray data collection, as well as structure solving and refinement. All figures were prepared by Christina Gagnon. The overall contribution of Christina Gagnon to this work is approximately 80-90% of all generated data in addition to writing the paper with the help of James Omichinski. Mathieu Lussier-Price taught Christina Gagnon how to perform cloning, protein purification and ITC experiments (sample preparation, data acquisition and analysis). Haytham Wahba taught Christina Gagnon how to conduct crystallography experiments (sample preparation, crystallization, data collection and analysis) and how to use the softwares for data collection and processing as well as graphical tools for figure making. Xavier H. Mascle produced and purified phosphorylated SIMs, which were used in the X-ray experiments. James G. Omichinski conceived the project, designed the study and supervised the writing of the paper.

Summary

Several PML-Nuclear Bodies (PML-NBs)-resident proteins are SUMOylated and/or contain a SUMO-Interacting Motif (SIM). The non-covalent interaction between SUMO proteins and SIMs is modulated by phosphorylation of SIMs and acetylation of key residues on SUMO proteins. SUMO acetylation at key sites interferes with PML-NBs functions and dynamics, but little is known about how this occurs at the molecular level. We examined the role SUMO1 acetylation plays on its binding to the phospho-SIMs of PML and Daxx. Binding affinity between SUMO1 and phospho-mimetic PML-SIMs was significantly reduced when SUMO1 was acetylated at positions K37, K39, K45 or K46, with K39 and K46 modifications having the most important impact. The crystal structures of the SUMO1 acetylated (Ac) variants bound to phospho-PML-SIM show dramatic disruptions in the K46Ac complex and subtler reorganization for the others. The crystal structure of phospho-Daxx-SIM in complex with SUMO1 K46Ac also displays important changes distinct from those observed with PML. Together, the results provide an atomic level description of how certain SUMO1 acetylation patterns provide means to regulate binding specificity to different phospho-SIMs in addition to demonstrating the great plasticity involved in SUMO/SIM interactions.

Introduction

Small Ubiquitin-like MOdifiers (SUMOs) are a family of proteins structurally related to ubiquitin, which are found predominantly in the nucleus of eukaryotic organisms (134). Humans express five SUMO paralogs (SUMO1-5) that can be divided into two subgroups based on their primary sequence similarities (135). The sequence of SUMO1 is highly homologous to that of both SUMO4 and SUMO5, whereas SUMO2 and SUMO3 share over 95% sequence identity, but only 45% with SUMO1. Like ubiquitin, all SUMO proteins can be conjugated to target lysine residues in substrates as a post-translational modification (PTM) through a process referred to as SUMOylation. The SUMOylation reaction requires an E1-activating enzyme (SAE1-SAE2), an E2-conjugating enzyme (UBC9), and in most cases an SUMO E3 ligase, such as the Protein Inhibitor of Activated STATs (PIAS) family members (17). Since the SUMO proteins seem to share highly overlapping functions in cells, it is still not yet clear why humans expressed multiple paralogs. However, it is well established that the SUMO proteins play important roles in regulating a wide array of fundamental cellular processes, including transcription, DNA repair, and ribosome biogenesis (136, 137). SUMO proteins typically carry out their regulatory functions by participating in protein-protein interactions either in their free, unconjugated form or following SUMOylation of a target protein. In both cases, the SUMO proteins are able to participate in non-covalent interactions with specialized SUMO-Binding Domains (SBDs) located in their binding partners (134).

To date, the best characterized SBD is the SUMO-Interacting Motif (SIM), which must minimally contain a short hydrophobic sequence of the form $\psi\psi\psi\psi$, $\psi\psi x\psi$, or $\psi x\psi\psi$ (where ψ is V, I, or L and x is typically D, E, S, or T) located in an intrinsically disordered region (IDR) of a protein (5). In addition, SIMs typically contain several negatively charged amino acids as well as potential phosphorylation sites (serine or threonine) located immediately adjacent to the hydrophobic core sequence either at its N- or C-terminal side. Upon binding to a SUMO-family protein, the SIM adopts a β strand conformation that is either in a parallel or an antiparallel orientation relative to the second β strand of the SUMO protein. It is believed that this binding orientation is dictated by whether the negatively charged amino acids are located on the N-terminal (anti-parallel) or C-terminal (parallel) side of the SIM hydrophobic core (5, 60).

In the 1960s, subnuclear spherical structures between 0.1-1.0 micron in diameter were observed by electronic microscopy in tumour cells. Subsequent studies revealed that the main component of these structures was the ProMyelocytic Leukemia (PML) protein, and these bodies are commonly referred to as PML-Nuclear Bodies (PML-NBs). The discovery of the PML protein as the major component of PML-NBs was in large part due to its implication in Acute Promyelocytic Leukemia (APL). In the context of APL, there is a dramatic change in the composition of PML-NBs, which is linked to the expression of a fusion protein resulting from chromosomal translocations between the *PML* gene on chromosome 15 and the *Retinoic Acid Receptor α* (*RAR α*) gene on chromosome 17 (PML-RAR α and RAR α -PML). Based on their cellular location, it was initially proposed that PML-NBs regulate nuclear functions, such as DNA replication and translation, as well as being involved in epigenetic silencing. Humans express seven different isoforms of the PML protein (PML-I to PML-VII) (93), and nucleation of PML-NBs is thought to occur via multimerization of PML proteins through their RBCC (Ring-finger, B-box, Coiled-Coil) motif, which is conserved in all seven isoforms. Following nucleation, additional components (109) are recruited to yield mature, heterogeneous ring-like shells (138). It is now known that PML-NBs function as recruitment and storage centers for a vast array of nuclear proteins and all of these proteins are either modified post-translationally by a SUMO protein or contain a SIM in their sequence.

PML itself is conjugated by SUMO proteins on at least three lysine residues (K65/160/490) (84) and six of its seven isoforms contain a SIM (56). The SIM of PML is composed of four hydrophobic residues (VVVI) followed by a cluster of acidic (D, E) and serine residues on the C-terminal side. Interestingly, PML functions can be regulated by changing the phosphorylation status of its SIM. This occurs in a Casein Kinase 2 (CK2)-dependent manner where the kinase phosphorylates PML at residues S560, S561, S562 and S565 (isoform I). In addition, CK2-dependent regulation has also been observed with other PML-NBs components that have SIMs containing phosphorylation sites such as Daxx and PIAS1 (116, 139-142). Structural studies with the phospho-SIMs of PML and Daxx demonstrated that the phosphorylated residues within their SIM make specific contacts with positively charged lysine residues that are highly conserved in SUMO family proteins and these interactions increase the binding affinity of SUMO1 towards these SIMs (60, 61, 63, 69, 91). In the SUMO1/PML-SIM interaction, phosphorylated serines (pS) 560, 561 and 562 of

the SIM form direct contacts with K46, K39 and H43 of SUMO1, respectively. In the case of the SUMO1/Daxx-SIM interaction, pS737 and pS739 make contacts with K46 and H43 of SUMO1 as seen in the PML-SIM complex, but there is no interaction observed between D738 of Daxx and K39 of SUMO1(85). These results suggested that different spacing patterns of the acidic and phosphorylated residues of the SIM play an important role in regulating the interaction with SUMO family proteins.

In addition to the effect of phosphorylation, it has been shown that SUMO/SIM interactions may be further regulated by acetylation of lysine residues in the SUMO protein (63, 89, 90). More specifically, it was shown *in vitro* that acetylation of SUMO1 at K37 interferes with its non-covalent interaction with several SIM-containing proteins (PML, Daxx, PIAS1), and that substituting K37 with a glutamine *in vivo* alters both the dynamics of PML-NBs formation as well as SUMO-mediated transcriptional repression. Therefore, it was postulated that acetylation of SUMO1 and phosphorylation of the SIMs by CK2 can form a complex regulatory network for regulating SUMO/SIM interactions (63). Given the key role of post-translational modifications in regulating SUMO/SIM interactions and the impact these pathways may have on biological functions, we have characterized the interactions between the phospho-SIMs of PML (PML-SIM-4PO₄) and Daxx (mDaxx-SIM-2PO₄) in complex with acetylated variants of SUMO1 (K37Ac, K39Ac, K45Ac, K46Ac) using a combination of Isothermal Titration Calorimetry (ITC) and X-ray crystallography. We establish that the interaction between SUMO1 and PML-SIM-4PO₄ is significantly reduced by acetylation of K37, K39, K45 and K46 of SUMO1, but that the variants have different effects depending on the acetylation site. Structural characterization of acetylated SUMO1 variants in complex with PML-SIM-4PO₄ or Daxx-SIM-2PO₄ indicates that acetylation of a specific lysine has a different impact on the positioning of phosphorylated residues of the SIMs relative to the positively charged residues of SUMO1. Together, these results provide an atomic level description of the role of both phosphorylation and acetylation in phospho-SIMs interaction with SUMO proteins, as well as important insights into how the acetylation of SUMO proteins impacts the affinity and specificity of the interaction with phospho-SIMs.

Results

Acetyl-mimetic variants of SUMO1 have lower affinities for phospho-mimetic PML-SIMs

Previous studies have demonstrated that phosphorylation of residues within SIM motifs enhances their affinity for SUMO proteins (85), whereas acetylation of lysine residues in SUMO proteins decreases their binding affinity to SIMs (63). In an attempt to biophysically characterize the opposing roles that these post-translational modifications have on the interaction between SUMO1 and the SIM of PML, we generated a series of acetyl-mimetic variants of Δ N-SUMO1 (residues 17-97 of human SUMO1) and phospho-mimetic PML-SIM peptides (residues 547-573 of human PML isoform I). To mimic lysine acetylation, the lysine (K) residues of interest were substituted by a glutamine (Q) residue, the natural amino acid possessing the closest structural homology to acetylated lysine. In the case of SUMO1, four lysine residues (**Figure 16A**) have been shown to be acetylated either *in vitro* or *in vivo* (63, 89, 90). Therefore, we generated four single K-to-Q variants of Δ N-SUMO1 (K37Q, K39Q, K45Q, K46Q). The four serine residues in the SIM of PML that undergo phosphorylation by CK2 (**Figure 16B**) were mimicked by substituting with either an aspartic acid (D) or a glutamic acid (E), the two natural amino acids that possess a negative charge at physiological pH. The four serine residues were substituted as a group to generate a PML-SIM-4SD (S560D/S561D/S562D/S565D) and a PML-SIM-4SE (S560E/S561E/S562E/S565E) peptide (**Figure 16B**). The Δ N-SUMO1 proteins were used to be consistent with our previous structural studies (85).

To biophysically characterize the effects that these mimetic substitutions have on the binding of SUMO1 to the SIM of PML, ITC experiments were performed. Control experiments were first conducted with the wild-type Δ N-SUMO1 protein and either the PML-SIM-4SD or the PML-SIM-4SE peptide to verify that the negatively charged amino acids are in fact mimicking phosphorylation. In these experiments, K_D values of 0.4 μ M (4SD) and 2.1 μ M (4SE) were obtained with the phospho-mimetic peptides (**Table I**) and these measured values are similar to the K_D value (1.2 μ M) previously reported between a tetra-phosphorylated PML-SIM peptide (PML-SIM-PO₄) and SUMO1 (85). Thus, we used the

phospho-mimetics for the ITC experiments due to difficulties in generating large quantities of the phosphorylated PML-SIM peptide. The ITC experiments were performed with the four Δ N-SUMO1 K-to-Q variants and the two phospho-mimetic peptides (**Figure 17**). In experiments with the PML-SIM-4SD peptide (**Figure 17A**), the K46Q and K39Q modifications in Δ N-SUMO1 reduce the binding affinity by a dramatic 32-fold and 17-fold, respectively, whereas a more modest decrease in binding affinity is observed with the K37Q (7-fold) and the K45Q (6-fold) modifications (**Table I**). Similarly, experiments with the PML-SIM-4SE peptide (**Figure 17B**) also show that there is a more dramatic decrease in binding affinity with the K46Q (21-fold) and the K39Q (10-fold) modifications of Δ N-SUMO1, than with the K37Q (6-fold) and K45Q (5-fold) modifications. Taken together, the results demonstrate that mimicking acetylation at any of the four lysine residues in the basic patches adjacent to the SUMO1/SIM binding site decreases the binding affinity to the phospho-mimetic PML-SIM peptides.

Acetylated Δ N-SUMO1 variants behave similarly to K-to-Q variants in ITC studies

Given that all four SUMO1 K-to-Q variants displayed lower binding to the phospho-mimetic PML-SIM peptides, we sought to determine whether acetylation at the same four sites causes similar changes in SUMO1 binding to the phospho-mimetic PML-SIM peptides. Four acetylated Δ N-SUMO1 variants (K37Ac, K39Ac, K45Ac, K46Ac) were produced using a bacterial system that allows for site selective incorporation of acetylated lysine (*143*) and the binding of the acetylated Δ N-SUMO1 variants to the PML-SIM-4SD and PML-SIM-4SE peptides was evaluated using identical ITC experiments. In studies with the PML-SIM-4SD peptide (**Figure 18A**), the K46Ac and K39Ac modifications of Δ N-SUMO1 reduce the binding affinity by 28-fold and 26-fold, respectively, whereas a more modest decrease in binding affinity is observed with the K37Ac (4-fold) and the K45Ac (6-fold) modifications (**Table II**). Similarly, experiments with the PML-SIM-4SE peptide (**Figure 18B**) also show that there is a more dramatic decrease in binding affinity with either the K46Ac (17-fold) or the K39Ac (17-fold) variant than with the K37Ac (4-fold) and the K45Ac (5-fold).

Taken together, the ITC experiments demonstrate that neutralizing the positive charge on the four lysine residues, either by substitution with a glutamine residue or acetylation, reduces significantly the affinity of SUMO1 for the phospho-mimetic PML-SIM peptides and the most dramatic reduction is observed when the positive charge is removed from either K46 or K39. These ITC results are consistent with our previously obtained structural data demonstrating that these two lysine residues make key electrostatic/ionic interactions with phosphorylated serine residues in the PML-SIM-4PO₄ peptide (85). In addition, these experiments demonstrate that the glutamine substituted SUMO1 proteins behave in a very similar manner as the acetylated SUMO1 proteins with respect to their relative binding to the phospho-mimetic PML-SIM peptides. This supports the use of glutamine to mimic acetylated lysine in cellular experiments that probe the interaction between SUMO1 and the SIM of PML.

Acetylation of SUMO1 does not alter binding to the hydrophobic core PML-SIM

To gain more detailed insights into the role that acetylation of lysine residues of SUMO1 have on its structure and binding to the SIM of PML, we initially crystallized the four acetylated Δ N-SUMO1 variants in complex with the unphosphorylated PML-SIM peptide (**Table III**). The crystals of the four complexes all belong to the P2₁2₁2₁ space group and contain either one or two copies of the complex in the asymmetric unit. The structures of the four acetylated Δ N-SUMO1 proteins in complex with the PML-SIM peptide are all virtually identical to the previously determined structure of the wild-type Δ N-SUMO1 in complex with the same PML-SIM peptide (**Figure 19**). These results are not unexpected since only the hydrophobic core of PML-SIM contributes to the binding interface in the absence of phosphorylation of the SIM of PML, as there are no interactions involving the positively charged lysine side chains of SUMO1. This is in contrast to what was observed with the PML-SIM-4PO₄ in complex with Δ N-SUMO1, where the positively charged K39 and K46 side chains form electrostatic interactions with two of the negatively charged phospho-serines. Based on these structures, we could observe that acetylation of SUMO1 causes only localized changes in the overall structure of the complex and this is consistent with the fact that these

four lysine residues are located on the surface of the protein. Thus, the structures of the four complexes will serve as controls for examining the structural effect that acetylation has on binding to the phosphorylated SIM of PML.

Crystallization trials and structure of the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complex

Given the dramatic reduction on binding affinity observed with the K46Ac and K39Ac SUMO1 variants in the ITC studies, we attempted to structurally characterize these acetylated proteins in complex with the PML-SIM-4PO₄ peptide. Despite numerous attempts, we were only able to obtain well diffracting crystals with the K39Ac variant. The Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ crystals (**Table IV**) belong to the P2₁2₁2₁ space group and diffract to a resolution of 1.45 Å. The asymmetric unit contains one copy of the complex (**Figure 20B**). Clear electron density is observed for residues 18-94 of SUMO1, as well as for residues 551-562 (**Figure 20A**) and 569-573 (**Figure 20C**) of PML. The structure of the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complex displays many similarities to the wild-type complex, but also several important differences. Most notably, the positions of the side chains of K46 and K45 change dramatically in comparison to their location in the wild-type complex. In the unacetylated form, the positively charged side chain of K46 is positioned 2.9 Å from the negatively charged side chain of pS560 whereas they are 5.5 Å apart in the K46Ac structure (**Figure 21**). In addition, K45 now appears to have the capacity to exist in two conformations where one of these conformations is equivalent to the position K45 adopts the wild-type SUMO1 structure. In the second conformation, the side chain of K45 moves in the direction of pS562 and the negative charge of pS562 is then positioned 4.2 Å from the positive charge of K45. This movement of K45 appears to be possible due to the loss of the ionic interaction between K46 and pS560. In contrast, the positions of the side chains of H43 and pS562 do not change significantly and the interaction between these two residues does not appear to be altered by acetylation of K46. Thus, pS562 appears to interact with both H43 and K45 in the structure with the SUMO1 K46Ac variant. In addition, K37 and K39 are located in virtually the same positions in both the wild-type and the K46Ac complexes. However, in the K46Ac complex, the negative charge from the side chain of pS561 moves 1.3 Å further away from the

positive charge on the side chain of K39, which results in a distance of 6.1 Å between these residues (**Table V**). Taken together, these movements in the basic loop of SUMO1, as well as the important displacements of the phospho-SIM, helps to explain the dramatic reduction of binding measured by ITC with the SUMO1 K46Ac variant.

Structure of the Δ N-SUMO1-K39Q:PML-SIM-4PO₄ complex

Well diffracting crystals were not obtained with the K39Ac variant and attempts were made to obtain crystals with the K39Q variant since ITC studies indicated that the K-to-Q SUMO1 variants can mimic acetylation. This method proved to be successful and permitted the generation of well diffracting crystals. The Δ N-SUMO1-K39Q:PML-SIM-4PO₄ crystals belong to the P2₁2₁2₁ space group and diffract to a resolution of 1.84 Å (**Table IV**). The asymmetric unit contains one copy of the complex (**Figure 22B**). Clear electron density is observed for PML residues 19-94 of SUMO1 as well as for residues 551-562, which make the interface with SUMO1 (**Figure 22A**) and residues 569-573 that participate in crystal contacts (**Figure 22C**). In the complex with the K39Q variant, residues K37, H43 and K46 are found at virtually the same positions as in the wild-type complex (**Figure 23**). Interactions between side chains of H43 and K46 with the phospho-serines 562 and 560 are preserved in terms of both their geometry and distances. However, Q39 is displaced 0.5 Å further away from pS561 when compared to K39 in the wild-type structure. As was observed in the K46Ac variant structure, K45 adopts an alternative conformation in the K39Q structure, which places it 4.0 Å from pS562 (**Table V**). Briefly, the K39Q structure displays less dramatic re-organization as the K46Ac structure, but nonetheless confirms the loss of interaction between residue 39 of SUMO1 and pS561 of PML-SIM as well as the added flexibility in the basic loop following the disruption of key electrostatic interactions.

Structures of the Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ and the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ complex

Since a more modest effect was observed with the K37Ac and K45Ac variants, attempts were made to structurally characterize these SUMO1 variants in order to explain these significant, but subtler changes in binding affinity. In both cases, crystals obtained with

the Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ and the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ complexes belong to the P2₁2₁2₁ space group and diffract to a resolution of 1.63 Å and 1.76 Å, respectively (**Table IV**). In addition, they both contain two copies of the corresponding complex in the asymmetric unit (**Figures 24 and 26, Figures S2 and S4**), where the peptide is either found close to the interface between two SUMO proteins or more exposed to the solvent (**Figures S1 and S3**). Therefore, for each structure, only the complex where the peptide is exposed to the solvent will be discussed in order to minimize interpretation of artefacts that may be the result of crystal packing.

In the Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ complex (**Figure 24D**), clear electron density is observed for residues 19-94 of SUMO1 and residues 553-562 of PML (**Figure 24C**). Overall, the position of K37 is not significantly altered in the acetylated form in comparison to the unacetylated form, which is consistent with the fact that it primarily forms hydrophobic contacts with V558 from the SIM of PML through the aliphatic portion of its side chain (**Figure 25**). Likewise, the positions of K45 and K46 are also unchanged when comparing the complexes of the acetylated and unacetylated form of K37. However, there are notable changes in the locations of the K39 and H43 side chains. An increase of 1.6 Å is observed for the distance between the side chain of H43 and the phosphate group of pS562 in the K37Ac complex. Also, the positive charge of K39 is 1 Å farther away from the negative charge of pS561 (**Table VI**).

In the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ complex (**Figure 26B**), clear electron density is observed for residues 19-94 of SUMO1 and residues 552-562 of PML (**Figure 26A**). Although the global structure of the complex with the K45Ac variant is very similar to the complex with the unacetylated form, an interesting phenomenon is observed when comparing the distances between the three phospho-serine residues and the side chains of K39, H43 and K46. Even though all distances are different from those of the wild-type structure, it seems that a stronger interaction with K39 may be compensating for the weakening of interactions with K46 and H43. In the K45Ac structure, K46 is now 0.5 Å further away from pS560 and H43 is 0.4 Å farther from pS562, but K39 is 1.9 Å closer to pS561 of PML (**Figure 27**). Therefore, the modest loss of binding affinity measured by ITC for both the K37Ac and

K45Ac variants may be the result of a more global effect of charge, rather than the disruption of a specific interaction between SUMO1 and PML.

Structure of the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complex

Given the dramatic structural changes observed in the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complex, we attempted to crystallize the K46Ac variant in complex with the phosphorylated SIM of Daxx. The rationale behind this comes from previous findings showing that the phospho-SIMs of Daxx and PML both interact with the side chain of K46 of SUMO, but in a slightly different manner. For these studies, the first four residues (DPEE) on the N-terminal side of the hydrophobic SIM core of the Daxx peptide (residues 729-732 of Daxx-SIM) were replaced with the four corresponding residues (AEER) of the PML peptide (residues 552-555 of PML-SIM) to generate a modified form of Daxx (mDaxx-SIM-2PO₄), which more readily crystallizes as previously described (85). Crystals of the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complex belong to the C2 space group and diffract to a resolution of 1.51 Å (**Table VII**). The asymmetric unit contains two copies of the complex (**Figure S5**). In both cases, clear electron density is observed for residues 20-94 of SUMO1 (**Figure 28B, D**) and residues 733-738 of Daxx (**Figure 28A, C**). For structural analysis, the first subunit was selected and this complex was compared to the corresponding complex in the Δ N-SUMO1:Daxx-SIM-2PO₄ structure (**Figure 29**). In complex with wild-type Δ N-SUMO1, pS739 and pS737 of Daxx-SIM interact with H43 and K46 of SUMO1, respectively, in a manner analogous to pS562 and pS560 of PML-SIM-4PO₄ (85). In complex with the K46Ac variant, the distance between the side chains of K46Ac and pS737 increases by 3.4 Å (**Table VIII**). However, the displaced phosphate group from pS737 is now located only 2.5 Å from K39 and is thus in position to form an alternate interaction with this positively charged residue. The movement of pS737 is also accompanied by a re-organization of the neighbouring residues of the SIM. In addition, there is no electron density for the side chain of pS739 suggesting that it is no longer able to interact with H43. However, this seems to be partially compensated for by a change in the location of the side chain of D738, which is now positioned only 4.1 Å from H43 (**Figure S6**). As observed for the phospho-SIM of PML with the K46Ac variant, there are dramatic differences in the recognition pattern at the binding interface. However, the details of the changes at the binding interface are significantly

different. As was seen with several acetylated variants (K37, K39 and K45) binding to the phospho-SIM of PML, changes in the overall global charge density can induce dramatic changes in the binding affinity and recognition pattern.

Discussion

SUMO proteins can interact non-covalently with SIMs either in their conjugated or unconjugated form and it is well established that phosphorylation at key sites within SIMs increases their affinity for SUMO proteins (61, 91). Structural characterization of the interaction between SUMO1 and the phosphorylated SIMs of PML and Daxx showed that they form distinct recognition patterns, but highlighted the fact that several common lysine residues in SUMO1 play key roles at the interface of both complexes (85). In addition, it has been recently demonstrated that several of these same lysine residues are modified by acetylation (63, 89, 90), but little is known about how acetylation of these lysine residues alters SUMO proteins binding to phospho-SIMs at the atomic level. In this work, we have characterized the role that acetylation plays in the binding of the phospho-SIM of PML to SUMO1 at the atomic level using a combination of structural and biophysical studies. We established, using ITC studies, that SUMO1 binding to the phospho-SIM of PML is significantly reduced by acetylation at lysine residues 37, 39, 45, and 46 with acetylation at K39 or K46 having a more dramatic effect than acetylation at K37 or K45. Consistent with the ITC results, structural characterization by X-ray crystallography permitted us to elucidate the mechanisms by which electrostatic interactions are disrupted and/or re-organized in each complex. In addition, we confirmed that there are dramatic changes in interactions at the binding interface with the SUMO1 K46Ac variant, but these changes are distinct when comparing the phospho-SIMs of PML and Daxx in complex with K46Ac.

It is clear from both the ITC and structural studies that acetylation of K46 has a dramatic effect on the recognition of the phospho-SIMs of both PML and Daxx. A common feature between the structures of the both complexes is the additional flexibility given to residue K46 following its acetylation. The additional degrees of freedom appear to be transmitted to K45, which adopts two conformations in the complexes with the K46Ac variant, whereas K45 is found at a single, well-defined position in other structures with SUMO1 (wild-type and other acetylated variants). In the second confirmation, there is now an electrostatic interaction between K45 and pS562, which appears to partially compensate for the weakening of other interactions (K46-pS560 and K39-pS561). However, we did not observe the alternative confirmation of K45 in the structure of the K46Ac in complex with the

phospho-SIM of Daxx. This may be due to the fact that the position of the first phospho-serine residue of the SIM relative to the side chain of K46 of SUMO is different in the complexes of unacetylated SUMO1 with the two SIMs. As for the position of the other basic residues of SUMO1, namely H43, K39 and K37, only minor changes are observed in their positions. In contrast, we observe slight repositioning of the phospho-serine residues in the PML-SIM, likely due to the disruption of the interaction between K46 and pS560. In the Daxx structure with the K46Ac variant, the SUMO1 residues are more or less positioned as in the unacetylated form, but alternative interactions are generated. For example, D738 is re-oriented towards H43, whereas pS737 is positioned in close proximity to K39 and more distant from its original location near K46. In addition, no electron density is observed for pS739 in the Daxx structure, indicating that acetylation of K46Ac somehow prevents the interaction between pS739 and H43. Taken together, our data suggests that acetylation SUMO1 at K46 has different effects depending on which SIM it is binding to and that the nature of the interaction depends on the combination of phosphorylated and/or acidic residues located in the first three positions following the hydrophobic core of the SIM.

Although the ITC experiments suggested that K39 acetylation has a very dramatic effect on binding affinity, we did not retrieve such a dramatic change in electrostatic interactions in the K39Q complex with the phospho-SIM of PML. Of note, we were unable to obtain well diffracting crystals with the K39Ac form, which may be due to the positioning of K39 in the crystal lattice. In the wild-type structure, K39 interacts with the C-terminus of PML-SIM of the adjacent asymmetric unit and this would not be possible without the positive charge (85). Interestingly, the one common feature between the K39Q and K46Ac variants in complex with the phospho-SIM of PML is the presence of an alternative conformation for K45. However, the other changes observed at the interface between the SIM of PML and the K39Q variant of SUMO1 are more similar to the changes observed with the K37Ac and K45Ac structures. These changes are best described as more global rearrangements of electrostatic interactions. In these complexes, interactions can be weakened or strengthened, while others are left unchanged. Nonetheless, the net result measured for these two variants is a decrease in binding affinity, which highlights the inherent plasticity in the SUMO/SIM interactions.

The plasticity observed in the structures of these complexes allows for an ideal mechanism for regulating SUMO/SIM interactions using a combination of phosphorylation and acetylation. A key factor resides in the fact that SIMs are located in intrinsically disordered regions of proteins. Thus, they have a natural built-in flexibility in addition to being in regions where extensive post-translational modifications typically occur (144, 145). This argument also applies to the SUMO proteins, since the four lysine residues at the binding interface targeted by acetylation are located in a flexible loop. Based on the structural data, it seems that both the position of an electrostatic interaction and its strength will determine the impact that the acetylation-induced disruption will have on the global SUMO/SIM interaction. Acetylation at position K46 and K39 prevents these residues from interacting with the phospho-SIM through electrostatic interactions (direct disruption), whereas acetylation of either K37 or K45 does not affect a distinct interaction, but rather alters the global electrostatic environment and disrupts interactions involving other positively charged residues. It is interesting to note that even though we identified residues K39 and K46 as having the most impact on binding affinity by ITC, it is the K37/K39 combination which is the most predominant *in vivo* (89). K37 acetylation could indeed have a cellular impact if it influences the global positive charge density in the basic loop of SUMO. Importantly, *in vivo* studies have shown that K37 (SUMO1)/K33 (SUMO2) acetylation affects the dynamics of PML-NBs (63) and this suggests that this global electrostatic environment plays a key role in regulating cellular SUMO/SIM interactions. Given the dramatic effect that the K46Q and K39Q variants have in the ITC experiments, it will be interesting to test these variants in similar *in vivo* studies to determine if an even more dramatic disruption of function is observed with these acetyl-mimetics.

Experimental procedures

Recombinant peptides and proteins constructs

The sequences encoding for the human PML-SIM peptides (residues 547-573; PML Isoform I) were ordered as oligonucleotides (Integrated DNA Technologies) with flanking BamHI and EcoRI restriction enzyme sites, 5'-phosphorylated, annealed and cloned as BamHI-EcoRI fragment into a modified pGEX-2T vector (Amersham) with a Tobacco Etch Virus (TEV) protease cut site replacing the original thrombin cut site. Δ N-SUMO1 protein (residues 17-97 of human SUMO1) constructs were generated using site-directed mutagenesis. For the acetylation mutants, the codon of the specific lysine residue (positions 37, 39, 45 and 46) was replaced by an *amber* codon (TAG) and cloned into pCDF PyIT vector, along with a pBK-AcKRS3 vector (gifts from Dr. Jason Chin)(146). For the acetyl-mimetic mutants, the codon of the specific lysine residue (positions 37, 39, 45 and 46) was replaced by a glutamine codon and cloned into pHTDuet vector (Novagen). Constructs were verified by DNA sequencing.

Peptides and proteins expression and purification

PML-SIM, PML-SIM-4SD and PML-SIM-4SE peptides were expressed as GST fusion in *E. coli* host strain *TOPP2* (Stratagene). Cells were grown in Luria Broth (LB) medium supplemented with 100 μ g/ml ampicillin overnight at 37°C. The next day, cells were diluted in LB plus antibiotics and grown until reaching O.D.₆₀₀ between 0.6-0.8. Protein expression was induced with 1 mM IsoPropyl- β -D-ThioGalactopyranoside (IPTG; Inalco) for 4h at 30°C. Cells were pelleted by centrifugation (10 min, 6000 rpm) and frozen at -20°C. Pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 7.4, 1M NaCl, 0.2 mM EDTA, 1 mM DTT) and passed twice through a French Press followed by a brief sonication. The resulting suspension was centrifuged (1h, 35,000 rpm, 4°C) and the supernatant incubated (1h, 4°C) with a Glutathione Sepharose 4B (GSH; GE Healthcare) resin. Following the incubation, the resin was centrifuged and washed several times with lysis buffer to remove non-specifically bound molecules. The resin was then rinsed three times with TEV buffer (10 mM NaH₂PO₄/Na₂HPO₄ pH 7.4, 125 mM NaCl, 5 mM DTT) and incubated overnight at room

temperature with 100 units of TEV protease. The next day, the supernatant was collected, filtered and dialyzed extensively in water. The resulting solution was then dialyzed in 5% acetic acid overnight prior to purification on a C₄-reverse phase HPLC column (Vydac) using a 0.05% trifluoroacetic acid (TFA) water-acetonitrile gradient. The organic solvent was removed from the collected fractions by roto-evaporation. The resulting solution was dialyzed in 20 mM Tris-HCl pH 7.5, then in water. The final purified solution was lyophilized and stored at -20°C in solid form.

ΔN-SUMO1 acetylated (K37Ac, K39Ac, K45Ac and K46Ac) proteins were expressed as His-tag fusion proteins in *E. coli* host strain *BL21(DE3)*. Cells were grown in LB supplemented with 35 μg/ml streptomycin and 50 μg/ml kanamycin overnight at 37°C. The next day, cells were diluted in LB plus antibiotics and grown until O.D₆₀₀ between 0.6-0.8. Then, 10 mM *N*-acetyl-lysine and 20 mM nicotinamide (NAM) were added to the bacterial suspension and cells were incubated 30 min at 37 °C, after which protein expression was induced for 3h at 37°C with 1 mM IPTG. The cells were harvested and lysed in a similar fashion as described above, with the exception of the lysis buffer composition (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 50 mM imidazole). After the French Press, the suspension was centrifuged 1h at 35,000 rpm (4°C) and the supernatant was incubated 1h (4°C) with nickel-charged Chelating Sepharose (GE Healthcare) resin. The bound resin was first washed several times with lysis buffer to remove non-specifically bound molecules and the acetylated SUMO1 proteins were eluted using a buffer consisting of 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 500 mM imidazole. The eluted protein solution was dialyzed into water and then into TEV buffer. The resulting protein solution was incubated overnight at room temperature with 100 units of TEV protease. After the TEV cleavage reaction, the protein solution was centrifuged at 3,500 rpm for 3 min and passed through a 0.45 μM filter. The filtered solution was dialyzed in water and then in lysis buffer prior to being reloaded on nickel-charged Chelating Sepharose resin to remove the cleaved His-tag. The purified proteins were finally dialyzed in water overnight, lyophilized and stored at -20°C in solid form. For the crystallography experiments, the acetylated proteins were additionally purified by HPLC as described above.

The Δ N-SUMO1 K-to-Q mutants (K37Q, K39Q, K45Q and K46Q) and wild-type Δ N-SUMO1 proteins were expressed as His-tag fusion proteins in *E. coli* host strain *BL21(DE3)*. Cells were grown in LB supplemented with 35 μ g/ml streptomycin overnight at 37°C. The next day, cells were diluted in LB plus antibiotics and grown until reaching O.D₆₀₀ between 0.6-0.8. Protein expression was induced for 3h at 37°C with 1 mM IPTG. The purification procedure is identical to the one described above for the Δ N-SUMO1 acetylated proteins.

***In vitro* phosphorylation of SIM peptides**

For preparation of the PML-SIM-4PO₄ peptide, 200-250 η mol of the PML-SIM peptide were resuspended in 1mL of 1X NEBuffer for Protein Kinase (50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 0.1 mM EDTA, 2 mM DTT, 0.01% Brij 35). 3,000 units of CKII enzyme (NEB) were added to the solution, which was then incubated for 72h at room temperature. Every 24h, the reaction was supplemented with fresh enzyme (500 units, 1 μ L) and fresh ATP (2 μ mol) twice daily for the entire reaction period. After 72h, acetic acid was added to a final concentration of 5% (v/v) to quench the reaction. The solution was then centrifuged 2 min at 13,000 rpm, filtered and purified by HPLC as described above.

ITC experiments

Lyophilized proteins and peptides were resuspended in ultrapure milliQ water and dialysed overnight at room temperature into 20 mM Tris-HCl pH 7.4. Protein and peptide concentrations were determined by UV absorbance at 280 nm. ITC measurements were performed at 25°C using a VP-ITC calorimeter (MicroCal). The concentration of the PML-SIM peptides in the syringe over the Δ N-SUMO1 proteins concentration in the sample cell were maintained at a molar ratio of 10:1. Data were analyzed using MicroCal Origin Software and all experiments fit the single binding site model with 1:1 stoichiometry. Errors in K_D values were estimated from duplicate measurements or more.

Crystallization and data collection

Lyophilized Δ N-SUMO1 proteins were resuspended in ultrapure milliQ water at a final concentration of 500 μ M-1 mM. The PML-SIM or PML-SIM-PO₄ peptides were added to a

final concentration of 600 μ M-1.2 mM, to yield a protein:peptide ratio of 1:1.2. Crystals of each complex were obtained by the hanging drop vapour diffusion method using a precipitant solution containing 100 mM sodium cacodylate pH 6.5, 22 to 34% (*w/v*) PEG3350 and 10 mM calcium chloride. Crystals were cryoprotected in a solution containing 10% glycerol. Diffraction data were collected using either a Rayonix MX300 or a Pilatus3S 6M detector at beamline 08-ID of the Canadian Light Source (CLS) or a Pilatus3S 6M detector at the beamline F1 of the Macromolecular Cornell High Energy Synchrotron Source (MacCHESS). Datasets were indexed, integrated and scaled using HKL2000 (HKL Research, Inc.).

Structure determination and refinement

Initial phases were obtained by molecular replacement using the crystal structure of Δ N-SUMO1 in complex with PML-SIM (Protein Data Bank [PDB] ID [4WJO](#)), PML-SIM-4PO₄ (Protein Data Bank PDB ID [4WJN](#)), or Daxx-SIM-2PO₄ (Protein Data Bank PDB ID [4WJP](#)) as search template. Phases were improved by iterative cycles of model building with Coot (*132, 147*) and refinement with PHENIX (*148*). Test data sets were randomly selected from the observed reflections prior to refinement. The figures were prepared with PyMOL.

Authors contribution

C. G., M. L.-P., X. H. M. and H. W. prepared reagents and conducted experiments. C. G. and J. G. O. designed experiments and wrote the manuscript.

Acknowledgements

We would like to thank the laboratory of Jason Chin for providing the system for production of acetylated proteins, as well as Laurent Cappadocia for helpful insights and discussions on the project.

Funding

This work was supported by grants from the Canadian Institutes for Health Research to J.G.O (#74739 and 130414). Research described in this work was carried out in part at the Macromolecular Diffraction Facility at Cornell High Energy Synchrotron Source (MacCHESS), which is supported by the NSF under grant DMR-1332208 and NIH/NIGMS under grant GM-103485 and at the Canadian Light Source (CLS) which is supported by the Natural Sciences and Engineering Research Council of Canada, the National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada, and the University of Saskatchewan.

Conflict of interest

The authors of this manuscript declare no conflict of interest.

Tables

Table I. ITC results for Δ N-SUMO1 K-to-Q variants binding to either PML-SIM-4SD or PML-SIM-4SE.

Δ N-SUMO1	PML-SIM-4SD		PML-SIM-4SE	
	K_D (μ M)	Fold Δ^*	K_D (μ M)	Fold Δ^*
WT	0.4	1	2.1	1
K37Q	2.7 \pm 0.2	7	13.5 \pm 0.6	6
K39Q	6.7 \pm 0.2	17	20.4 \pm 1.5	10
K45Q	2.5 \pm 0.3	6	9.9 \pm 0.4	5
K46Q	12.8 \pm 0.7	32	44.0 \pm 6.8	21

*Relative change in binding affinity compared to WT Δ N-SUMO1. The errors listed represent the values around mean.

Table II. ITC results for Δ N-SUMO1 K-to-Ac variants binding to either PML-SIM-4SD or PML-SIM-4SE.

Δ N-SUMO1	PML-SIM-4SD		PML-SIM-4SE	
	K_D (μ M)	Fold Δ^*	K_D (μ M)	Fold Δ^*
WT	0.4	1	2.1	1
K37Ac	1.8 \pm 0.1	4	7.6 \pm 0.3	4
K39Ac	11.2 \pm 0.7	28	34.8 \pm 6.9	17
K45Ac	2.6 \pm 0.2	6	10.1 \pm 0.1	5
K46Ac	10.6 \pm 2.4	26	34.7 \pm 1.7	17

*Relative change in binding affinity compared to WT Δ N-SUMO1. The errors listed represent the values around mean.

Table III. Crystal data collection and refinement statistics for acetylated SUMO variants in complex with PML-SIM.

Dataset	Δ N-SUMO1- K37Ac:PML-SIM	Δ N-SUMO1- K39Ac:PML-SIM	Δ N-SUMO1- K45Ac:PML-SIM	Δ N-SUMO1- K46Ac:PML-SIM
Data collection				
Beamline	08-ID, CLS	08-ID, CLS	08-ID, CLS	08-ID, CLS
Wavelength (Å)	0.98	0.98	0.98	0.98
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell parameters	a = 38.3, b = 38.6, c = 142.9 $\alpha = \beta = \gamma = 90$	a = 38.3, b = 47.4, c = 62.2 $\alpha = \beta = \gamma = 90$	a = 38.2, b = 47.1, c = 63.5 $\alpha = \beta = \gamma = 90$	a = 38.4, b = 47.1, c = 63.2 $\alpha = \beta = \gamma = 90$
Resolution range (Å)	36.97 – 1.71 (1.761 – 1.71)	32.64 – 1.45 (1.5 – 1.45)	31.76 – 1.54 (1.60 – 1.54)	29.74 – 1.35 (1.40 – 1.35)
No. of unique reflections	23837 (2085)	20636 (1949)	17300 (1656)	25729 (2496)
Multiplicity	4.6 (3.8)	5.8 (4.7)	7.1 (6.0)	6.2 (5.2)
Completeness (%)	98.67 (88.16)	99.21 (96.34)	98.72 (95.78)	99.76 (98.62)
R _{merge}	0.061 (0.895)	0.052 (0.979)	0.075 (0.876)	0.057 (0.957)
CC1/2	0.999 (0.626)	0.999 (0.531)	0.999 (0.651)	0.999 (0.578)
I/ σ (I)	14.17 (1.08)	20.47 (1.87)	17.43 (1.75)	18.76 (1.72)
Refinement statistics				
Resolution (Å)	36.97 – 1.71	32.64 – 1.45	31.76 – 1.54	29.74 – 1.35
Reflections (total/test) ^a	23818/2003	20636/2000	17300/1731	25736/2000
R _{work} /R _{free} (%)	17.31/20.68	15.34/18.09	14.67/18.52	14.97/18.07
No. atoms (excluding hydrogens)				
Protein	1417	754	750	763
Water	150	113	139	143
B-factors				
Protein	37.59	23.67	16.28	17.11
Water	43.98	40.51	35.66	36.86
Rmsd				
Bond length (Å)	0.010	0.009	0.009	0.008
Bond angle (°)	1.25	1.22	1.17	1.21
Ramachandran ^b				
Favored (%)	100	100	100	100
Outliers (%)	0	0	0	0

Values in parentheses are for highest resolution shell. $R_{sym} = \sum hkl \sum i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum hkl \sum i I_{hkl,i}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices hkl and $\langle I_{hkl} \rangle$ is the mean intensity of the reflection. $R_{work} = \sum hkl ||F_o| - |F_c|| / \sum hkl |F_o|$, where $|F_o|$ is the observed structure-factor amplitude and $|F_c|$ is the calculated structure-factor amplitude. R_{free} is the R factor based on at least 500 test reflections that were excluded from the refinement.

^a Reflection for $F_o > 0$

^b MolProbity analysis

Table IV. Crystal data collection and refinement statistics for acetylated SUMO variants in complex with PML-SIM-4PO₄.

Dataset	Δ N-SUMO1-K37Ac:PML-SIM-4PO ₄	Δ N-SUMO1-K39Q:PML-SIM-4PO ₄	Δ N-SUMO1-K45Ac:PML-SIM-4PO ₄	Δ N-SUMO1-K46Ac:PML-SIM-4PO ₄
Data collection				
Beamline	F1, macCHESS	08ID-1, CLS	08ID-1, CLS	F1, macCHESS
Wavelength (Å)	0.98	0.98	0.98	0.98
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell parameters	a = 38.3, b = 38.5, c = 144.0 $\alpha = \beta = \gamma = 90$	a = 38.3, b = 47.2, c = 63.3 $\alpha = \beta = \gamma = 90$	a = 38.3, b = 62.7, c = 90.9 $\alpha = \beta = \gamma = 90$	a = 38.2, b = 46.9, c = 63.6 $\alpha = \beta = \gamma = 90$
Resolution range (Å)	30.04-1.63 (1.68-1.63)	32.75 – 1.84 (1.91 – 1.84)	45.45 – 1.76 (1.82 – 1.76)	37.76-1.45 (1.50-1.45)
No. of unique reflections	27527 (2640)	10279 (882)	22351 (2104)	20884 (2038)
Multiplicity	5.8 (5.3)	6.3 (6.1)	6.1 (5.0)	6.3 (5.5)
Completeness	99.33 (96.63)	97.60 (86.39)	99.32 (94.10)	99.85 (99.41)
R _{merge}	0.073 (0.830)	0.127 (0.871)	0.121 (0.870)	0.082 (0.882)
CC1/2	0.998 (0.750)	0.996 (0.757)	0.996 (0.572)	0.999 (0.656)
I/ σ (I)	12.94 (1.37)	11.71 (2.30)	11.58 (2.10)	9.47 (1.14)
Refinement statistics				
Resolution (Å)	30.04-1.63	32.75 – 1.84	45.45 – 1.76	37.76-1.45
Reflections (total/test) ^a	27520/2000	10161/1017	22351/1999	20883/2000
R _{work} /R _{free} (%)	17.65/20.09	15.44/18.49	16.68/20.27	16.26/18.57
No. atoms (excluding hydrogens)				
Protein	1444	790	1451	802
Water	194	98	189	112
B-factors				
Protein	31.24	24.65	27.88	20.65
Water	40.78	37.94	39.53	34.67
Rmsd				
Bond length (Å)	0.009	0.009	0.010	0.008
Bond angle (°)	1.06	0.99	1.07	1.02
Ramachandran ^b				
Favored (%)	100	100	99.35	100
Outliers (%)	0	0	0	0

Values in parentheses are for highest resolution shell. $R_{sym} = \sum hkl \sum i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum hkl \sum i I_{hkl,i}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices hkl and $\langle I_{hkl} \rangle$ is the mean intensity of the reflection. $R_{work} = \sum hkl ||F_o| - |F_c|| / \sum hkl |F_o|$, where $|F_o|$ is the observed structure-factor amplitude and $|F_c|$ is the calculated structure-factor amplitude. R_{free} is the R factor based on at least 500 test reflections that were excluded from the refinement.

^a Reflection for $F_o > 0$

^b MolProbity analysis

Table V. Summary of distances between phosphoserines (pS) of PML-SIM-4PO₄ and key residues of ΔN-SUMO1, ΔN-SUMO1-K46Ac and ΔN-SUMO1-K39Q.

Interaction	wt-4PO ₄	K46Ac-4PO ₄	K39Q-4PO ₄
K46-pS560	2.9*	5.5	3.3
K39-pS561	4.8	6.1	5.3
H43-pS562	2.8	2.9	2.6
K45-pS562	8.6	4.2	4.0

*Distances are given in Angstroms (Å) and are calculated between heteroatoms (nitrogen and oxygen) of involved residues. Values in *italic* are for the additional alternate conformer.

Table VI. Summary of distances between phosphoserines (pS) of PML-SIM-4PO₄ and key residues of ΔN-SUMO1, ΔN-SUMO1-K37Ac and ΔN-SUMO1-K45Ac.

Interaction	wt-4PO ₄	K37Ac-4PO ₄	K37Ac-4PO ₄	K45Ac-4PO ₄	K45Ac-4PO ₄
		A:B	C:D	A:B	C:D
K46-pS560	2.9*	3.8	2.8	2.4	3.3
K39-pS561	4.8	3.0	5.8	2.9	3.1
H43-pS562	2.8	5.4	4.4	2.4	3.4

*Distances are given in Angstroms (Å) and are calculated between heteroatoms (nitrogen and oxygen) of involved residues.

Table VII. Data collection and refinement statistics for Δ N-SUMO1-K46Ac in complex with mDaxx-SIM-2PO₄.

Dataset	Δ N-SUMO1-K46Ac:Daxx-SIM-2PO ₄
Data collection	
Beamline	08ID-1, CLS
Wavelength (Å)	0.98
Spa group	C2
Unit cell parameters	a = 100.2, b = 38.3, c = 67.7 $\alpha = \gamma = 90, \beta = 120$
Resolution range (Å)	48.28 – 1.51 (1.56 – 1.51)
No. of unique reflections	34518 (3027)
Multiplicity	3.2 (2.7)
Completeness	97.20 (85.64)
R _{merge}	0.0789 (0.5164)
CC1/2	0.996 (0.624)
I/ σ (I)	10.97 (2.14)
Refinement statistics	
Resolution (Å)	48.28 – 1.51
Reflections (total/test) ^a	34518/2000
R _{work} /R _{free} (%)	16.49/19.23
No. atoms (excluding hydrogens)	
Protein	1430
Water	282
B-factors	
Protein	19.31
Water	32.83
Rmsd	
Bond length (Å)	0.008
Bond angle (°)	1.08
Ramachandran ^b	
Favored (%)	100
Outliers (%)	0
<p>Values in parentheses are for highest resolution shell. $R_{sym} = \frac{\sum hkl \sum i I_{hkl,i} - \langle I_{hkl} \rangle }{\sum hkl \sum i I_{hkl,i}}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices hkl and I_{hkl} is the mean intensity of the reflection. $R_{work} = \frac{\sum hkl F_o - F_c }{\sum hkl F_o }$, where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude. R_{free} is the R factor based on at least 500 test reflections that were excluded from the refinement.</p> <p>^a Reflection for $F_o > 0$</p> <p>^b MolProbity analysis</p>	

Table VIII. Summary of distances between phosphoserines (pS) of Daxx-SIM and key residues of Δ N-SUMO1 and Δ N-SUMO1-K46Ac.

Interaction	wt-2PO₄		K46Ac-2PO₄	
	A:B	C:D	A:B	C:D
K46-pS737	3.5	4.2	6.9	8.9
K39-pS737	9.3	10.5	2.5	7.0
K39-D738	4.4	6.5	8.0	5.7
H43-pS739	-	2.5	-	-
H43-D738	7.3	6.4	4.1	7.3

*Distances are given in Angstroms (Å) and are calculated between heteroatoms (nitrogen and oxygen) of involved residues. Values are absent when peptide density was missing.

Figures

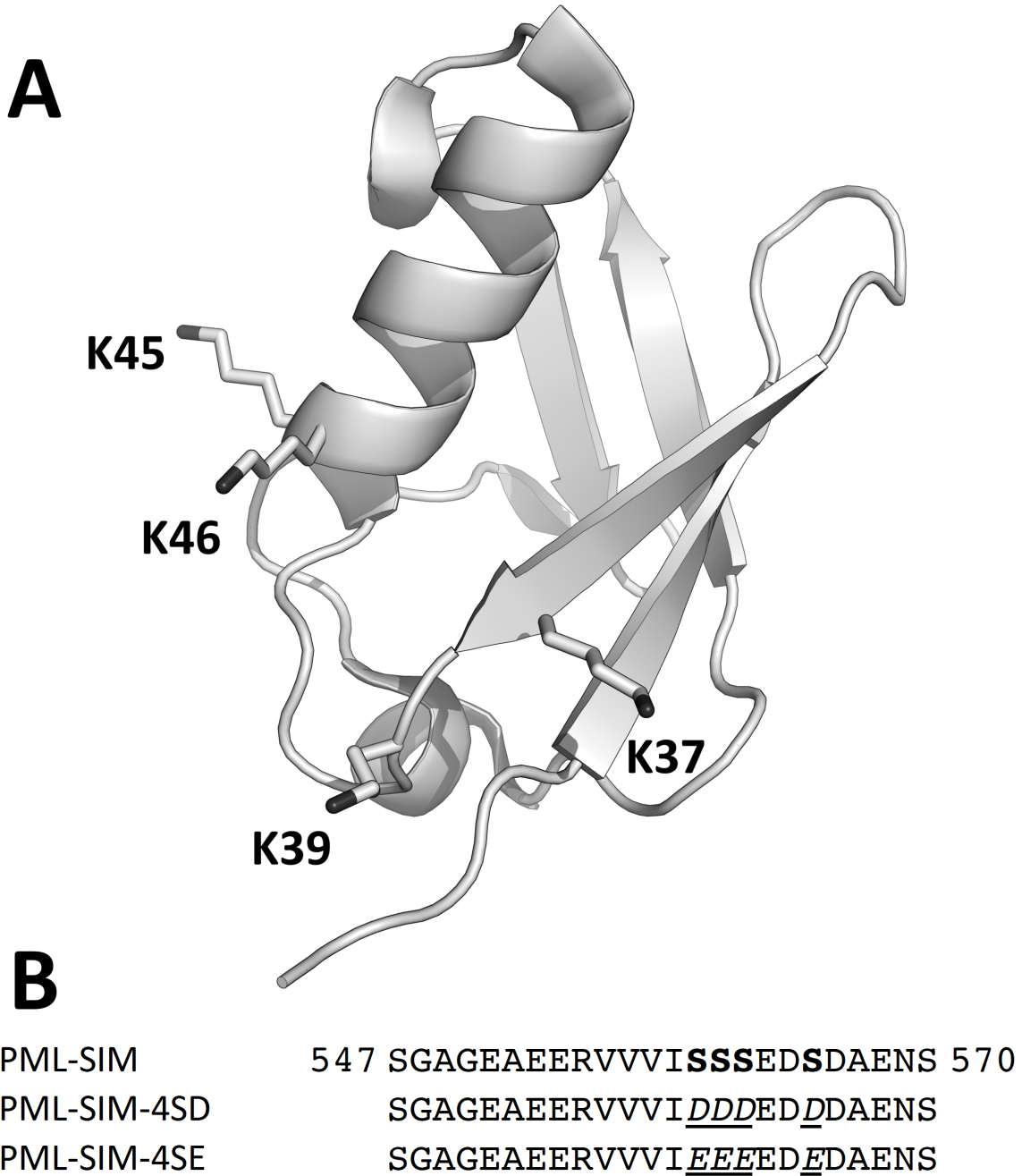


Figure 16. Representation of important post-translational modification sites on SUMO1 and PML-SIM. (A) Ribbon model of the 3D structure of Δ N-SUMO1 (light gray, PDB ID [4WJN](#)) with the four lysine residues (K37, K39, K45, K46) at the SIM binding interface, which undergo acetylation highlighted as sticks. (B) Amino acid sequences of wild-type PML-SIM and phospho-mimetic variants (PML-SIM-4SD, PML-SIM-4SE) used in the ITC studies. Serine residues that are phosphorylated by CK2 are in bold and negatively charged substitutions are in italics as well as underlined.

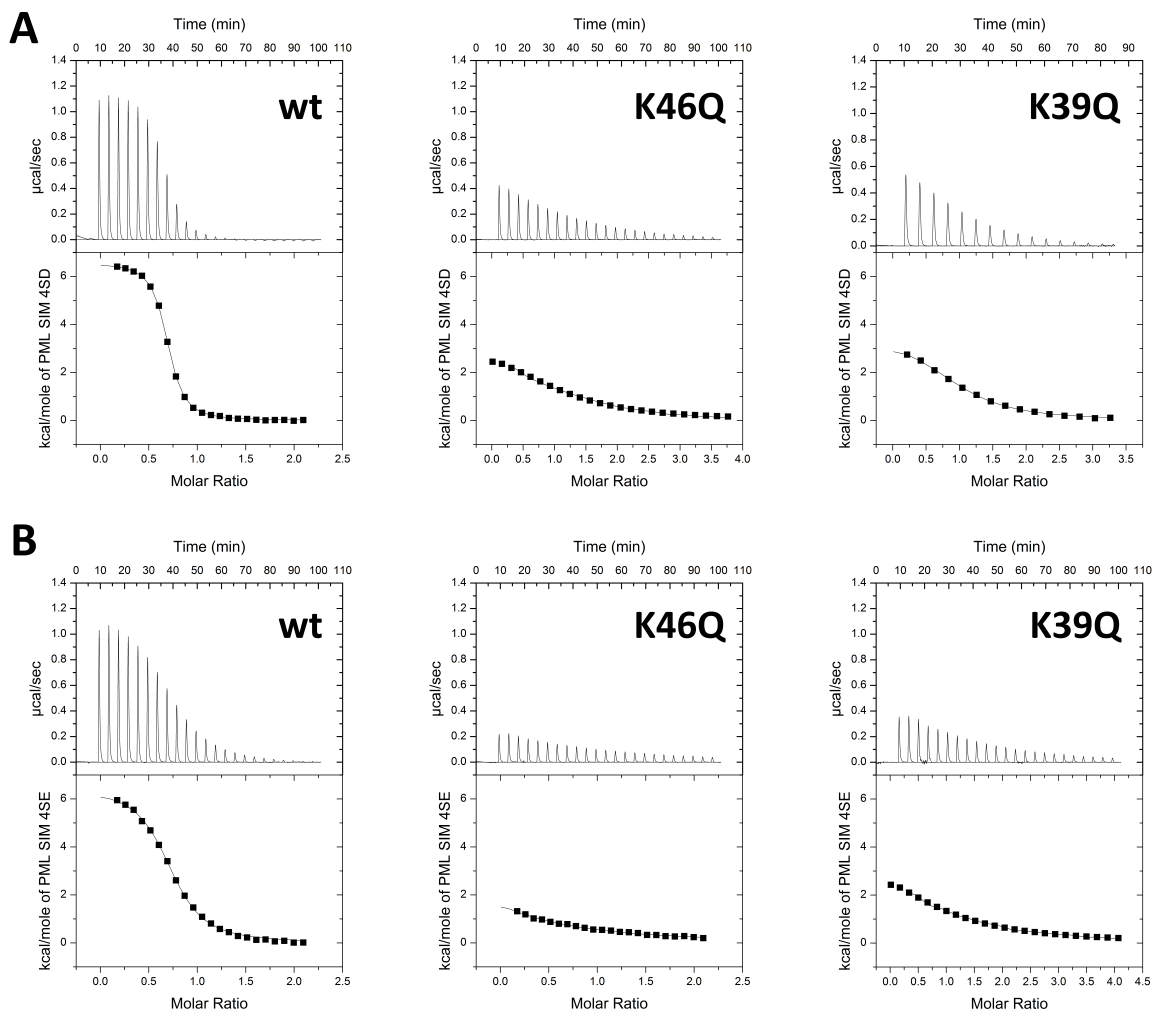


Figure 17. SUMO1 K-to-Q variants dramatically decrease binding to PML-SIM peptides. Representative ITC thermograms for the interaction between Δ N-SUMO1 acetyl-mimetic variants (K46Q, K39Q) and the PML-SIM-4SD (A panel) or PML-SIM-4SE (B panel) phospho-mimetic peptide.

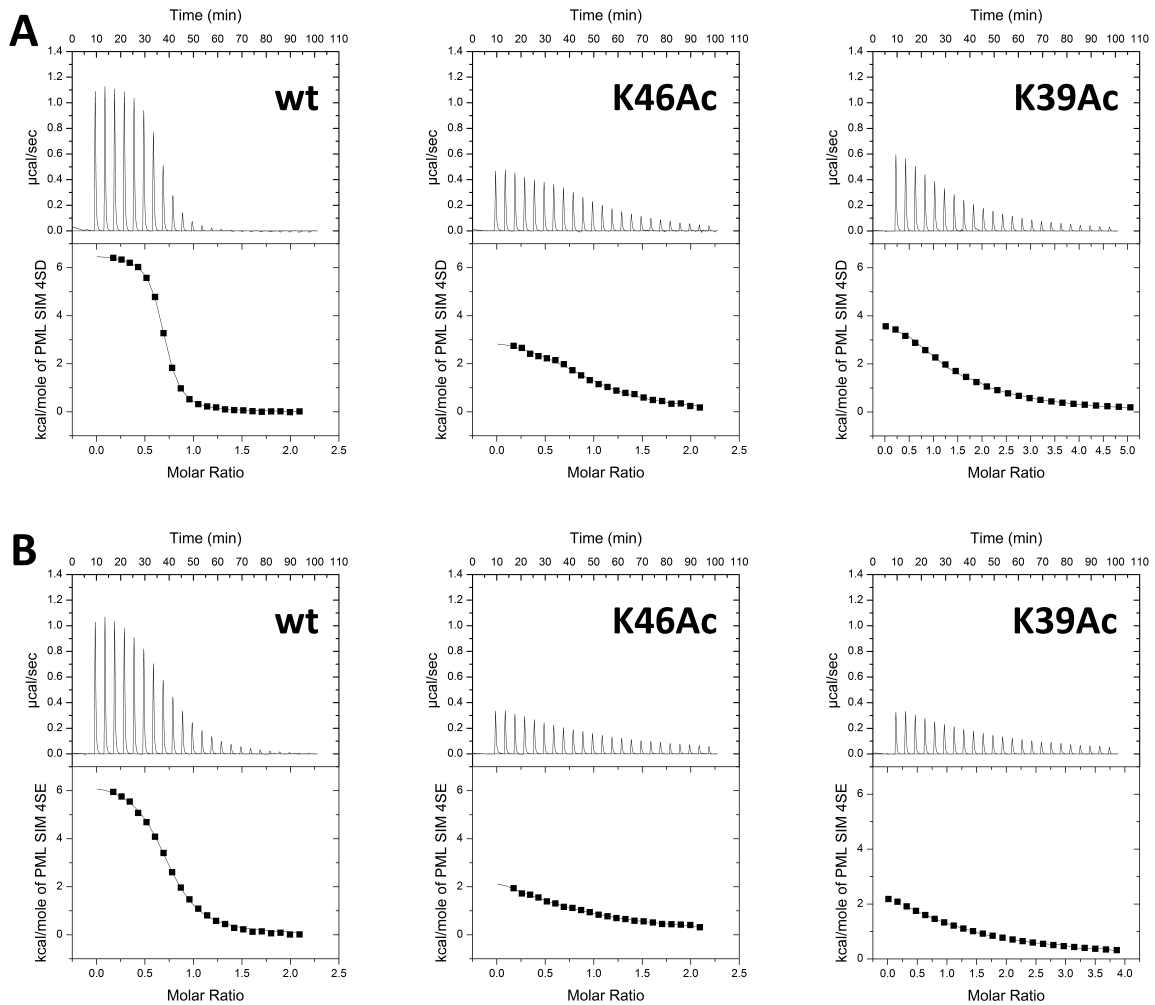


Figure 18. Acetylated SUMO1 variants bind to PML-SIM peptides in a similar manner as K-to-Q variants. Representative ITC thermograms for the interaction between the Δ N-SUMO1 acetylated variants (K46Ac, K39Ac) and the PML-SIM-4SD (A panel) or PML-SIM-4SE (B panel) phospho-mimetic peptide.

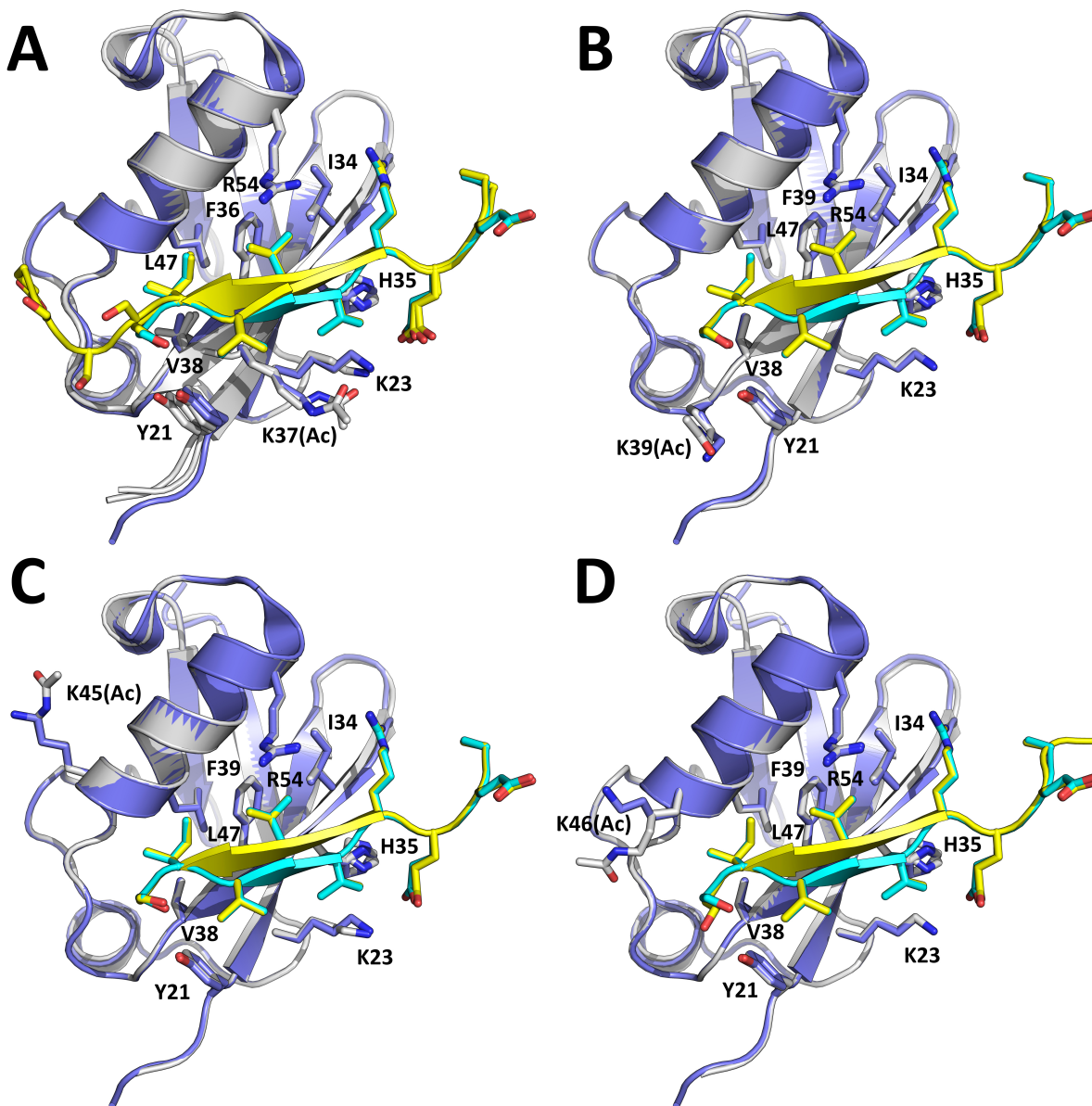


Figure 19. Acetylation of SUMO1 does not affect interactions with the hydrophobic core of the PML-SIM. Overlay of the Δ N-SUMO1:PML-SIM structure (PDB entry [4WJO](#)) with Δ N-SUMO1-K37Ac:PML-SIM (A), Δ N-SUMO1-K39Ac:PML-SIM (B), Δ N-SUMO1-K45Ac:PML-SIM (C), and Δ N-SUMO1-K46Ac:PML-SIM (D). Δ N-SUMO1 (carbon atoms slate) and the acetyl variants (carbon atoms grey) are shown in cartoon representation. The SIM of PML is shown in cartoon and sticks in either cyan (Δ N-SUMO1) or yellow (acetyl variants). Important residues at the binding interface are represented as sticks and select SUMO1 residues are labeled.

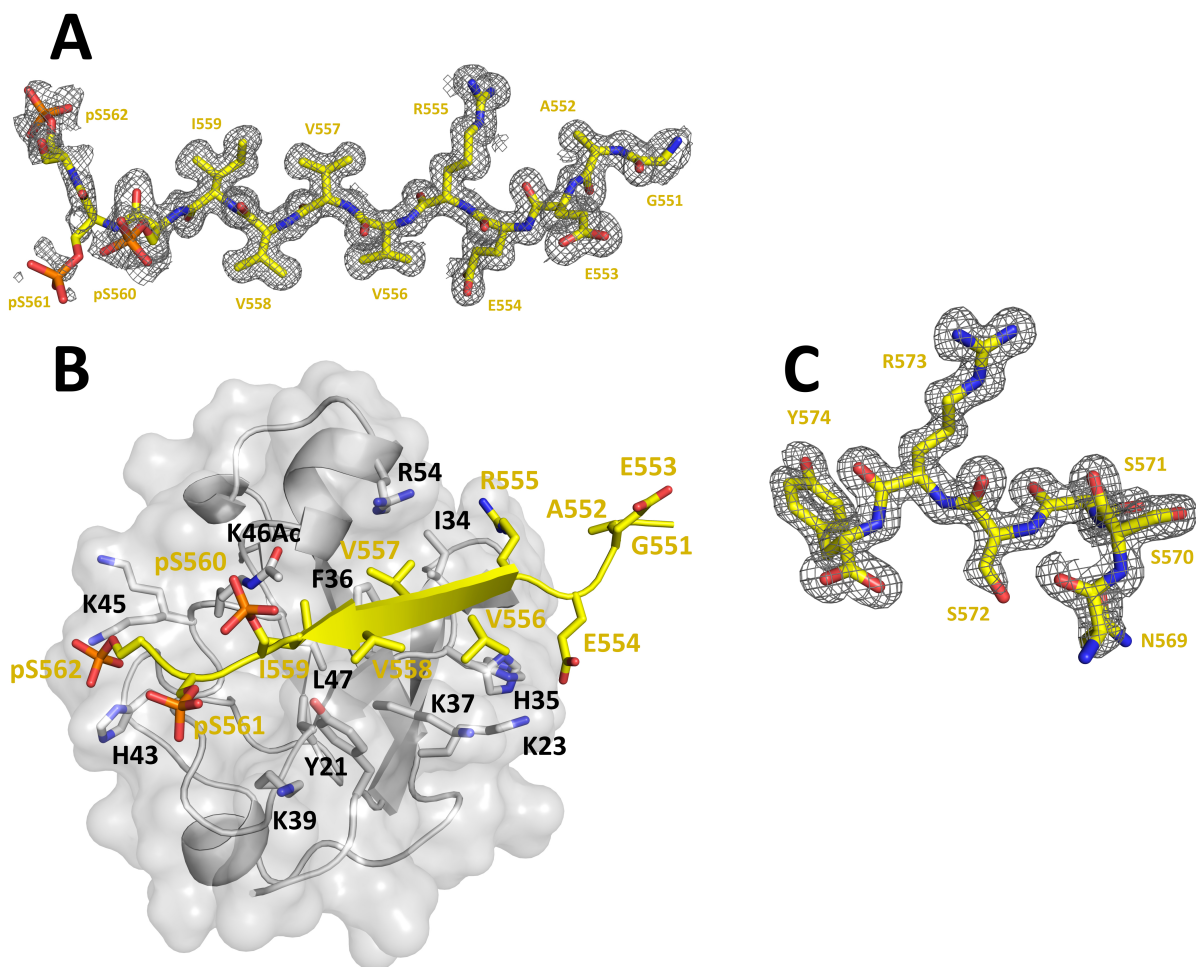


Figure 20. Crystal structure of the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complex. (A) Electron density map corresponding to residues 551-562 of PML-SIM-4PO₄. The peptide is in stick representation with carbon atoms colored yellow. The $2F_oF_c$ electron density map, contoured at 0.6σ , is represented by a gray mesh. (B) Overall structure of the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complex highlighting the interaction between SUMO1 and residues 551-562 of PML. The proteins are in cartoon representation with Δ N-SUMO1-K46Ac carbon atoms (gray) and PML-SIM-4PO₄ peptide carbon atoms (yellow); Δ N-SUMO1-K46Ac residues interacting with the PML-SIM-4PO₄ peptide are in stick representation. (C) Electron density map corresponding to residues 569-574 of the PML-SIM-4PO₄ peptide that are stabilized in between asymmetric units. The peptide is in stick representation with carbon atoms colored yellow. The $2F_oF_c$ electron density map, contoured at 1.2σ , is represented by a gray mesh.

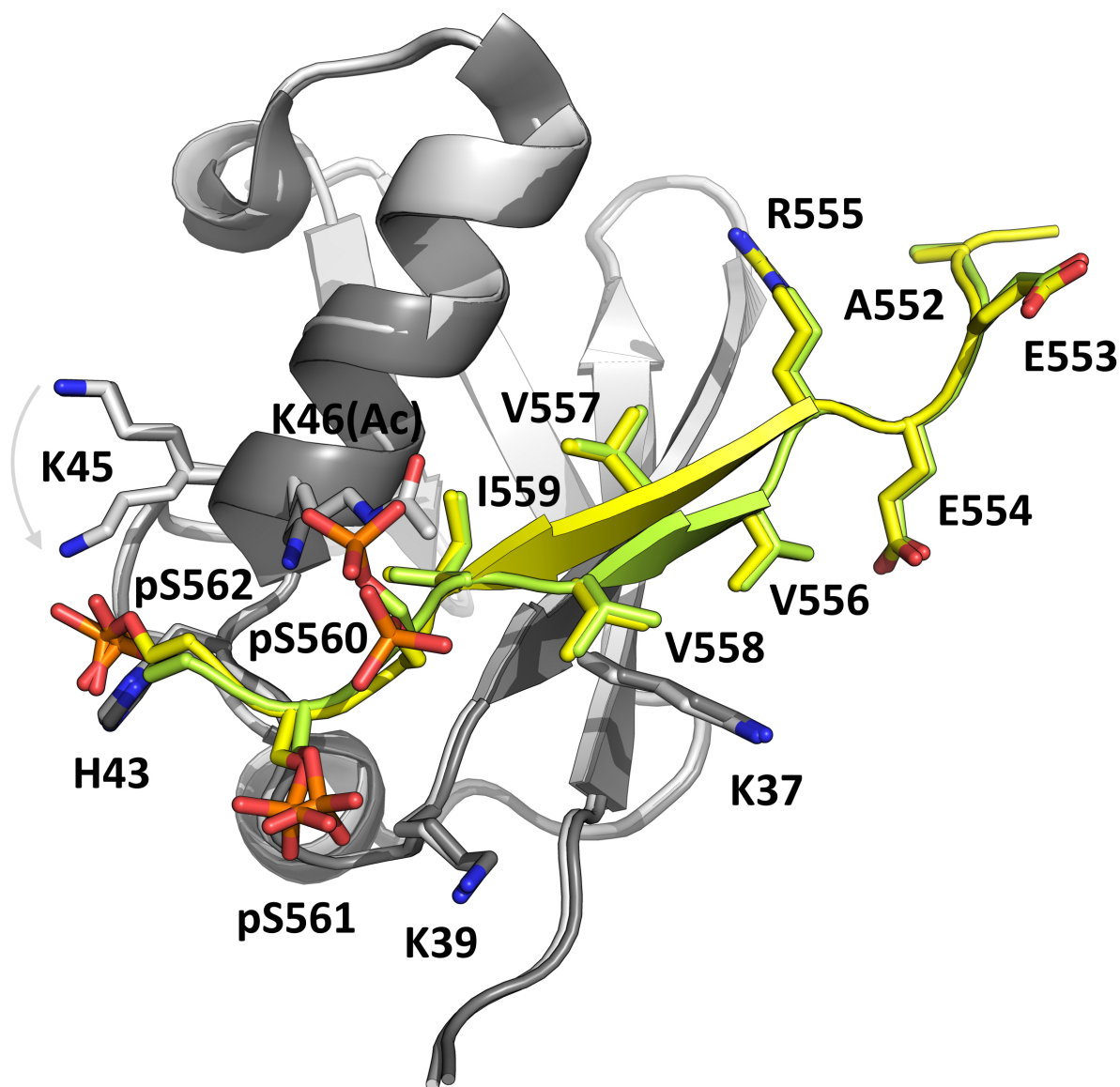


Figure 21. Comparison of Δ N-SUMO1:PML-SIM-4PO₄ (PDB entry [4WJN](#)) and Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complexes. Overlay of the Δ N-SUMO1:PML-SIM-4PO₄ (carbon atoms dark gray:lime) and the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ structures (carbon atoms light gray:yellow). Proteins are represented in cartoon; Δ N-SUMO1 basic residues and residues 551-562 of PML-SIM-4PO₄ are in stick representation. The light gray arrow highlights the difference in the positioning of the two conformations of K45 in the Δ N-SUMO1-K46Ac:PML-SIM-4PO₄ complex.

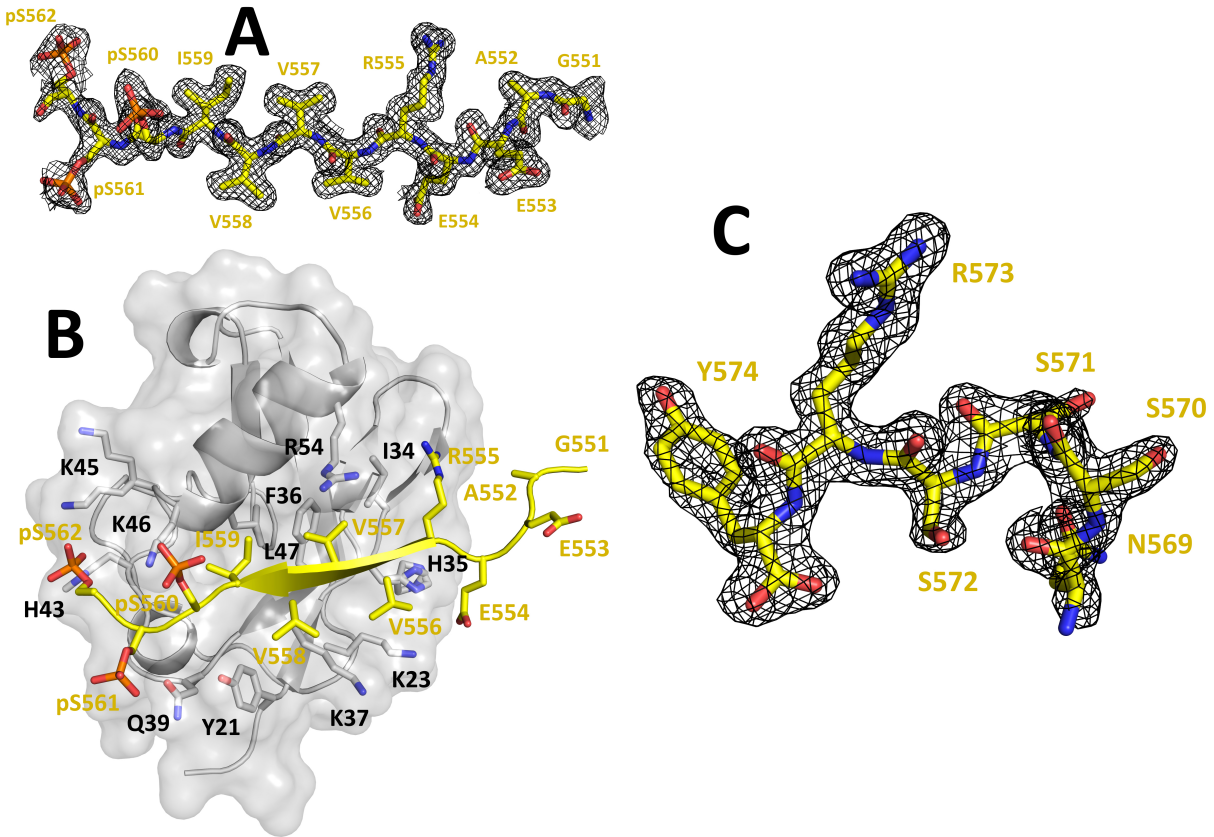


Figure 22. Crystal structure of the Δ N-SUMO1-K39Q:PML-SIM-4PO₄ complex. (A) Electron density map corresponding to residues 551-562 of PML-SIM-4PO₄. The peptide is in stick representation with carbon atoms colored yellow. The 2F_oF_c electron density map, contoured at 0.7 σ , is represented by a gray mesh. (B) Overall structure of the Δ N-SUMO1-K39Q:PML-SIM-4PO₄ complex highlighting the interaction between SUMO1 and residues 551-562 of PML. The proteins are in cartoon representation with Δ N-SUMO1-K39Q carbon atoms (gray) and PML-SIM-4PO₄ peptide carbon atoms (yellow); Δ N-SUMO1-K39Q residues interacting with the PML-SIM-4PO₄ peptide are in stick representation. (C) Electron density map corresponding to residues 569-574 of the PML-SIM-4PO₄ peptide that are stabilized in between asymmetric units. The peptide is in stick representation with carbon atoms colored in yellow. The 2F_oF_c electron density map, contoured at 1.0 σ , is represented by a gray mesh.

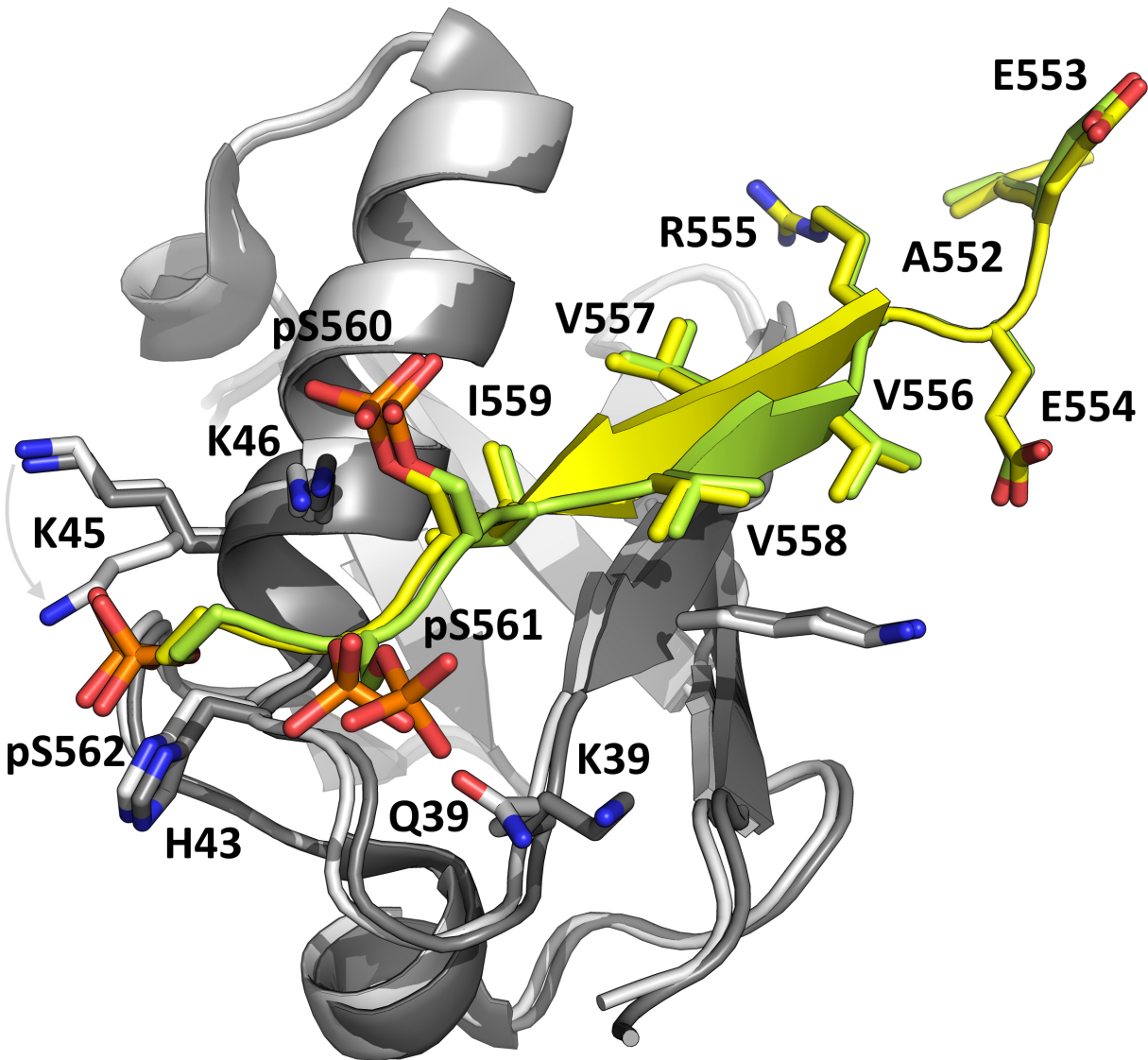


Figure 23. Comparison of Δ N-SUMO1:PML-SIM-4PO₄ (PDB entry [4WJN](#)) and Δ N-SUMO1-K39Q:PML-SIM-4PO₄ complexes. Overlay of the Δ N-SUMO1:PML-SIM-4PO₄ (carbon atoms dark gray:lime) and the Δ N-SUMO1-K39Q:PML-SIM-4PO₄ structures (carbon atoms light gray:yellow). Proteins are represented in cartoon; Δ N-SUMO1 basic residues and residues 551-562 of PML-SIM-4PO₄ are in stick representation. The light gray arrow highlights the difference in the positioning of the two conformations of K45 in the Δ N-SUMO1-K39Q:PML-SIM-4PO₄ complex.

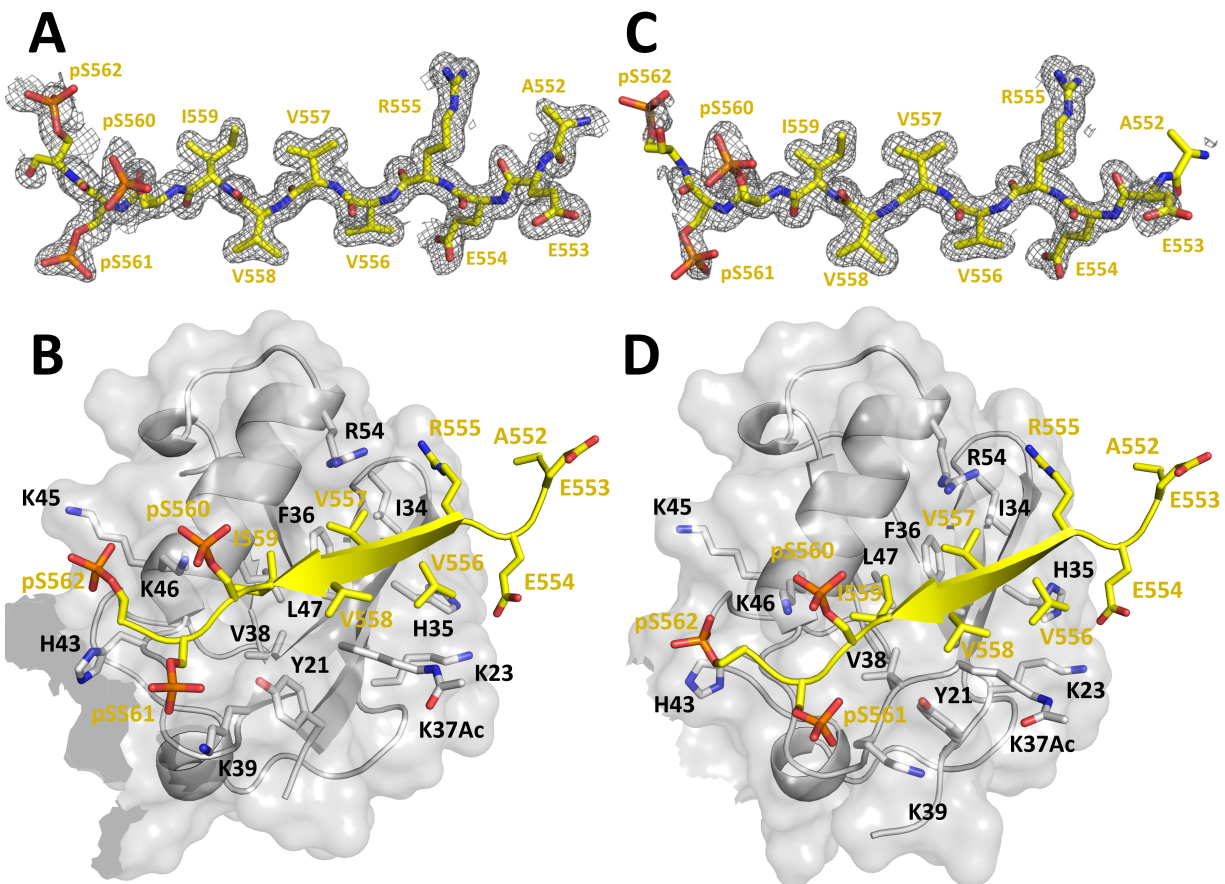


Figure 24. Crystal structure of Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ complexes. (A) Electron density map corresponding to residues 552-562 of PML-SIM-4PO₄ (chain B). The peptide is in stick representation with carbon atoms colored yellow. The $2F_oF_c$ electron density map, contoured at 0.8σ , is represented by a gray mesh. (B) Overall structure of the Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ (chains A:B) complex highlighting the interaction between SUMO1 and residues 552-562 of PML. The proteins are in cartoon representation with Δ N-SUMO1-K37Ac carbon atoms (gray, chain A) and PML-SIM-4PO₄ peptide carbon atoms (yellow, chain B); Δ N-SUMO1-K37Ac residues interacting with the PML-SIM-4PO₄ peptide are in stick representation. (C) Electron density map corresponding to residues 552-562 of PML-SIM-4PO₄ (chain D). The peptide is in stick representation with carbon atoms colored yellow. The $2F_oF_c$ electron density map, contoured at 0.8σ , is represented by a gray mesh. (D) Overall structure of the Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ (chains C:D) complex highlighting the interaction between SUMO1 and residues 552-562 of PML. The proteins are in cartoon representation with Δ N-SUMO1-K37Ac carbon atoms (gray, chain C) and PML-SIM-4PO₄ peptide carbon atoms (yellow, chain D); Δ N-SUMO1-K37Ac residues interacting with the PML-SIM-4PO₄ peptide are in stick representation.

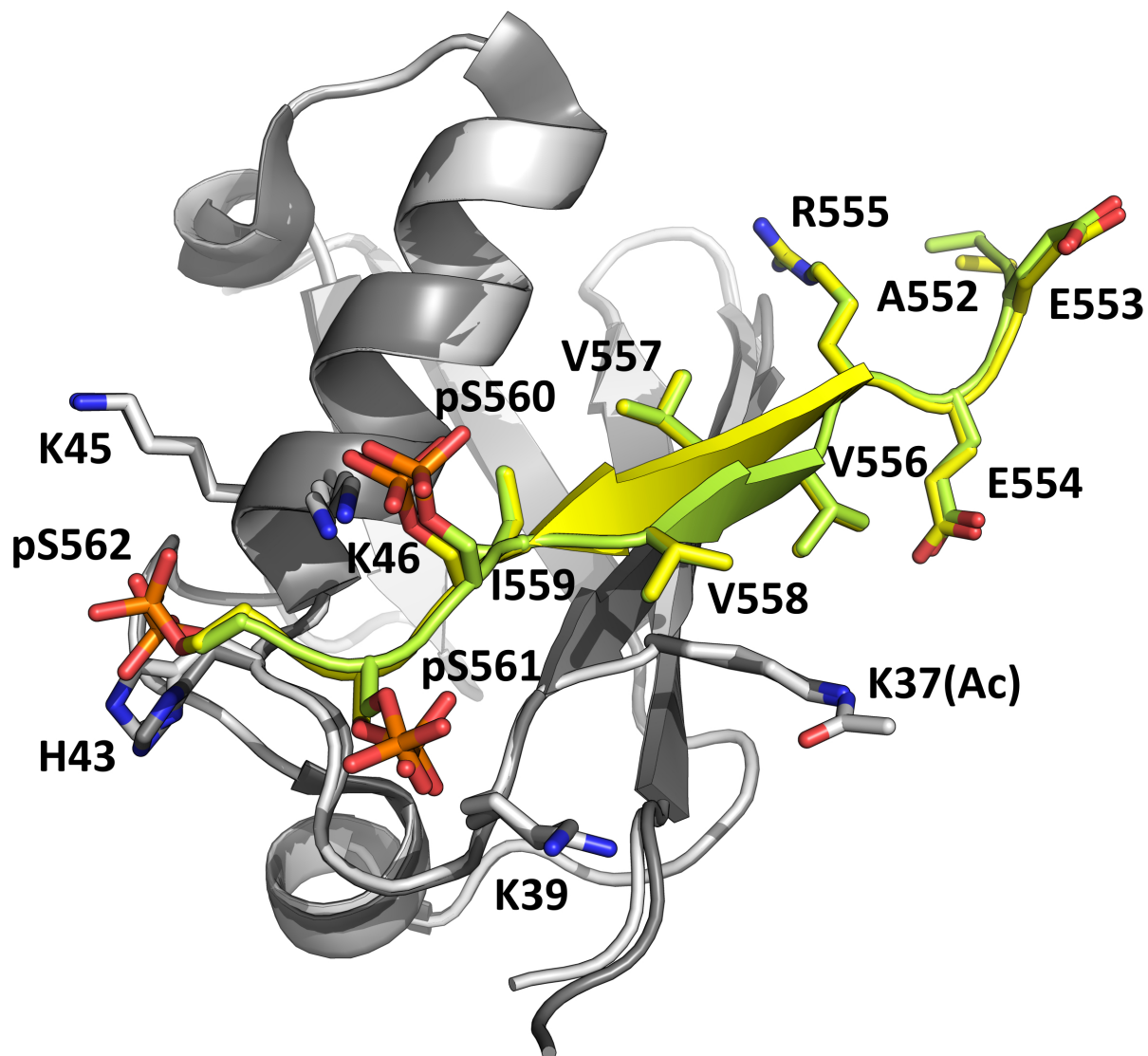


Figure 25. Comparison of Δ N-SUMO1:PML-SIM-4PO₄ (PDB entry [4WJN](#)) and Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ complexes. Overlay of the Δ N-SUMO1:PML-SIM-4PO₄ (carbon atoms dark gray:lime) and the Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ structures (carbon atoms light gray:yellow). Proteins are represented in cartoon; Δ N-SUMO1 basic residues and residues 552-562 of PML-SIM-4PO₄ are in stick representation.

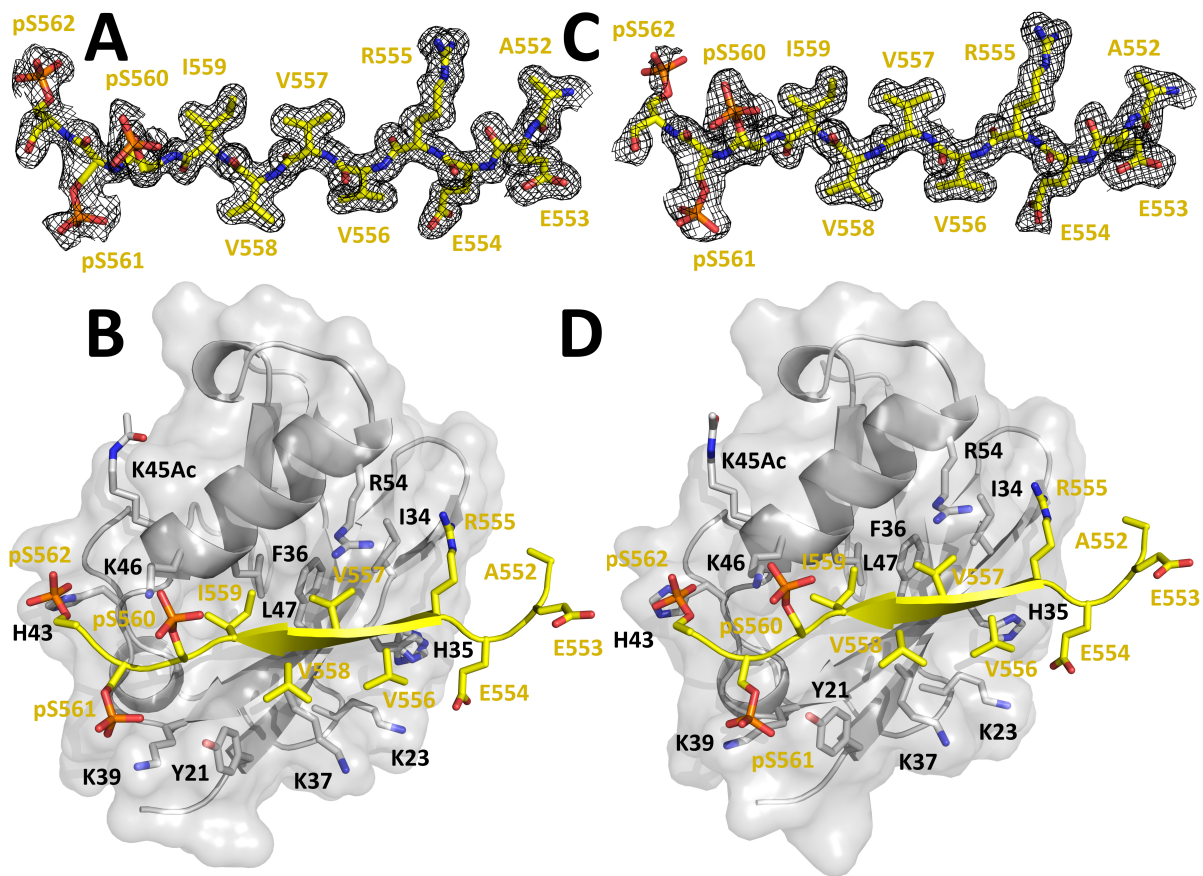


Figure 26. Crystal structures of Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ complexes. (A) Electron density map corresponding to residues 552-562 of PML-SIM-4PO₄ (chain B). The peptide is in stick representation with carbon atoms colored yellow. The $2F_oF_c$ electron density map, contoured at 0.7σ , is represented by a gray mesh. (B) Overall structure of the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ (chains A:B) complex highlighting the interaction between SUMO1 and residues 552-562 of PML. The proteins are in cartoon representation with Δ N-SUMO1-K45Ac carbon atoms (gray, chain A) and PML-SIM-4PO₄ peptide carbon atoms (yellow, chain B); Δ N-SUMO1-K45Ac residues interacting with the PML-SIM-4PO₄ peptide are in stick representation. (C) Electron density map corresponding to residues 552-562 of PML-SIM-4PO₄ (chain D). The peptide is in stick representation with carbon atoms colored yellow. The $2F_oF_c$ electron density map, contoured at 0.6σ , is represented by a gray mesh. (D) Overall structure of the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ (chains C:D) complex highlighting the interaction between SUMO1 and residues 552-562 of PML. The proteins are in cartoon representation with Δ N-SUMO1-K45Ac carbon atoms (gray, chain C) and PML-SIM-4PO₄ peptide carbon atoms (yellow, chain D); Δ N-SUMO1-K45Ac residues interacting with the PML-SIM-4PO₄ peptide are in stick representation.

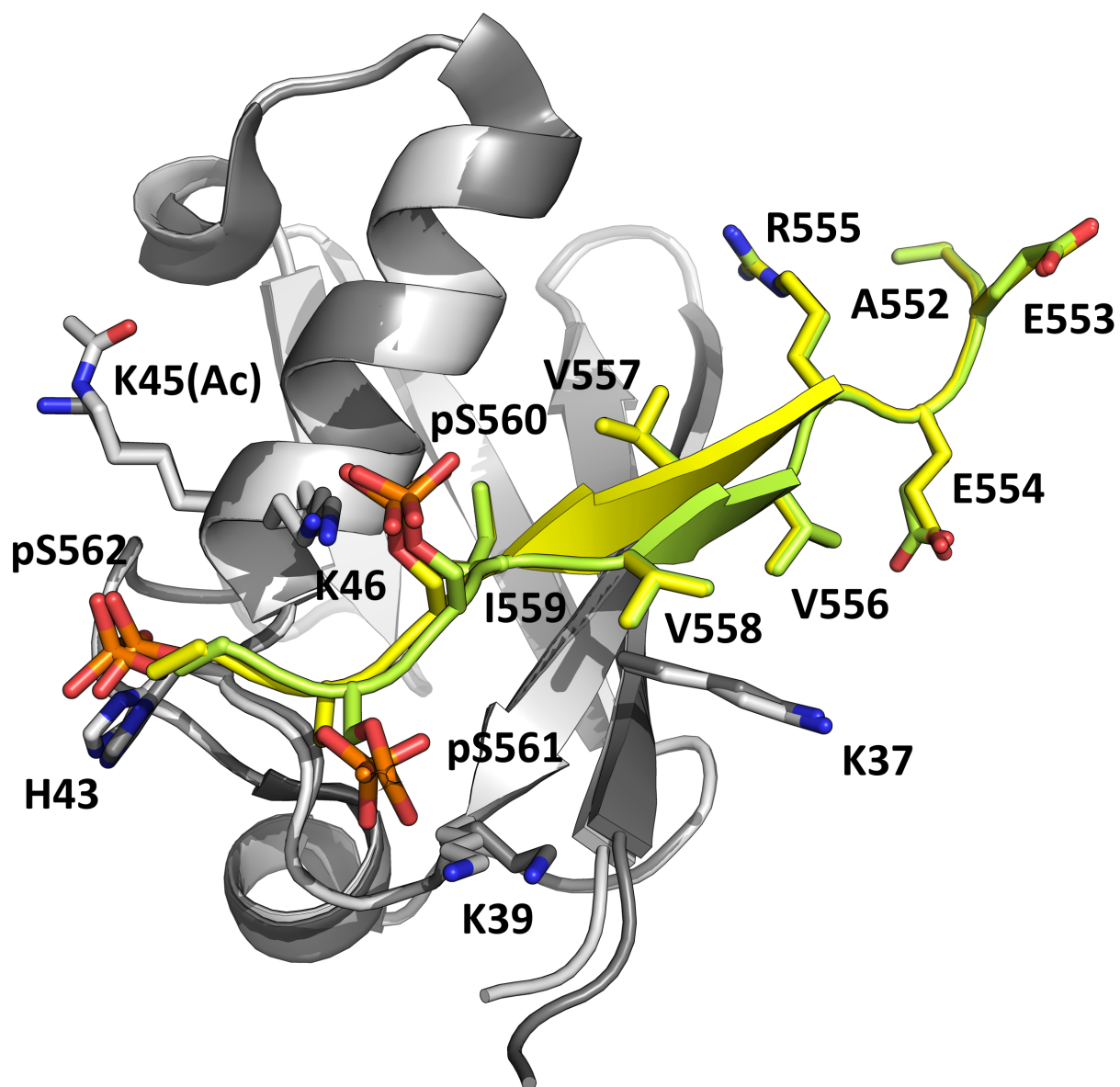


Figure 27. Comparison of Δ N-SUMO1:PML-SIM-4PO₄ (PDB entry [4WJN](#)) and Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ complexes. Overlay of the Δ N-SUMO1:PML-SIM-4PO₄ (carbon atoms dark gray:lime) and the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ structures (carbon atoms light gray:yellow). Proteins are represented in cartoon; Δ N-SUMO1 basic residues and residues 551-562 of PML-SIM-4PO₄ are in stick representation.

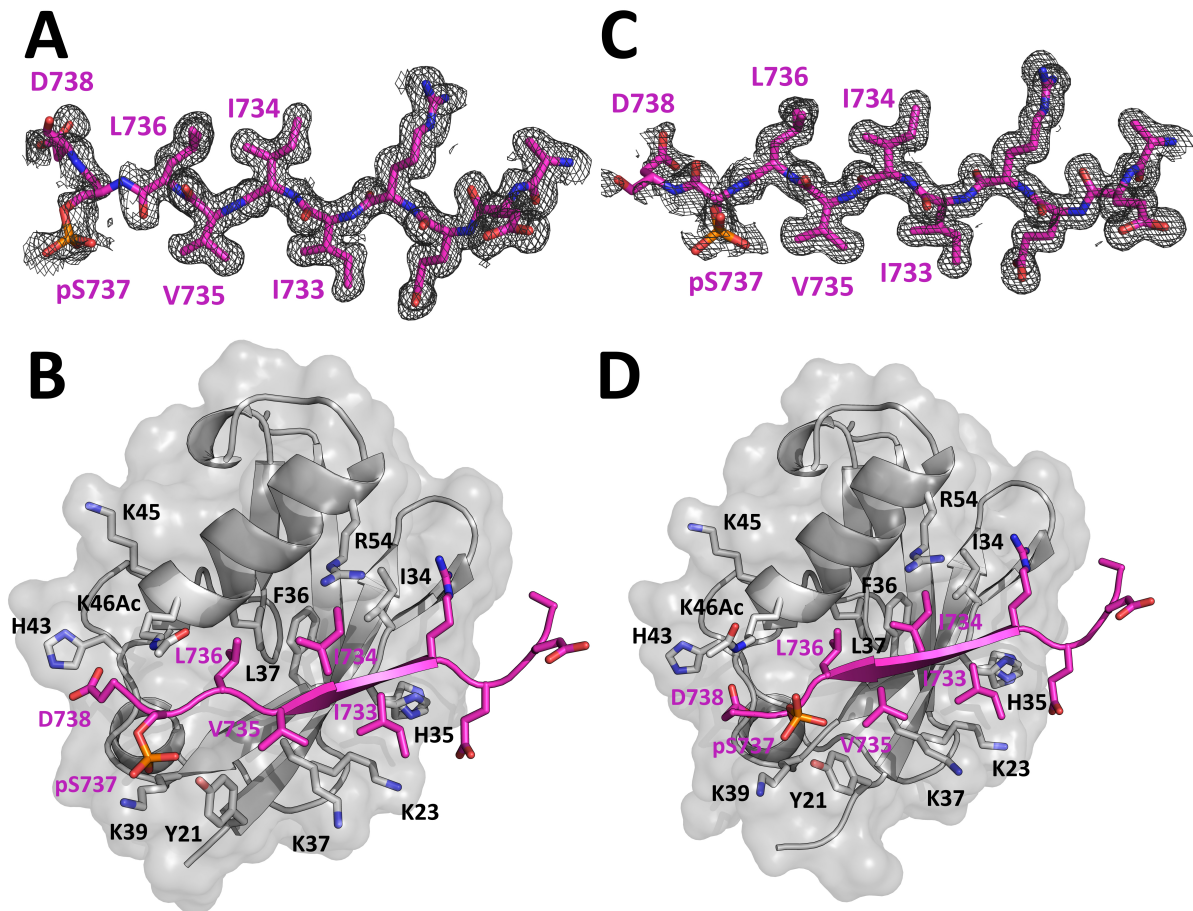


Figure 28. Crystal structure of Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complexes. (A) Electron density map corresponding to residues 733-738 of mDaxx-SIM-2PO₄ (chain B). The peptide is in stick representation with carbon atoms colored magenta. The $2F_oF_c$ electron density map, contoured at 0.6σ , is represented by a gray mesh. (B) Overall structure of the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ (chains A:B) complex highlighting the interaction between SUMO1 and residues 733-738 of Daxx. The proteins are in cartoon representation with Δ N-SUMO1-K46Ac carbon atoms (gray, chain A) and mDaxx-SIM-2PO₄ peptide carbon atoms (magenta, chain B); Δ N-SUMO1-K46Ac residues interacting with mDaxx-SIM-2PO₄ peptide are in stick representation. (C) Electron density map corresponding to residues 733-738 of mDaxx-SIM-2PO₄ (chain D). The peptide is in stick representation with carbon atoms colored magenta. The $2F_oF_c$ electron density map, contoured at 0.6σ , is represented by a gray mesh. (D) Overall structure of the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ (chains C:D) complex highlighting the interaction between SUMO1 and residues 733-738 of Daxx. The proteins are in cartoon representation with Δ N-SUMO1-K46Ac carbon atoms (gray, chain C) and mDaxx-SIM-2PO₄ peptide carbon atoms (magenta, chain D); Δ N-SUMO1-K46Ac residues interacting with the mDaxx-SIM-2PO₄ peptide are in stick representation. Unlabeled residues in N-terminus of mDaxx-SIM-2PO₄ correspond to residues 552-555 of PML-SIM.

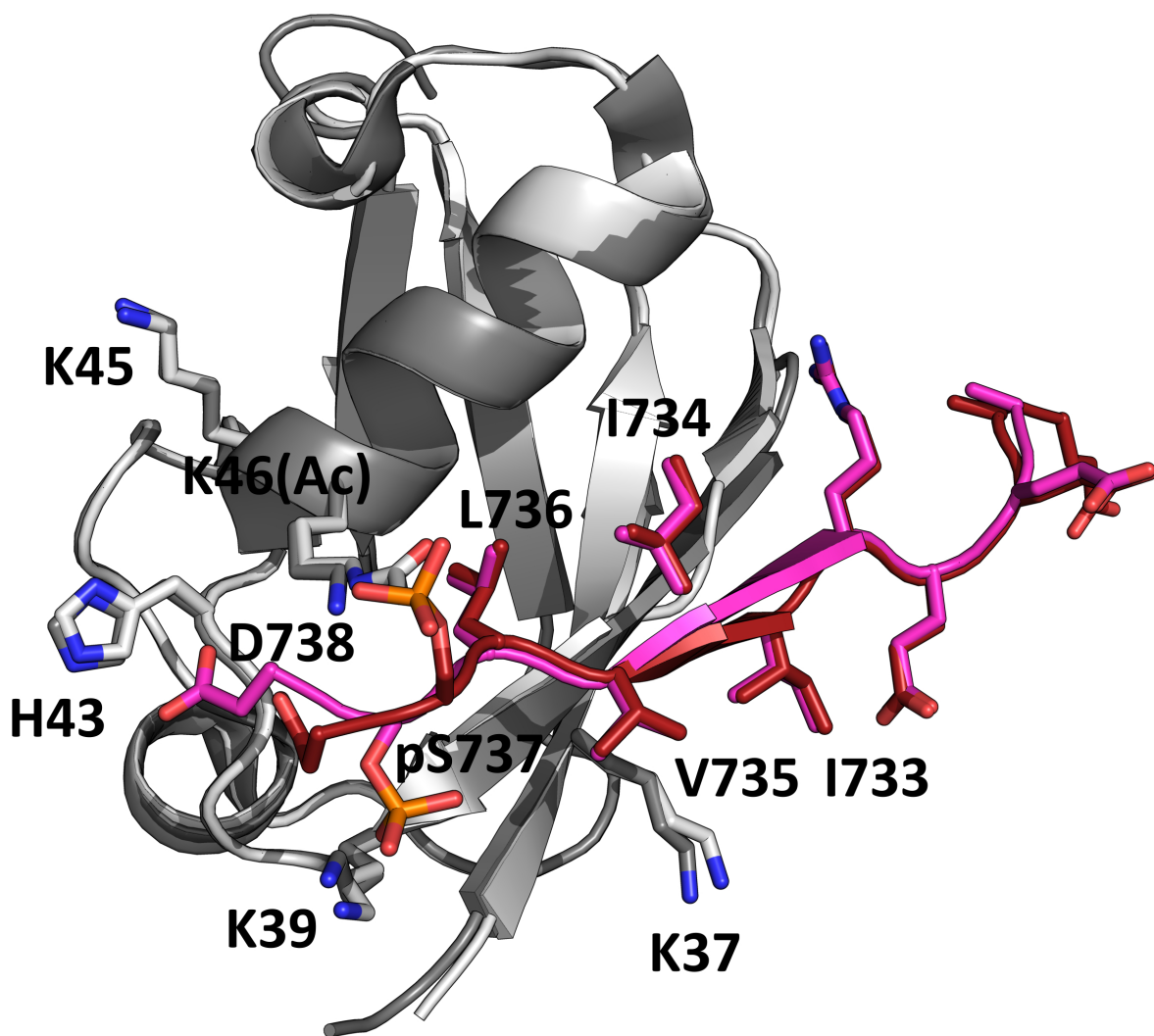


Figure 29. Comparison of Δ N-SUMO1:mDaxx-SIM-2PO₄ (PDB entry [4WJP](#)) and Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complexes. Overlay of the Δ N-SUMO1:Daxx-SIM-2PO₄ (carbon atoms dark gray:red) and the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ structures (carbon atoms light gray:magenta). Proteins are represented in cartoon; Δ N-SUMO1 basic residues and residues 733-738 of mDaxx-SIM-2PO₄ are in stick representation. Unlabeled residues at the N-terminus of mDaxx-SIM-2PO₄ correspond to residues 552-555 of PML-SIM.

Discussion générale et conclusion

Notre étude s'est intéressée à l'étude de l'interaction entre SUMO1 et le SIM de PML et plus précisément à l'impact concomitant de l'acétylation de SUMO1 et de la phosphorylation du SIM de PML sur cette interaction. Cependant, nous savons que la famille des protéines SUMO compte à ce jour cinq paralogues (SUMO1-5) (135). Il pourrait donc être pertinent d'évaluer l'effet de l'acétylation des quatre autres protéines SUMO (SUMO2-5) sur leur interaction avec le SIM de PML. Ainsi, il a été montré que la protéine SUMO2 est acétylée sur la lysine K33 (K37 chez SUMO1) et que cette modification altère son interaction avec les phospho-SIMs de PML, Daxx et des protéines de la famille PIAS (63). De plus, les résidus K39 et K46 connus pour être acétylés chez SUMO1 sont conservés chez les paralogues SUMO2/3 et SUMO4 (résidus K35 et K42). Il serait donc intéressant de générer les variants acétylés des différents paralogues SUMO pour des études thermodynamiques et structurales. Ceci permettrait de mieux comprendre les divers types d'interactions qu'adopte chaque protéine SUMO lors de son interaction avec des SIMs. Néanmoins, il est concevable d'anticiper que les modes d'interaction entre le phospho-SIM de PML et les variants acétylés des autres paralogues SUMO soient grandement similaires à ceux caractérisés pour les complexes entre le phospho-SIM de PML et les variants acétylés de SUMO1.

Les SIMs de PML et de Daxx étant très similaires, des études préliminaires ont aussi été entreprises pour caractériser l'interaction entre les variants acétylés de SUMO1 et le phospho-SIM de Daxx. Il a été possible d'identifier structurellement les particularités de chaque SIM en complexe avec les variants acétylés de SUMO1. Ainsi, une caractéristique commune ressort de nos études structurales. Il apparaît que l'acétylation du résidu K46 de SUMO engendre une importante réorganisation structurale pour les deux phospho-SIMs étudiés. Ceci corrobore notre hypothèse quant au rôle primordial de K46 dans l'établissement de contacts électrostatiques entre SUMO1 et les phospho-SIMs de PML et Daxx (85). De plus, les structures du phospho-SIM de Daxx en complexe avec les autres variants acétylés de SUMO1 ont été obtenues et quelques changements subtils ont pu être remarqués. Néanmoins, pour déterminer si ces modifications ont un impact significatif sur l'affinité de liaison entre chaque variant acétylé de SUMO1 et le phospho-SIM de Daxx des expériences d'ITC seront

nécessaires. Compte tenu de la modeste réorganisation structurale observée pour ces complexes il est probable que l'acétylation à ces autres lysines (K37, K39, K45) n'ait aucun effet sur l'interaction globale, ce qui remet en question leur pertinence dans un contexte *in vivo*. Effectivement, les lysines K37, K39 et K45 ne sont pas impliquées dans des liaisons électrostatiques avec le phospho-SIM de Daxx dans notre structure native (85), ce qui suggère que leur acétylation n'aura que peu ou pas d'impact sur les interactions SUMO/SIM.

Dans le but de déterminer si l'acétylation des protéines SUMO est utilisée comme un mécanisme général pour réguler les interactions SUMO/SIM, d'autres investigations ciblant les différentes familles de SIM seront nécessaires. En effet, la séquence primaire de la partie hydrophobe de chaque SIM est unique, mais la plus grande variation se retrouve incontestablement au niveau des résidus adjacents. En effet, le nombre de sites de phosphorylation, de résidus acides, ou encore l'espacement entre ces résidus et leur distance relative à la partie hydrophobe sont autant de variables à prendre en compte. Par exemple, les phospho-SIMs de PML, Daxx et SP100 contiennent des sérines phosphorylées immédiatement en C-terminal de la partie hydrophobe, alors que les sites de phosphorylation dans les SIMs des protéines PIAS se retrouvent quelques résidus plus loin (59). Néanmoins, les SIMs de PML et des protéines PIAS présentent tous deux une série de plusieurs sérines phosphorylables, alors que pour d'autres SIMs ces sérines sont parfois remplacées par des résidus acides, plutôt espacées au lieu d'être en bloc, ou encore trouvées en moins grand nombre. En plus de toutes ces subtilités structurales, la régulation des interactions SUMO/SIM mène à des conséquences cellulaires très variées. L'altération d'interactions SUMO/SIM peut avoir des issues multiples dépendamment des protéines spécifiques (paralogues de SUMO et phospho-SIM) impliquées, de la localisation spécifiques de ces facteurs à travers le noyau, du stade de la division cellulaire dans lequel se trouve la cellule, de la présence de transcription active à proximité, etc (5). Sachant ceci, il est possible d'envisager des études dans un contexte cellulaire afin d'obtenir des pistes de réponse à ces questions. L'utilisation des variants acétyl-mimétiques (K-to-Q) pour des études en cellules de levure et humaines peut être considérée. La levure est un organisme intéressant du fait de sa simplicité d'utilisation pour évaluer l'impact des variants acétylés de SUMO sur le contrôle de la transcription et l'expression génique (149). Les cellules humaines sont particulièrement utiles pour étudier des

phénomènes tels que la formation des corps nucléaires contenant PML, ou encore l'induction de la sénescence dépendante de PML (150) ainsi que l'implication de tous les autres facteurs d'intérêt dans le contexte des PML-NBs (109).

Demeure également la question des enzymes responsables de l'acétylation des protéines SUMO *in vivo*. Des études ont établi que l'acétylation des protéines SUMO est régulée par les histones déacétylases de classe I (63). Quant à l'identité des acétylases responsables de la modification des protéines SUMO, le même groupe a démontré que le facteur de transcription p300 peut acétyler SUMO2 dans des essais *in vitro* et que le bromodomaine (BRD) de p300 est en mesure d'interagir avec SUMO2 acétylée à la position K33. Cependant, p300 BRD n'interagit pas avec SUMO1 acétylée à la position K37, suggérant une certaine spécificité de reconnaissance des lysines acétylées par les protéines contenant des bromodomains. Néanmoins, p300 semble pouvoir interagir avec SUMO1 via son domaine doigt de zinc de type ZZ aussi retrouvé chez la protéine CBP. La surface d'interaction de SUMO1 impliquée dans la reconnaissance du domaine ZZ est distincte de celle requise pour la liaison aux SIMs, mais son implication fonctionnelle reste à déterminer (151). L'acétylation peut interférer avec les interactions SUMO/SIM et donner accès à des partenaires d'interactions alternatifs. En plus de reconnaître des patrons d'acétylation, la protéine p300 est également en mesure d'effectuer l'acétylation de ses partenaires d'interactions (histones et substrats autres). Donc, même si l'activité catalytique acétyltransférase de p300 envers les protéines de la famille SUMO n'a pas été prouvée hors de tout doute, il est fortement possible qu'elle soit en mesure d'acétyler *in vivo* minimalement SUMO2. Il est aussi probable que plusieurs acétylases différentes reconnaissent les protéines SUMO en tant que substrat. Le niveau de spécificité de ces enzymes présumées envers les divers paralogues SUMO est encore à déterminer.

Les données recueillies lors de nos études nous indiquent que les interactions SUMO/SIM sont d'une très grande plasticité, celles-ci pouvant être finement régulées par des modifications post-traductionnelles telles que la phosphorylation et l'acétylation. De plus, cette adaptabilité est renforcée par la capacité qu'ont les paralogues SUMO à accommoder un grand nombre de motifs SIM intrinsèquement différents dans leur séquence primaire mais dont les propriétés physico-chimiques (flexibilité, hydrophobicité, acidité) sont similaires. Ainsi, la

flexibilité des SIMs peut s'expliquer par leur localisation au sein de régions intrinsèquement désordonnées dans la séquence de leur protéine d'appartenance. Par opposition, l'identification de séquences de type SIM typiques dans des régions structurées n'est guère pertinente puisque ces régions sont souvent enfuies et non accessibles pour interagir avec les protéines SUMO. En addition, les lysines ciblées par l'acétylation de SUMO se trouvent aussi dans une région plus flexible chez ces protéines. Ces données concordent avec les nombreuses études relatant l'existence de multiples modifications post-traductionnelles ciblant les acides aminés présents dans les régions non-structurées des protéines (144, 145). De plus, il semble que l'orientation et la force des interactions électrostatiques à l'interface de l'interaction SUMO1/PML-SIM-4PO₄ dictent l'impact que l'acétylation aura sur le complexe global. Il est possible de concevoir que l'acétylation de K46 ou K39 nuise directement aux liaisons électrostatiques, alors que celle de K37 ou K45 cause plutôt des réarrangements subtils au niveau des interactions hydrophobes rendant cette région de SUMO1 légèrement plus flexible. Il a été suggéré que l'acétylation de K37 (K33 chez SUMO2) *in vivo* modifie la dynamique des corps nucléaires contenant PML (63). Sachant que plusieurs protéines transitant à l'intérieur des corps nucléaires contiennent des séquences SIMs, il est très probable que l'acétylation de K37 affecte l'interaction de SUMO1 avec plusieurs protéines autres que PML qui sont importantes pour l'homéostasie des corps nucléaires. Par exemple, la protéine PIAS1 est un E3 ligase qui interagit avec SUMO1 afin d'accomplir ses fonctions. Si l'interaction avec SUMO1 et PIAS1 est contrainte, les niveaux de SUMOylation globaux dans la cellule seront diminués et l'intégrité des corps nucléaires en sera affectée puisque leur structure dépend de la SUMOylation (109). Ainsi, il sera intéressant d'utiliser nos variants K39Q et K46Q dans des études similaires afin de déterminer l'ampleur de leur effet en cellules. Ultiment, l'introduction des véritables protéines recombinantes acétylées est voulue, puisque ceci permettrait d'étudier *in vivo* les impacts et la dynamique de l'acétylation en modulant l'activité enzymatique d'enzymes (dé)acétylases, ou encore la capacité de liaisons des protéines contenant des bromodomains par des traitements avec des inhibiteurs spécifiques.

Nos résultats démontrent clairement que l'acétylation des protéines SUMO et la phosphorylation des SIMs agissent de façon réciproque, mais antagoniste, pour réguler les interactions SUMO/SIM. L'acétylation des protéines SUMO diminue leur affinité pour les phospho-SIMs et ceci survient via diverses réorganisations structurales qui permettent d'expliquer les différents effets observés pour chaque variant.

Bibliographie

1. C. Walsh, *Posttranslational Modification of Proteins. Expanding Nature's Inventory.*, (Roberts & Company Publishers, Greenwood Village, 2005).
2. Y. L. Deribe, T. Pawson, I. Dikic, Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666-672 (2010).
3. B. T. Seet, I. Dikic, M. M. Zhou, T. Pawson, Reading protein modifications with interaction domains. *Nature reviews. Molecular cell biology* **7**, 473-483 (2006).
4. G. A. Khoury, R. C. Baliban, C. A. Floudas, Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific reports* **1**, (2011).
5. J. R. Gareau, C. D. Lima, The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature reviews. Molecular cell biology* **11**, 861-871 (2010).
6. R. T. Hay, SUMO: a history of modification. *Molecular cell* **18**, 1-12 (2005).
7. K. O. Alegre, D. Reverter, Swapping small ubiquitin-like modifier (SUMO) isoform specificity of SUMO proteases SENP6 and SENP7. *The Journal of biological chemistry* **286**, 36142-36151 (2011).
8. H. L. Su, S. S. Li, Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. *Gene* **296**, 65-73 (2002).
9. X. D. Zhang *et al.*, SUMO-2/3 modification and binding regulate the association of CENP-E with kinetochores and progression through mitosis. *Molecular cell* **29**, 729-741 (2008).
10. C. Y. Wang, J. X. She, SUMO4 and its role in type 1 diabetes pathogenesis. *Diabetes Metab. Res. Rev.* **24**, 93-102 (2008).
11. S. E. Schwarz, K. Matuschewski, D. Liakopoulos, M. Scheffner, S. Jentsch, The ubiquitin-like proteins SMT3 and SUMO-1 are conjugated by the UBC9 E2 enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 560-564 (1998).

12. E. Urena *et al.*, Evolution of SUMO Function and Chain Formation in Insects. *Molecular biology and evolution* **33**, 568-584 (2016).
13. P. Bayer *et al.*, Structure determination of the small ubiquitin-related modifier SUMO-1. *Journal of molecular biology* **280**, 275-286 (1998).
14. I. Matic *et al.*, In vivo identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an in vitro to in vivo strategy. *Mol Cell Proteomics* **7**, 132-144 (2008).
15. O. Kerscher, R. Felberbaum, M. Hochstrasser, Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**, 159-180 (2006).
16. J. J. Palvimo, PIAS proteins as regulators of small ubiquitin-related modifier (SUMO) modifications and transcription. *Biochemical Society transactions* **35**, 1405-1408 (2007).
17. D. Schmidt, S. Muller, PIAS/SUMO: new partners in transcriptional regulation. *Cellular and molecular life sciences : CMLS* **60**, 2561-2574 (2003).
18. D. M. Wenzel, K. E. Stoll, R. E. Klevit, E2s: structurally economical and functionally replete. *The Biochemical journal* **433**, 31-42 (2011).
19. R. T. Hay, SUMO-specific proteases: a twist in the tail. *Trends in cell biology* **17**, 370-376 (2007).
20. D. Mukhopadhyay, M. Dasso, Modification in reverse: the SUMO proteases. *Trends in biochemical sciences* **32**, 286-295 (2007).
21. R. D. Everett, C. Boutell, B. G. Hale, Interplay between viruses and host sumoylation pathways. *Nat. Rev. Microbiol.* **11**, 400-411 (2013).
22. M. S. Rodriguez, C. Dargemont, R. T. Hay, SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *The Journal of biological chemistry* **276**, 12654-12659 (2001).
23. D. A. Sampson, M. Wang, M. J. Matunis, The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *The Journal of biological chemistry* **276**, 21664-21669 (2001).
24. V. Bernier-Villamor, D. A. Sampson, M. J. Matunis, C. D. Lima, Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell* **108**, 345-356 (2002).

25. D. Lin *et al.*, Identification of a substrate recognition site on Ubc9. *The Journal of biological chemistry* **277**, 21740-21748 (2002).
26. S. H. Yang, A. Galanis, J. Witty, A. D. Sharrocks, An extended consensus motif enhances the specificity of substrate modification by SUMO. *The EMBO journal* **25**, 5083-5093 (2006).
27. X. J. Yang, S. Gregoire, A recurrent phospho-sumoyl switch in transcriptional repression and beyond. *Molecular cell* **23**, 779-786 (2006).
28. F. Mohideen *et al.*, A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. *Nat Struct Mol Biol* **16**, 945-952 (2009).
29. D. Komander, M. Rape, The ubiquitin code. *Annu Rev Biochem* **81**, 203-229 (2012).
30. R. Mahajan, C. Delphin, T. Guan, L. Gerace, F. Melchior, A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88**, 97-107 (1997).
31. M. J. Matunis, E. Coutavas, G. Blobel, A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* **135**, 1457-1470 (1996).
32. G. Gill, Something about SUMO inhibits transcription. *Curr Opin Genet Dev* **15**, 536-541 (2005).
33. S. P. Jackson, D. Durocher, Regulation of DNA damage responses by ubiquitin and SUMO. *Molecular cell* **49**, 795-807 (2013).
34. H. D. Ulrich, H. Walden, Ubiquitin signalling in DNA replication and repair. *Nature reviews. Molecular cell biology* **11**, 479-489 (2010).
35. J. Schimmel *et al.*, Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Molecular cell* **53**, 1053-1066 (2014).
36. T. Buschmann, S. Y. Fuchs, C. G. Lee, Z. Q. Pan, Z. Ronai, SUMO-1 modification of Mdm2 prevents its self-ubiquitination and increases Mdm2 ability to ubiquitinate p53. *Cell* **101**, 753-762 (2000).
37. C. Klenk, J. Humrich, U. Qitterer, M. J. Lohse, SUMO-1 controls the protein stability and the biological function of phosphatidylinositol 3-kinase. *The Journal of biological chemistry* **281**, 8357-8364 (2006).

38. H. Saitoh, J. Hinchey, Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *The Journal of biological chemistry* **275**, 6252-6258 (2000).
39. W. Wei *et al.*, A stress-dependent SUMO4 sumoylation of its substrate proteins. *Biochemical and biophysical research communications* **375**, 454-459 (2008).
40. Y. Azuma, A. Arnaoutov, M. Dasso, SUMO-2/3 regulates topoisomerase II in mitosis. *J Cell Biol* **163**, 477-487 (2003).
41. D. Guo *et al.*, A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. *Nature genetics* **36**, 837-841 (2004).
42. C. D. Chung *et al.*, Specific inhibition of Stat3 signal transduction by PIAS3. *Science* **278**, 1803-1805 (1997).
43. A. Pichler, A. Gast, J. S. Seeler, A. Dejean, F. Melchior, The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell* **108**, 109-120 (2002).
44. M. H. Kagey, T. A. Melhuish, D. Wotton, The polycomb protein Pc2 is a SUMO E3. *Cell* **113**, 127-137 (2003).
45. M. M. Rytinki, S. Kaikkonen, P. Pehkonen, T. Jaaskelainen, J. J. Palvimo, PIAS proteins: pleiotropic interactors associated with SUMO. *Cellular and molecular life sciences : CMLS* **66**, 3029-3041 (2009).
46. T. Makhnevych, C. Ptak, C. P. Lusk, J. D. Aitchison, R. W. Wozniak, The role of karyopherins in the regulated sumoylation of septins. *J Cell Biol* **177**, 39-49 (2007).
47. A. Roscic *et al.*, Phosphorylation-dependent control of Pc2 SUMO E3 ligase activity by its substrate protein HIPK2. *Molecular cell* **24**, 77-89 (2006).
48. M. Ihara, H. Yamamoto, A. Kikuchi, SUMO-1 modification of PIASy, an E3 ligase, is necessary for PIASy-dependent activation of Tcf-4. *Molecular and cellular biology* **25**, 3506-3518 (2005).
49. R. J. Deshaies, C. A. Joazeiro, RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**, 399-434 (2009).
50. A. Reindle *et al.*, Multiple domains in Siz SUMO ligases contribute to substrate selectivity. *Journal of cell science* **119**, 4749-4757 (2006).
51. M. M. Dawlaty *et al.*, Resolution of sister centromeres requires RanBP2-mediated SUMOylation of topoisomerase IIalpha. *Cell* **133**, 103-115 (2008).

52. A. Shalizi *et al.*, PIASx is a MEF2 SUMO E3 ligase that promotes postsynaptic dendritic morphogenesis. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **27**, 10037-10046 (2007).
53. A. A. Yunus, C. D. Lima, Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA. *Molecular cell* **35**, 669-682 (2009).
54. A. Flotho, F. Melchior, Sumoylation: a regulatory protein modification in health and disease. *Annu Rev Biochem* **82**, 357-385 (2013).
55. S. Zhu *et al.*, Protection from isopeptidase-mediated deconjugation regulates paralogue-selective sumoylation of RanGAP1. *Molecular cell* **33**, 570-580 (2009).
56. O. Kerscher, SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO reports* **8**, 550-555 (2007).
57. G. J. Praefcke, K. Hofmann, R. J. Dohmen, SUMO playing tag with ubiquitin. *Trends in biochemical sciences* **37**, 23-31 (2012).
58. J. T. Hannich *et al.*, Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae*. *The Journal of biological chemistry* **280**, 4102-4110 (2005).
59. A. Minty, X. Dumont, M. Kaghad, D. Caput, Covalent modification of p73alpha by SUMO-1. Two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. *The Journal of biological chemistry* **275**, 36316-36323 (2000).
60. J. Song, L. K. Durrin, T. A. Wilkinson, T. G. Krontiris, Y. Chen, Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14373-14378 (2004).
61. P. Stehmeier, S. Muller, Phospho-regulated SUMO interaction modules connect the SUMO system to CK2 signaling. *Molecular cell* **33**, 400-409 (2009).
62. D. Reverter, C. D. Lima, Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* **435**, 687-692 (2005).

63. R. Ullmann, C. D. Chien, M. L. Avantaggiati, S. Muller, An acetylation switch regulates SUMO-dependent protein interaction networks. *Molecular cell* **46**, 759-770 (2012).
64. J. J. Perry, J. A. Tainer, M. N. Boddy, A simultaneous role for SUMO and ubiquitin. *Trends in biochemical sciences* **33**, 201-208 (2008).
65. R. Bernardi, P. P. Pandolfi, Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nature reviews. Molecular cell biology* **8**, 1006-1016 (2007).
66. C. M. Hecker, M. Rabiller, K. Haglund, P. Bayer, I. Dikic, Specification of SUMO1- and SUMO2-interacting motifs. *The Journal of biological chemistry* **281**, 16117-16127 (2006).
67. J. Zhu *et al.*, Small ubiquitin-related modifier (SUMO) binding determines substrate recognition and paralog-selective SUMO modification. *The Journal of biological chemistry* **283**, 29405-29415 (2008).
68. J. R. Mullen, S. J. Brill, Activation of the Slx5-Slx8 ubiquitin ligase by poly-small ubiquitin-like modifier conjugates. *The Journal of biological chemistry* **283**, 19912-19921 (2008).
69. N. Sekiyama *et al.*, Structural basis for regulation of poly-SUMO chain by a SUMO-like domain of Nip45. *Proteins* **78**, 1491-1502 (2010).
70. V. Lallemand-Breitenbach *et al.*, Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nature cell biology* **10**, 547-555 (2008).
71. M. H. Tatham *et al.*, RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nature cell biology* **10**, 538-546 (2008).
72. S. K. Olsen, A. D. Capili, X. Lu, D. S. Tan, C. D. Lima, Active site remodelling accompanies thioester bond formation in the SUMO E1. *Nature* **463**, 906-912 (2010).
73. L. M. Lois, C. D. Lima, Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *The EMBO journal* **24**, 439-451 (2005).
74. M. H. Tatham *et al.*, Unique binding interactions among Ubc9, SUMO and RanBP2 reveal a mechanism for SUMO paralog selection. *Nat Struct Mol Biol* **12**, 67-74 (2005).

75. X. H. Mascle *et al.*, Identification of a non-covalent ternary complex formed by PIAS1, SUMO1, and UBC9 proteins involved in transcriptional regulation. *The Journal of biological chemistry* **288**, 36312-36327 (2013).
76. E. S. Lander *et al.*, Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
77. J. C. Venter *et al.*, The sequence of the human genome. *Science* **291**, 1304-1351 (2001).
78. P. Cohen, The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends in biochemical sciences* **25**, 596-601 (2000).
79. T. Hunter, Signaling--2000 and beyond. *Cell* **100**, 113-127 (2000).
80. E. G. Krebs, The growth of research on protein phosphorylation. *Trends in biochemical sciences* **19**, 439 (1994).
81. S. K. Hanks, T. Hunter, Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **9**, 576-596 (1995).
82. S. Sarno, P. Vaglio, F. Meggio, O. G. Issinger, L. A. Pinna, Protein kinase CK2 mutants defective in substrate recognition. Purification and kinetic analysis. *The Journal of biological chemistry* **271**, 10595-10601 (1996).
83. E. Gresko *et al.*, PML tumor suppressor is regulated by HIPK2-mediated phosphorylation in response to DNA damage. *Oncogene* **28**, 698-708 (2009).
84. T. Kamitani *et al.*, Identification of three major sentrinization sites in PML. *The Journal of biological chemistry* **273**, 26675-26682 (1998).
85. L. Cappadocia *et al.*, Structural and functional characterization of the phosphorylation-dependent interaction between PML and SUMO1. *Structure* **23**, 126-138 (2015).
86. K. K. Lee, J. L. Workman, Histone acetyltransferase complexes: one size doesn't fit all. *Nature reviews. Molecular cell biology* **8**, 284-295 (2007).
87. M. A. Glozak, N. Sengupta, X. Zhang, E. Seto, Acetylation and deacetylation of non-histone proteins. *Gene* **363**, 15-23 (2005).

88. I. Matic, B. Macek, M. Hilger, T. C. Walther, M. Mann, Phosphorylation of SUMO-1 occurs in vivo and is conserved through evolution. *J Proteome Res* **7**, 4050-4057 (2008).
89. A. Cheema *et al.*, Functional mimicry of the acetylated C-terminal tail of p53 by a SUMO-1 acetylated domain, SAD. *Journal of cellular physiology* **225**, 371-384 (2010).
90. C. Choudhary *et al.*, Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840 (2009).
91. C. C. Chang *et al.*, Structural and functional roles of Daxx SIM phosphorylation in SUMO paralog-selective binding and apoptosis modulation. *Molecular cell* **42**, 62-74 (2011).
92. A. Rosendorff *et al.*, NXP-2 association with SUMO-2 depends on lysines required for transcriptional repression. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 5308-5313 (2006).
93. K. Jensen, C. Shiels, P. S. Freemont, PML protein isoforms and the RBCC/TRIM motif. *Oncogene* **20**, 7223-7233 (2001).
94. A. Reymond *et al.*, The tripartite motif family identifies cell compartments. *The EMBO journal* **20**, 2140-2151 (2001).
95. M. Buschbeck *et al.*, PML4 induces differentiation by Myc destabilization. *Oncogene* **26**, 3415-3422 (2007).
96. V. Fogal *et al.*, Regulation of p53 activity in nuclear bodies by a specific PML isoform. *The EMBO journal* **19**, 6185-6195 (2000).
97. H. Yoshida *et al.*, PML-retinoic acid receptor alpha inhibits PML IV enhancement of PU.1-induced C/EBPepsilon expression in myeloid differentiation. *Molecular and cellular biology* **27**, 5819-5834 (2007).
98. M. T. Daniel *et al.*, PML protein expression in hematopoietic and acute promyelocytic leukemia cells. *Blood* **82**, 1858-1867 (1993).
99. J. Zhu, V. Lallemand-Breitenbach, H. de The, Pathways of retinoic acid- or arsenic trioxide-induced PML/RARalpha catabolism, role of oncogene degradation in disease remission. *Oncogene* **20**, 7257-7265 (2001).

100. S. Tessier, N. Martin-Martin, H. de The, A. Carracedo, V. Lallemand-Breitenbach, PML, a protein at the cross road of oxidative stress and metabolism. *Antioxid. Redox Signal.*, (2016).
101. A. M. Ishov *et al.*, PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. *J Cell Biol* **147**, 221-234 (1999).
102. N. Stuurman *et al.*, The nuclear matrix from cells of different origin. Evidence for a common set of matrix proteins. *The Journal of biological chemistry* **265**, 5460-5465 (1990).
103. J. Wang *et al.*, Promyelocytic leukemia nuclear bodies associate with transcriptionally active genomic regions. *J Cell Biol* **164**, 515-526 (2004).
104. E. de Stanchina *et al.*, PML is a direct p53 target that modulates p53 effector functions. *Molecular cell* **13**, 523-535 (2004).
105. R. D. Everett *et al.*, PML contributes to a cellular mechanism of repression of herpes simplex virus type 1 infection that is inactivated by ICP0. *Journal of virology* **80**, 7995-8005 (2006).
106. G. Dellaire, D. P. Bazett-Jones, PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *BioEssays : news and reviews in molecular, cellular and developmental biology* **26**, 963-977 (2004).
107. S. Zhong, P. Salomoni, P. P. Pandolfi, The transcriptional role of PML and the nuclear body. *Nature cell biology* **2**, E85-90 (2000).
108. T. Kamitani, H. P. Nguyen, K. Kito, T. Fukuda-Kamitani, E. T. Yeh, Covalent modification of PML by the sentrin family of ubiquitin-like proteins. *The Journal of biological chemistry* **273**, 3117-3120 (1998).
109. V. Lallemand-Breitenbach, H. de The, PML nuclear bodies. *Cold Spring Harb Perspect Biol* **2**, a000661 (2010).
110. P. P. Scaglioni *et al.*, A CK2-dependent mechanism for degradation of the PML tumor suppressor. *Cell* **126**, 269-283 (2006).
111. O. G. Engelhardt *et al.*, The homeodomain-interacting kinase PKM (HIPK-2) modifies ND10 through both its kinase domain and a SUMO-1 interaction motif and alters the posttranslational modification of PML. *Exp Cell Res* **283**, 36-50 (2003).

112. F. Hayakawa, M. L. Privalsky, Phosphorylation of PML by mitogen-activated protein kinases plays a key role in arsenic trioxide-mediated apoptosis. *Cancer cell* **5**, 389-401 (2004).
113. G. Dellaire, D. P. Bazett-Jones, Beyond repair foci: subnuclear domains and the cellular response to DNA damage. *Cell cycle* **6**, 1864-1872 (2007).
114. P. Salomoni, M. Dvorkina, D. Michod, Role of the promyelocytic leukaemia protein in cell death regulation. *Cell Death Dis* **3**, e247 (2012).
115. S. Zhong *et al.*, Promyelocytic leukemia protein (PML) and Daxx participate in a novel nuclear pathway for apoptosis. *The Journal of experimental medicine* **191**, 631-640 (2000).
116. G. Cho, Y. Lim, J. A. Golden, SUMO interaction motifs in Sizn1 are required for promyelocytic leukemia protein nuclear body localization and for transcriptional activation. *The Journal of biological chemistry* **284**, 19592-19600 (2009).
117. A. Rabellino *et al.*, The SUMO E3-ligase PIAS1 regulates the tumor suppressor PML and its oncogenic counterpart PML-RARA. *Cancer research* **72**, 2275-2284 (2012).
118. P. P. Scaglioni *et al.*, CK2 mediates phosphorylation and ubiquitin-mediated degradation of the PML tumor suppressor. *Molecular and cellular biochemistry* **316**, 149-154 (2008).
119. J. W. Chin, Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem* **83**, 379-408 (2014).
120. H. Neumann *et al.*, A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Molecular cell* **36**, 153-163 (2009).
121. S. J. Elsasser, R. J. Ernst, O. S. Walker, J. W. Chin, Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nature methods* **13**, 158-164 (2016).
122. K. Kamieniarz, R. Schneider, Tools to tackle protein acetylation. *Chemistry & biology* **16**, 1027-1029 (2009).
123. M. M. Pierce, C. S. Raman, B. T. Nall, Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**, 213-221 (1999).

124. Y. Nozaki, C. Tanford, The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *The Journal of biological chemistry* **246**, 2211-2217 (1971).
125. G. I. Makhatadze, P. L. Privalov, Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *Journal of molecular biology* **213**, 375-384 (1990).
126. E. F. Garman, Developments in x-ray crystallographic structure determination of biological macromolecules. *Science* **343**, 1102-1108 (2014).
127. C. Nave, E. F. Garman, Towards an understanding of radiation damage in cryocooled macromolecular crystals. *J Synchrotron Radiat* **12**, 257-260 (2005).
128. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242 (2011).
129. P. V. Afonine *et al.*, Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica. Section D, Biological crystallography* **68**, 352-367 (2012).
130. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
131. A. C. Pike, E. F. Garman, T. Krojer, F. von Delft, E. P. Carpenter, An overview of heavy-atom derivatization of protein crystals. *Acta Crystallogr D Struct Biol* **72**, 303-318 (2016).
132. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).
133. H. M. Berman *et al.*, The Protein Data Bank. *Nucleic acids research* **28**, 235-242 (2000).
134. R. Geiss-Friedlander, F. Melchior, Concepts in sumoylation: a decade on. *Nature reviews. Molecular cell biology* **8**, 947-956 (2007).
135. Y. C. Liang *et al.*, SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies. *Scientific reports* **6**, 26509 (2016).
136. M. Haindl, T. Harasim, D. Eick, S. Muller, The nucleolar SUMO-specific protease SENP3 reverses SUMO modification of nucleophosmin and is required for rRNA processing. *EMBO reports* **9**, 273-279 (2008).

137. S. Muller, A. Ledl, D. Schmidt, SUMO: a regulator of gene expression and genome integrity. *Oncogene* **23**, 1998-2008 (2004).
138. M. J. Matunis, X. D. Zhang, N. A. Ellis, SUMO: the glue that binds. *Developmental cell* **11**, 596-597 (2006).
139. K. S. Sung *et al.*, Role of the SUMO-interacting motif in HIPK2 targeting to the PML nuclear bodies and regulation of p53. *Exp Cell Res* **317**, 1060-1070 (2011).
140. Z. A. Rasheed, A. Saleem, Y. Ravee, P. P. Pandolfi, E. H. Rubin, The topoisomerase I-binding RING protein, topors, is associated with promyelocytic leukemia nuclear bodies. *Exp Cell Res* **277**, 152-160 (2002).
141. D. Negorev, A. M. Ishov, G. G. Maul, Evidence for separate ND10-binding and homooligomerization domains of Sp100. *Journal of cell science* **114**, 59-68 (2001).
142. D. Y. Lin *et al.*, Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Molecular cell* **24**, 341-354 (2006).
143. H. Neumann, S. Y. Peak-Chew, J. W. Chin, Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **4**, 232-234 (2008).
144. L. M. Iakoucheva *et al.*, The importance of intrinsic disorder for protein phosphorylation. *Nucleic acids research* **32**, 1037-1049 (2004).
145. J. C. Hansen, X. Lu, E. D. Ross, R. W. Woody, Intrinsic protein disorder, amino acid composition, and histone terminal domains. *The Journal of biological chemistry* **281**, 1853-1856 (2006).
146. H. Neumann, J. L. Hazen, J. Weinstein, R. A. Mehl, J. W. Chin, Genetically encoding protein oxidative damage. *Journal of the American Chemical Society* **130**, 4028-4033 (2008).
147. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-2132 (2004).
148. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221 (2010).
149. A. Rienzo, A. Pascual-Ahuir, M. Proft, The use of a real-time luciferase assay to quantify gene expression dynamics in the living yeast cell. *Yeast* **29**, 219-231 (2012).

150. G. Ferbeyre *et al.*, PML is induced by oncogenic ras and promotes premature senescence. *Genes & development* **14**, 2015-2027 (2000).
151. C. Diehl *et al.*, Structural Analysis of a Complex between Small Ubiquitin-like Modifier 1 (SUMO1) and the ZZ Domain of CREB-binding Protein (CBP/p300) Reveals a New Interaction Surface on SUMO. *The Journal of biological chemistry* **291**, 12658-12672 (2016).

Supplementary figures

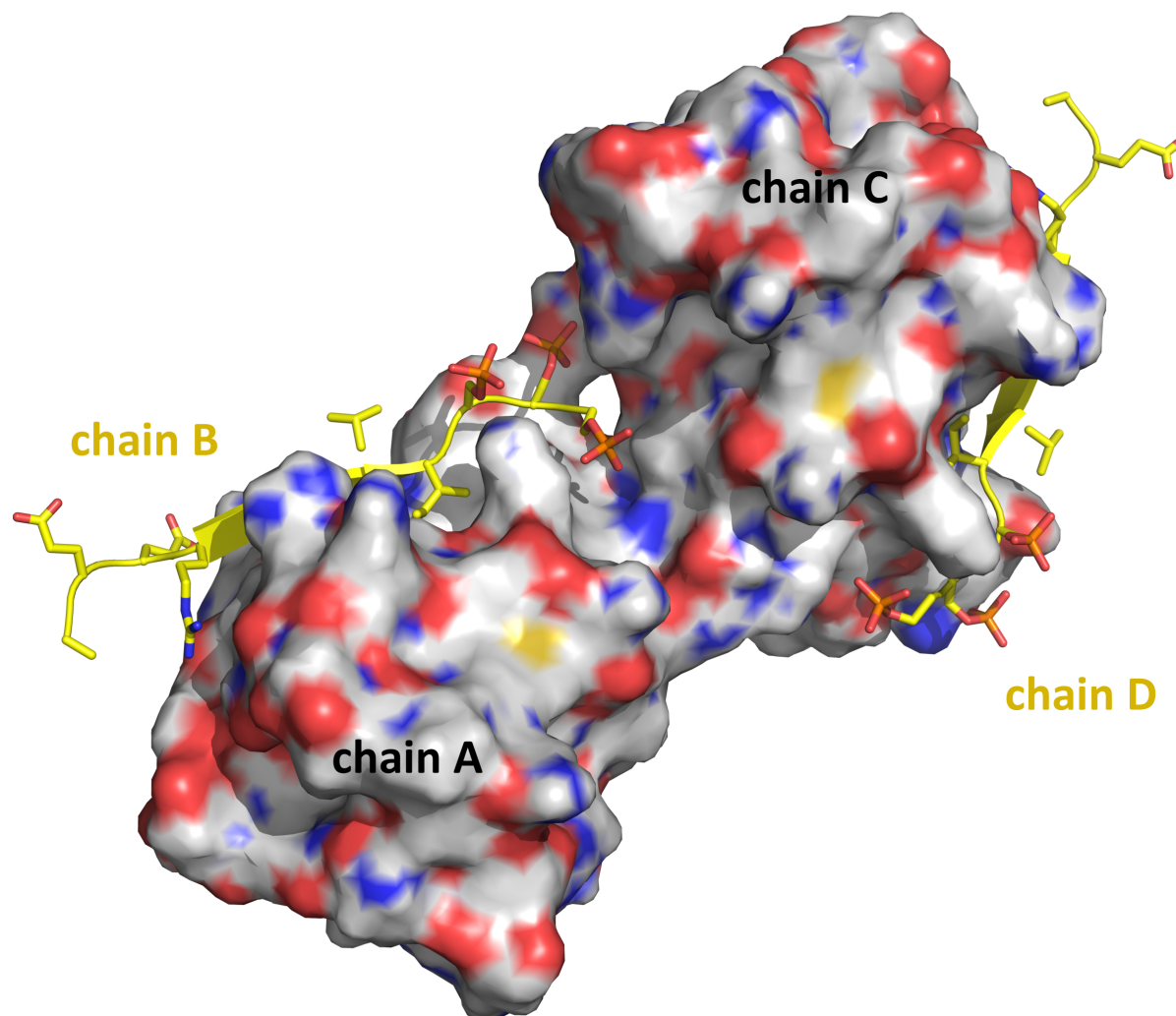


Figure S1. Structure of the two complexes present in the asymmetric unit of the ΔN -SUMO1-K37Ac:PML-SIM-4PO₄ crystal. The SUMO1 proteins (chain A and chain C) are shown as a surface representation where carbon atoms are colored gray, sulphur atoms yellow, acidic residues red, and basic residues blue. The PML-SIM-4PO₄ peptides (chain B and chain D) are shown in cartoon and stick representations where carbon atoms are colored yellow.

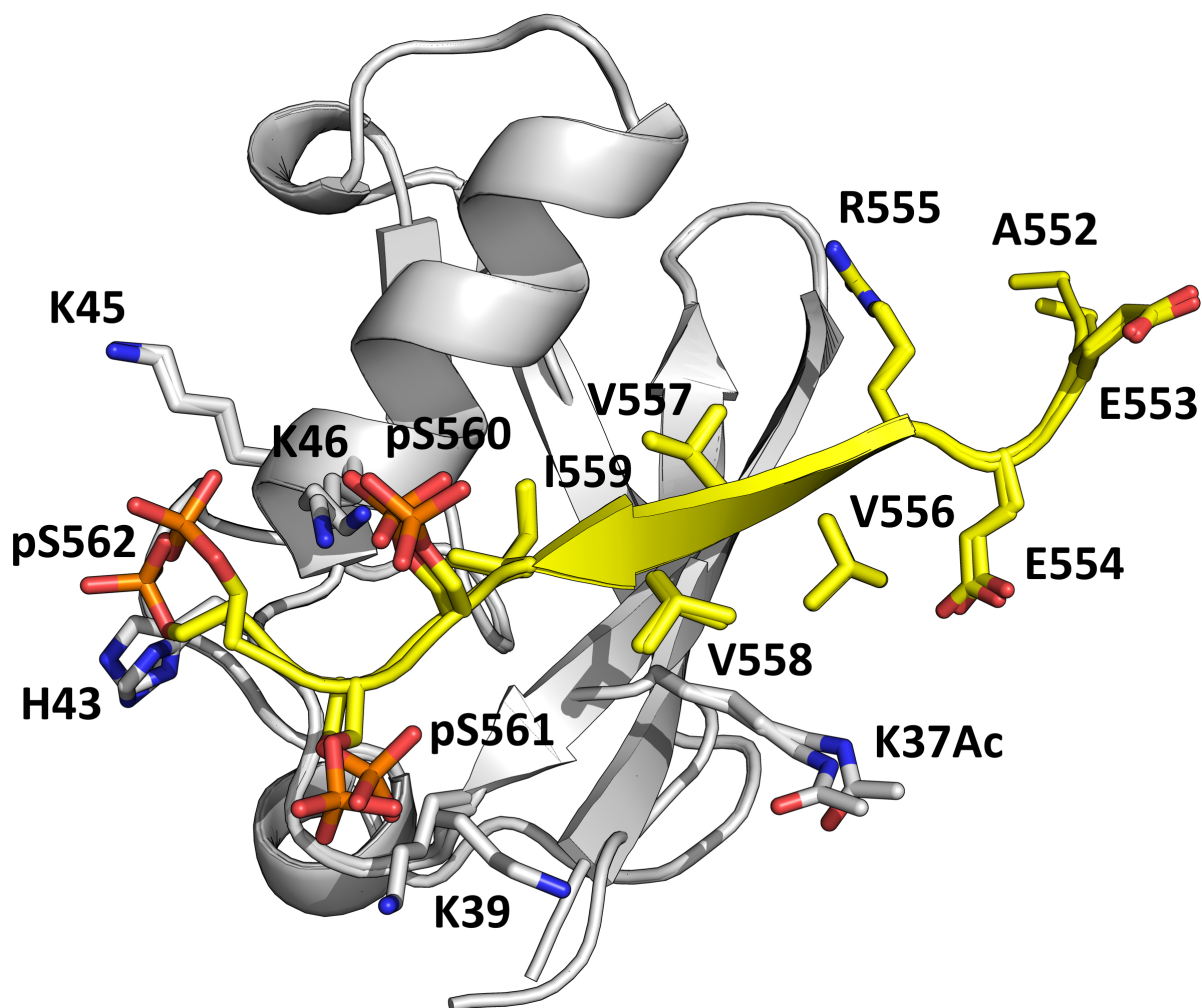


Figure S2. Overlay of the two Δ N-SUMO1-K37Ac:PML-SIM-4PO₄ complexes present in the asymmetric unit. The proteins are represented in cartoon in either grey (Δ N-SUMO1) or yellow (PML-SIM-4PO₄). The key residues of Δ N-SUMO1 (K37Ac, K39, H43, K45 and K46) as well as residues 552-562 of PML-SIM-4PO₄ are in stick representation.

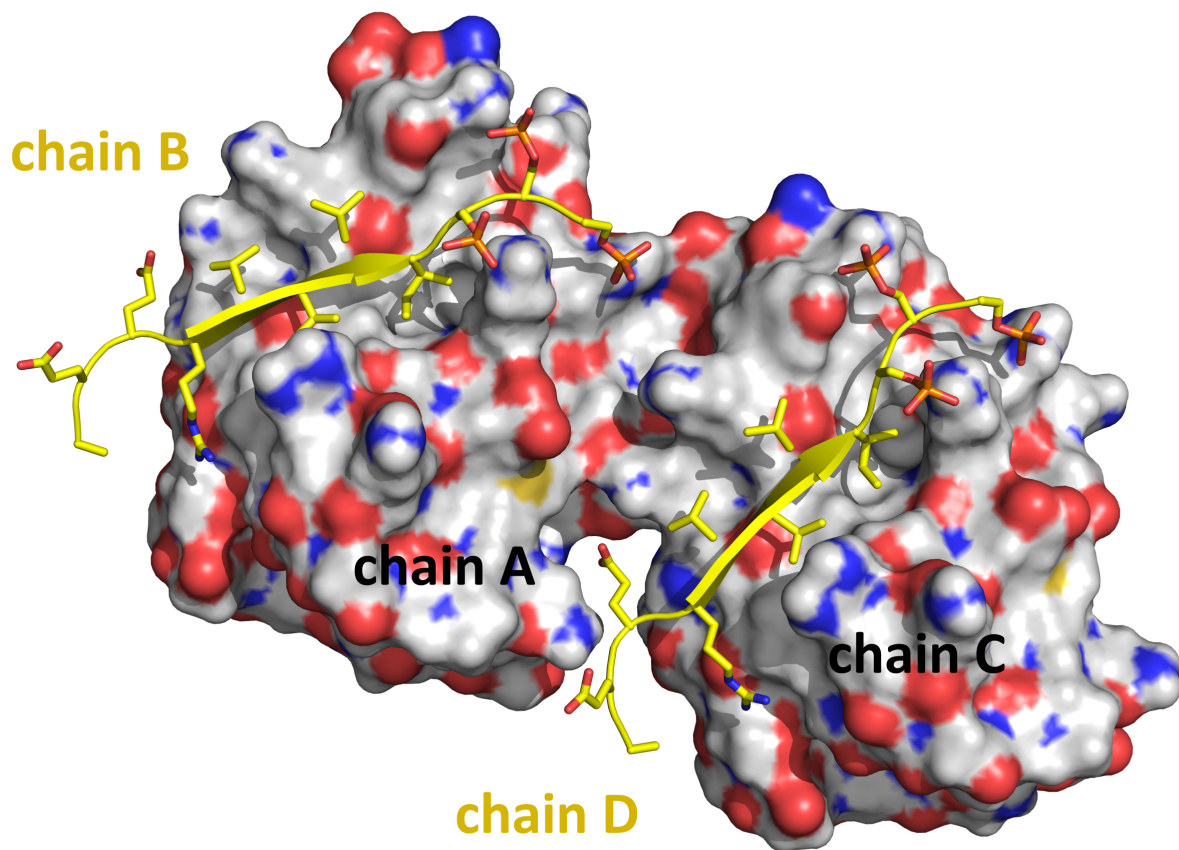


Figure S3. Structure of the two complexes present in the asymmetric unit of the Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ crystal. The SUMO1 proteins (chain A and chain C) are shown as a surface representation where carbon atoms are colored gray, sulphur atoms yellow, acidic residues red, and basic residues blue. The PML-SIM-4PO₄ peptides (chain B and chain D) are shown in cartoon and stick representations where carbon atoms are colored yellow.

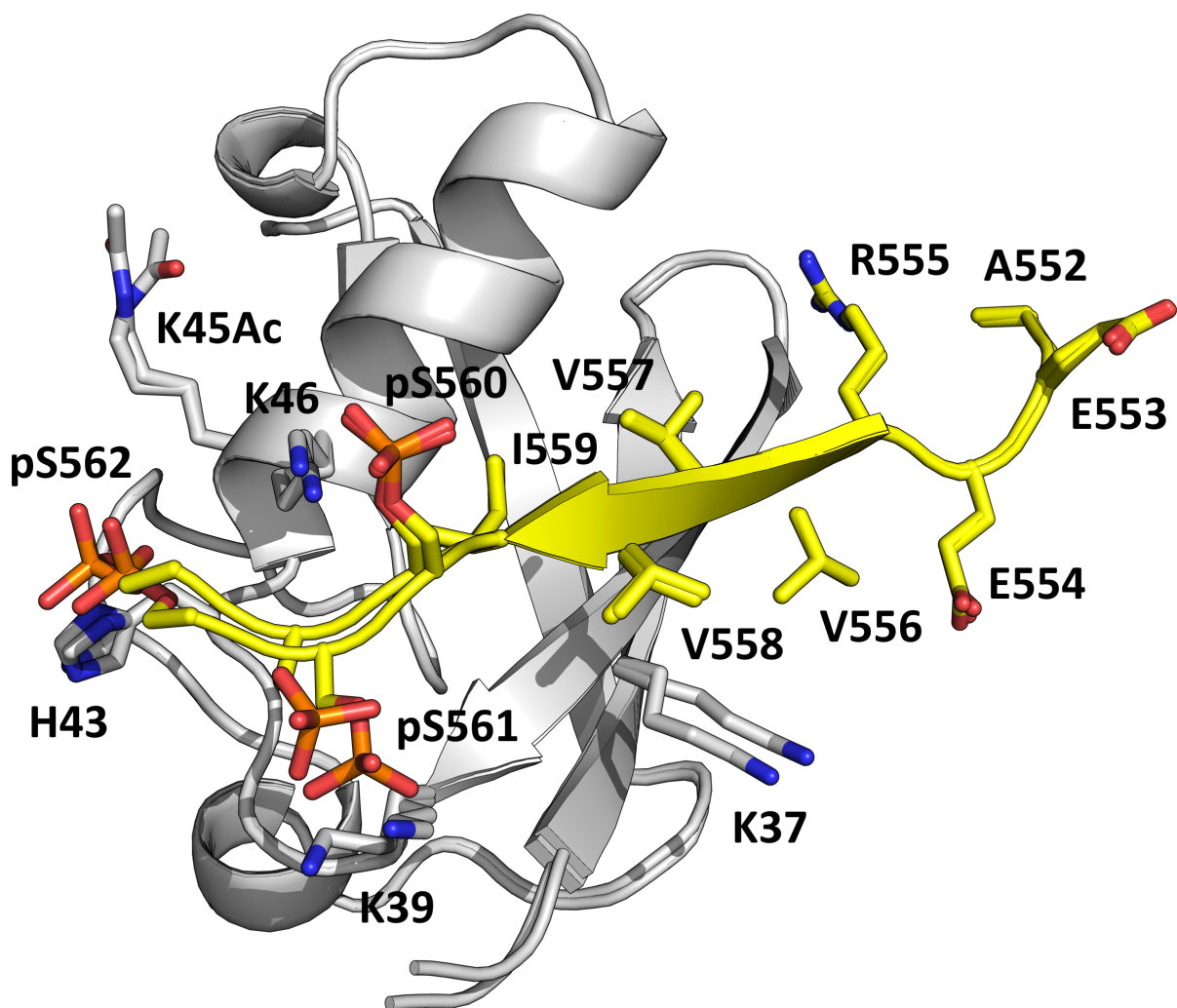


Figure S4. Overlay of the two Δ N-SUMO1-K45Ac:PML-SIM-4PO₄ complexes present in the asymmetric unit. The proteins are represented in cartoon in either grey (Δ N-SUMO1) or yellow (PML-SIM-4PO₄). The key residues of Δ N-SUMO1 (K37, K39, H43, K45Ac and K46) as well as residues 552-562 of PML-SIM-4PO₄ are in stick representation.

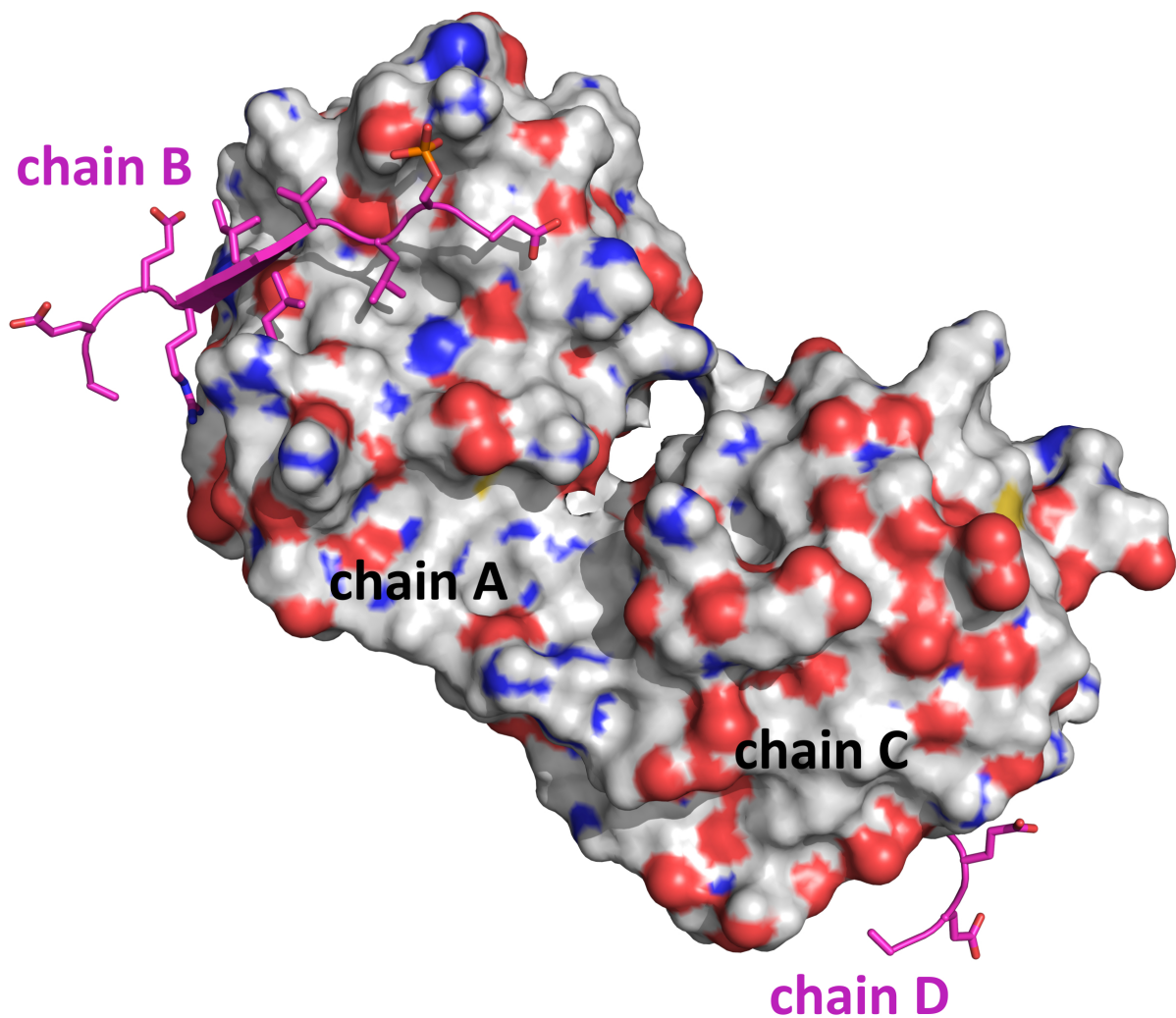


Figure S5. Structure of the two complexes present in the asymmetric unit of the Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ crystal. The SUMO1 proteins (chain A and chain C) are shown as a surface representation, where carbon atoms are colored gray, sulphur atoms yellow, acidic residues red, and basic residues blue. The mDaxx-SIM-2PO₄ peptides (chain B and chain D) are shown in cartoon and stick representations, where carbon atoms are colored magenta.

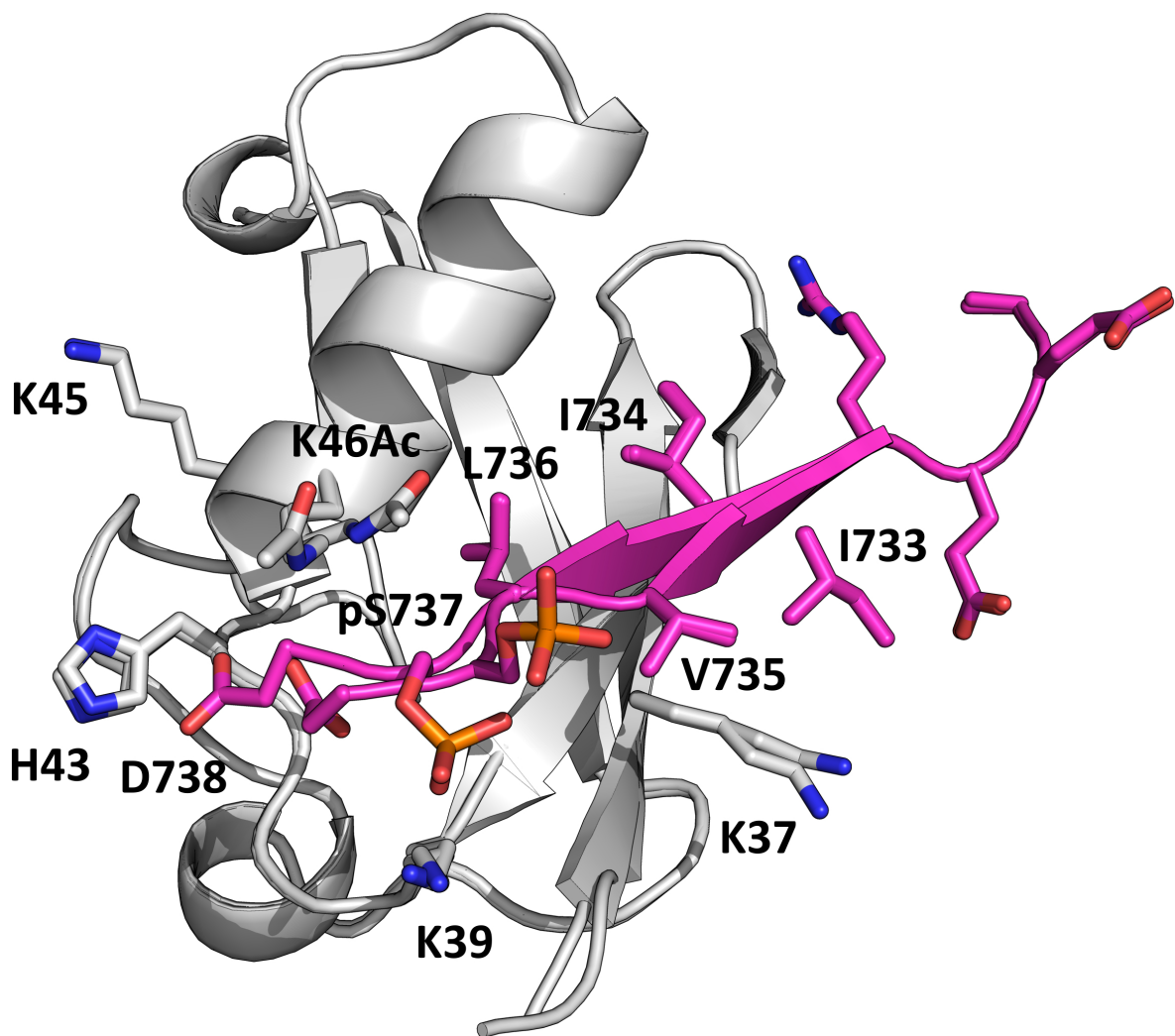


Figure S6. Overlay of the two Δ N-SUMO1-K46Ac:mDaxx-SIM-2PO₄ complexes present in the asymmetric unit. The proteins are represented in cartoon in either grey (Δ N-SUMO1) or magenta (mDaxx-SIM-2PO₄). The key residues of Δ N-SUMO1 (K37, K39, H43, K45 and K46Ac) as well as residues 733-738 of mDaxx-SIM-2PO₄ are in stick representation. Unlabeled residues at the N-terminus of mDaxx-SIM-2PO₄ correspond to residues 552-555 of PML-SIM.

