

**APPEL A PROJETS DE THÈSE 2026
CONCOURS DE L'ED 658 – Sciences du vivant**

Formulaire à compléter et déposer par les porteurs du projet de thèse HDR via l'adresse mail suivante :
ed658@univ-amu.fr OU/ET romane.haddadi@univ-amu.fr

MERCI DE METTRE EN SUJET DU MAIL : Concours ED 658 2026 porteur de projet

Ces dossiers seront publiés pour que les candidats postulent.

au plus tard le 3 mars 2026.

INTITULÉ DU PROJET

Étude *in silico* de l'activation des intégrines par O-glycosylation dans les carcinomes

In silico analysis of integrin activation by O-glycosylation in carcinomas

1- Directrice ou directeur de thèse HDR

NOM	Bard	Prénom	Frederic
Téléphone portable	06 35 50 62 00	Courriel	Frederic.bard@inserm.fr

Unité de recherche (UR) de rattachement

Code (acronyme et n°) de l'UR à AMU :	UM 105
Nom de l'UR :	CRCM

Directrice ou directeur de l'UR

Nom	Borg	Prénom	Jean-Paul
-----	------	--------	-----------

2- Codirectrice ou codirecteur de thèse, le cas échéant

NOM	Poli	Prénom	Giulio
Téléphone portable		Courriel	giulio.poli@unipi.it

Unité de recherche (UR) de rattachement

Code (acronyme et n°) de l'UR à AMU :	Italy
Nom de l'UR :	Molecular Modeling and Virtual Screening laboratory Department of Pharmacy University of Pisa

Directrice ou directeur de l'UR

Nom		Prénom	
-----	--	--------	--

SPECIALITE : (Rayer les mentions inutiles)

Bio-informatique et Génomique / Biochimie Structurale / Biologie du Cancer /

SOUS JURY POUR LE CONCOURS : (Rayer les mentions inutiles)

JURY 1 : BIOLOGIE CELLULAIRE

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

MERCI DE REMPLIR TOUTES LES INFORMATIONS EN FRANÇAIS ET EN ANGLAIS

Description de la problématique de recherche - Project description

La croissance tumorale et la formation de métastases dans les carcinomes dépendent de la capacité des cellules cancéreuses à activer leur adhésion à la matrice extracellulaire (MEC), un processus largement contrôlé par les intégrines, récepteurs transmembranaires de la MEC. Les mécanismes d'activation des intégrines dans les cellules cancéreuses restent à ce jour incomplètement élucidés. Notre laboratoire a récemment identifié un nouveau mécanisme d'activation des intégrines, contrôlé par la reprogrammation spatiale de la O-glycosylation de type GalNAc. [1,2,3]

La voie GALA, qui consiste en la relocalisation des enzymes de O-glycosylation GALNTs depuis l'appareil de Golgi vers le réticulum endoplasmique, est un processus observé dans plusieurs tumeurs solides, notamment les cancers du foie et du pancréas. Des analyses glycoprotéomiques récentes ont révélé que cette relocalisation entraîne une augmentation marquée de la O-glycosylation sur de multiples sites des intégrines. Les intégrines sont des hétérodimères composés de sous-unités alpha et bêta, formant des récepteurs à différents composants de la MEC. La sous-unité bêta1, l'une des plus exprimées, est capable de s'associer à de nombreuses sous-unités alpha. Nous observons la glycosylation de cette sous-unité bêta1 ainsi que des sous-unités alpha qui lui sont associées. [2,4]

En parallèle, nous observons une augmentation des propriétés adhésives des cellules après reprogrammation spatiale, accompagnée d'une activation des intégrines. Les cellules adhèrent plus rapidement à la fibronectine et migrent plus rapidement dans des tests de cicatrisation (*scratch wound assays*). Ce processus et ce profil de glycosylation sont conservés dans les tumeurs humaines et murines, suggérant un mécanisme hautement conservé au cours de l'évolution.

Les dimères d'intégrines peuvent adopter plusieurs états conformationnels : replié (inactif), étendu-fermé, et étendu-ouvert (actif) [5]. L'une des hypothèses est que la glycosylation modifie l'équilibre entre ces différentes conformations, et pourrait même induire de nouvelles conformations. Compte tenu du nombre de résidus glycosylés, de la présence de deux sous-unités et de la complexité des changements conformationnels des intégrines, il est difficile d'aborder cette question uniquement par des approches *in vitro*.

Ce projet vise à explorer, par des approches *in silico*, quelles modifications sont fonctionnellement déterminantes. Comprendre comment les O-glycanes portés par des résidus spécifiques modifient cet équilibre conformationnel à l'échelle atomique est en effet essentiel pour élucider la régulation de l'adhésion par la voie GALA. Les approches computationnelles sont particulièrement adaptées pour explorer ces effets de manière systématique. Ce projet sera complété par des travaux *in vitro* conduits par les autres membres de l'équipe.

Version anglaise :

Metastasis and tumor growth in carcinoma depend on the ability by cancer cells to activate adhesion to the extracellular matrix (ECM), a process highly dependent on the ECM receptor integrins. How integrins are activated in cancer cells remains incompletely understood. Recently, our lab has identified a novel mechanism of integrin activation controlled by the spatial reprogramming of O-GalNac glycosylation. [1,2,3]

**APPEL A PROJETS DE THÈSE 2026
CONCOURS DE L'ED 658 – Sciences du vivant**

The GALA pathway, that consists in the relocation from the Golgi to the endoplasmic reticulum of the O-glycosylation enzymes GALNTs, is a process occurring in several solid tumors such as liver and pancreatic cancer. Recent glycoproteomic analysis has revealed that this relocation results in a marked increase of O-glycosylation on multiple sites in integrins. Integrins are composed of dimers of alpha and beta subunits that combine to form receptors to various elements of the ECM. The integrin beta1 subunit is one of the most expressed and is able to combine with various alpha subunits. We observe glycosylation of this integrin beta1 subunit as well as alpha subunits that bind to beta1. [2,4]

In parallel, we observe an increase in the adhesive properties of cells after spatial reprogramming and an activation of integrins. Cells adhere faster on Fibronectin and migrate faster in scratch wound assays. The process and glycopattern is conserved in human and mouse tumors, suggesting a highly conserved mechanism. Integrins dimers can adopt multiple conformational states, bent (inactive), extended-closed, and extended-open (active) [5]. One hypothesis is that glycosylation affects the balance between these different conformations and might even induce new conformations. Given the number of residues glycosylated, the presence of two subunits and the complexity of integrin conformational changes, it is challenging to approach this question purely in vitro.

This project endeavour to explore predicting in silico which modifications are important. Thus, understanding how O-glycans on specific residues shift this conformational equilibrium at the atomic level is key to understand GALA regulation of adhesion. Computational approaches are particularly well suited to explore these effects systematically. This project will be complemented by in vitro work conducted by other members of the lab.

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

Mots clés – Keywords

O-glycosylation, Intégrines, MEC, Dynamique Moléculaire, PDAC, Pathway GALA, Docking Moléculaire, Bioinformatique Structurale, *Machine Learning*

O-glycosylation, Integrins, ECM, Molecular Dynamics, PDAC, Gala Pathway, Molecular Docking, Structural Bioinformatics, Machine Learning

Thématique / Domaine / Contexte - Theme / Field / Context

Les intégrines jouent un rôle central dans de nombreux processus biologiques, avec des implications en physiopathologie allant du développement embryonnaire à l'inflammation et au cancer. Notre projet propose d'explorer un nouvel aspect de la régulation des intégrines : l'addition régulée de glycanes. Pour développer ce nouveau concept, nous nous appuierons sur la littérature abondante et les connaissances accumulées en biologie des intégrines.

Ce projet s'articule à l'interface de la bioinformatique structurale basée sur l'IA, de la glycobiologie et de la biologie cellulaire. Il mobilisera des approches de biologie structurale computationnelle : incluant des simulations de dynamique moléculaire tout-atome, le docking moléculaire et l'apprentissage automatique : afin d'étudier comment la O-glycosylation induite par la voie GALA modifie la structure des intégrines, leur dynamique conformationnelle et leur liaison aux ligands.

Les simulations de dynamique moléculaire ont démontré leur efficacité pour étudier l'impact des modifications post-traductionnelles sur la conformation et la dynamique des protéines [6,7]. Les avancées récentes dans les outils de modélisation des glycoprotéines (GlycoSHAPE, CHARMM-GUI Glycan Modeler) permettent désormais de construire et de simuler des structures protéiques glycosylées réalistes [8,9]. L'apprentissage automatique appliqué aux données de simulation permettra ensuite de hiérarchiser les sites fonctionnellement les plus pertinents [10], en fournissant des prédictions testables pour la validation expérimentale par l'équipe. Ces travaux constitueront un appui précieux pour guider les expériences menées au sein du laboratoire par un post-doctorant et un ingénieur de recherche.

Integrins play many roles, with implications in physiopathology, ranging from development to inflammation and cancer. Our project proposes to explore a new aspect of integrin regulation: by the regulated addition of glycan. To develop this new concept, we will leverage on the abundant literature and knowledge accumulated on integrin biology.

**APPEL A PROJETS DE THÈSE 2026
CONCOURS DE L'ED 658 – Sciences du vivant**

This project is articulated at the interface of structural bioinformatics using AI, glyco- and cell biology and it will use computational structural biology approaches, including all-atom molecular dynamics simulations, molecular docking and machine learning, to investigate how GALA-induced O-glycosylation alters integrin structure, conformational dynamics and ligand binding.

Molecular dynamics simulations have proven effective for studying how post-translational modifications alter protein conformation and dynamics [6,7], and recent advances in glycoprotein modeling tools (GlycoSHAPE, CHARMM-GUI Glycan Modeler) now make it possible to build and simulate realistic glycosylated protein structures [8,9]. Machine learning applied to the simulation data will then prioritize the functionally most relevant sites [10], providing testable predictions for experimental validation by the team. This work will be greatly helpful to guide the experimental work in the lab pursued by a postdoc and a research engineer.

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

Questions principales du projet

- Pourquoi les sous-unités d'intégrines sont-elles davantage glycosylées lors de la relocalisation des GALNTs vers le réticulum endoplasmique ?
- Comment les glycanes modifient-ils la fonction des intégrines ?

Objectifs

- Cartographie des O-glycanes sur les structures 3D des sous-unités individuelles et des dimères
- Analyse de l'effet des O-glycanes sur les monomères et les dimères par simulation de dynamique moléculaire
- Effet des O-glycanes sur la formation des dimères et leur adoption de différents états conformationnels
- Effet des O-glycanes sur la liaison à la fibronectine par docking moléculaire

The two main questions of the project are:

- Why are integrins subunits more glycosylated upon relocation of GalNTs to the ER?
- How do glycans alter integrin function?

Objectifs :

- Mapping O-glycans on 3D structures of individual subunits and of the dimers
- Analysis of the effect of O-glycans on the monomers and dimers by Molecular dynamics
- Effect of O-glycans on the formation of dimers and their adoption of various conformation
- Effect of O-glycans on binding of fibronectin by Molecular Docking

Méthode – Method

Modélisation structurale → Simulations de DM → Analyse structurale → Amarrage avec les ligands → ML pour identifier les sites de glycosylation les plus pertinents → Validation expérimentale en laboratoire

Ce projet combinera des approches computationnelles complémentaires selon le pipeline suivant :

1. Modélisation structurale: Modélisation par homologie (RoseTTAFold) et/ou prédiction de structure par IA (AlphaFold2, Boltz-2) pour obtenir ou générer les modèles 3D initiaux des sous-unités et dimères d'intégrines, glycosylés et non glycosylés.
2. Préparation des systèmes glycoprotéiques: Reconstruction des liaisons O-glycane–protéine à l'aide de CHARMM-GUI Glycan Modeler avec les paramètres validés du champ de force CHARMM36m, afin de générer des systèmes glycosylés prêts pour la simulation.
3. Simulations de dynamique moléculaire tout-atome: Échantillonnage conformationnel et analyse de la dynamique des systèmes glycosylés et non glycosylés avec GROMACS ou AMBER, permettant une comparaison directe de leur comportement structural.
4. Analyse structurale des trajectoires de DM: Analyse systématique incluant le RMSD (stabilité structurale), le RMSF (flexibilité par résidu), l'analyse en composantes principales (ACP), les réseaux de liaisons

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

hydrogène, la surface accessible au solvant (SASA) et les cartes de contacts, afin de caractériser l'effet des O-glycanes sur la conformation et la dynamique des intégrines.

5. Docking moléculaire: Docking protéine–protéine des intégrines glycosylées et non glycosylées avec la fibronectine à l'aide de HADDOCK, guidé par les états conformationnels identifiés à l'étape 4, suivi d'un affinement par DM des complexes les mieux classés et d'une estimation de l'énergie libre de liaison (MM-PBSA/GBSA).
6. Apprentissage automatique: Hiérarchisation des sites de glycosylation fonctionnellement pertinents: Des descripteurs moléculaires extraits des trajectoires de DM et des résultats de docking seront utilisés pour entraîner des modèles prédictifs (forêt aléatoire, régression logistique) reliant des profils spécifiques de O-glycosylation aux modifications de la dynamique conformationnelle des intégrines et de leur affinité pour la fibronectine, permettant une hiérarchisation systématique des sites les plus pertinents sur le plan fonctionnel.

Les prédictions computationnelles seront confrontées aux données expérimentales d'adhésion et de migration cellulaires générées par le laboratoire.

Structural modeling -> MD Simulations -> Structural analysis -> Docking with ligands -> ML to identify more relevant glycosylation site -> In lab experimental validation

This project will combine complementary computational approaches in the following pipeline:

1. **Structural Modeling** : Template-based modeling (RoseTTAFold) and/or AI-based structure prediction (AlphaFold2, Boltz-2) to retrieve or generate initial 3D models of glycosylated and non-glycosylated integrin subunits and dimers.
1. **Glycoprotein System Preparation** : Reconstruction of O-glycan–protein linkages using CHARMM-GUI Glycan Modeler with validated CHARMM36m force field parameters, generating simulation-ready glycosylated systems.
1. **All-Atom Molecular Dynamics Simulations** : Conformational sampling and dynamics analysis of glycosylated and non-glycosylated systems using GROMACS or AMBER, enabling direct comparison of their structural behavior.
1. **Structural Analysis of MD Trajectories** : Systematic analysis including RMSD (structural stability), RMSF (per-residue flexibility), principal component analysis (PCA), hydrogen bond networks, solvent-accessible surface area (SASA), and contact maps to characterize the effect of O-glycans on integrin conformation and dynamics.
1. **Molecular Docking**: Protein–protein docking of glycosylated and non-glycosylated integrins with fibronectin using HADDOCK, informed by the conformational states identified in step 4, followed by MD-based refinement of top-ranked complexes and binding free energy estimation (MM-PBSA/GBSA).
1. **Machine Learning** : **Prioritization of Functionally Relevant Glycosylation Sites** — Molecular descriptors extracted from MD trajectories and docking results will be used to train predictive models (random forest,

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

logistic regression) linking specific O-glycosylation patterns to changes in integrin conformational dynamics and fibronectin binding affinity, enabling systematic prioritization of the most functionally relevant sites.

Computational results will be compared with experimental adhesion and migration generated by the lab. Machine learning-based prediction of glycosylation impact: molecular descriptors extracted from MD trajectories will be used to build a predictive model that links specific O-glycosylation patterns to integrin-fibronectin binding affinity changes.

Références bibliographiques

- [1] Hynes RO. Integrins: bidirectional, allosteric signaling machines. *Cell*. 2002 Sep 20;110(6):673-87.
- [2] Nguyen AT, Chia J, Ros M, et al. Organelle specific O-glycosylation drives MMP14 activation, tumor growth and metastasis. *Cancer Cell*. 2017;32(5):639–653.
- [3] Gill DJ, Chia J, Senewiratne J, Bard F. Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. *J Cell Biol*. 2010;189(5):843–858.
- [4] Tran LS, Chia J, Le Guezennec X, et al. ER O-glycosylation in synovial fibroblasts drives cartilage degradation. *Nat Commun*. 2025;16:2535.
- [5] Li J, Jo MH, Yan J, et al. Ligand binding initiates single-molecule integrin conformational activation. *Cell*. 2024;187(12):2990–3005.e17
- [6] Mehrbod M, et al. Integrin $\alpha\text{IIb}\beta\text{3}$ intermediates: From molecular dynamics to adhesion assembly. *Biophys J*. 2023;122(3):392–406.
- [7] Lee HS, Qi Y, Im W. Effects of N-glycosylation on protein conformation and dynamics: PDB analysis and MD simulation study. *Sci Rep*. 2015;5:8926.
- [8] Iacob RE, et al. Restoring protein glycosylation with GlycoShape. *Nat Methods*. 2024;21:2117–2127.
- [9] Park SJ, Lee J, Qi Y, et al. CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology*. 2019;29(4):320–331.
- [10] Brownless ALR, Rheaume E, Kuo KM, et al. Using Machine Learning to Analyze Molecular Dynamics Simulations of Biomolecules. *J Phys Chem B*. 2025;129(22):5375–5385.

Conditions scientifiques matérielles et financières du projet de recherche - Scientific, material, and financial conditions of the research project

Le groupe de Frédéric Bard est soutenu par plusieurs financements, dont une subvention Amidex et une bourse Leader en Oncologie de la Fondation ARC. Une ingénieure en bioinformatique, Eugénie Lohmann, sera en mesure d'accompagner les travaux de l'étudiant(e) en thèse. Par ailleurs, le CRCM dispose d'une plateforme bioinformatique solide, offrant un accès à des ressources de calcul suffisantes. Le groupe de Xavier Morelli et Yves Colette au CRCM utilise divers outils bioinformatiques et sera en mesure de guider l'étudiant(e) tout au long du projet. Par ailleurs, Dr. Giulio Poliu, de l'Université de Pavia est un spécialiste des études par molecular dynamics et a accepté de co-diriger la thèse.

The group of Frederic Bard is supported by various grants from Amidex and a Leader in Oncology grant from the ARC foundation. A bioinformatics engineer, Eugenie Lohman, will be able to support the work of the PhD student.

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

In addition, there is a strong bioinformatics platform at the CRCM, with access to sufficient computing resources. The group of Xavier Morelli and Yves Colette at the CRCM uses various bioinformatics tools and will be able to guide the student. In addition, Giulio Poli, a specialist in molecular dynamics, is collaborating with us and has accepted to co-supervise the thesis.

Profil et compétences recherchées - Profile and skills required

Le/la candidat(e) idéal(e) devra être titulaire d'un Master (ou équivalent) en bioinformatique structurale, biologie computationnelle, biophysique ou dans un domaine connexe. Une expérience préliminaire en modélisation de glycoprotéines (par exemple lors d'un stage de Master) constitue un atout. Le pipeline du projet implique la modélisation structurale de glycoprotéines, des simulations de dynamique moléculaire tout-atome, le docking moléculaire et l'analyse par apprentissage automatique de trajectoires de DM. Les compétences suivantes sont donc requises ou fortement souhaitées :

- Solide formation en biologie structurale et/ou biologie computationnelle, avec une bonne maîtrise de la structure 3D des protéines, des champs de force et des principes de modélisation moléculaire.
- Capacité à interpréter de manière critique les métriques de confiance des modèles IA (pLDDT, PAE, ipTM/pTM).
- Une expérience dans le déploiement et l'utilisation d'outils de prédiction de structure par IA (Chai-1, Boltz-2) serait un plus.
- Expérience des logiciels de dynamique moléculaire (GROMACS, NAMD ou AMBER) et des outils de préparation de systèmes glycoprotéiques (CHARMM-GUI, GlycoSHAPE).
- Familiarisé avec les approches de docking protéine-protéine (HADDOCK, ClusPro) et les méthodes d'estimation de l'énergie libre de liaison (MM-PBSA/GBSA).
- Maîtrise de Python (MDAnalysis, NumPy/SciPy) et Bash, incluant la soumission de jobs sur clusters HPC (SLURM).
- Notions d'apprentissage automatique avec scikit-learn (forêt aléatoire, régression logistique, importance des variables).
- Engagé(e) dans une science ouverte et reproductible : application des principes FAIR (Trouvable, Accessible, Interopérable, Réutilisable) à la gestion des données et du code.
- Maîtrise de Git/GitLab pour un suivi rigoureux des versions, facilitant la collaboration et la traçabilité des analyses.
- Familiarité avec les outils de visualisation moléculaire (PyMOL, VMD, ChimeraX) et les environnements Linux/HPC.
- Des connaissances en glycobiologie, modifications post-traductionnelles ou biologie du cancer constituent un atout supplémentaire.

APPEL A PROJETS DE THÈSE 2026 CONCOURS DE L'ED 658 – Sciences du vivant

- Intérêt pour le travail à l'interface entre recherche computationnelle et expérimentale, avec une volonté d'interagir étroitement avec l'équipe expérimentale GLYNThERA pour la validation des prédictions computationnelles.
- Bonne maîtrise de l'anglais (oral et écrit) ; le français n'est pas obligatoire mais apprécié.
- Pratique d'une veille scientifique régulière sur les dernières avancées en IA, bioinformatique, glycobiologie et biologie cellulaire (alertes PubMed, bioRxiv, flux de littérature ciblés).

The ideal candidate should hold a Master's degree (or equivalent) in structural bioinformatics, computational biology, biophysics, or a related field. Preliminary experience in glycoprotein modeling (e.g. during a Master internship) is an advantage. The project pipeline involves glycoprotein structural modeling, all-atom MD simulations, molecular docking, and machine learning-based analysis of MD trajectories. The following skills are therefore required or highly desirable:

- Strong background in structural biology and/or computational chemistry, with knowledge of protein 3D structure, force fields, and molecular modeling principles.
- Able to critically interpret AI model confidence metrics.
- Experienced in deploying and running AI-based protein structure prediction tools (Chai-1, Boltz-2) would be a plus.
- Experience with molecular dynamics simulation software (GROMACS, NAMD, or AMBER) and glycoprotein system preparation tools (CHARMM-GUI, GlycoSHAPE).
- Familiarity with protein-protein docking approaches (HADDOCK, ClusPro) and binding free energy estimation methods (MM-PBSA/GBSA).
- Proficiency in scripting languages, including Python (MDAnalysis, NumPy/SciPy) and Bash, with experience in writing and submitting job scripts on HPC clusters (SLURM).
- Applied machine learning using scikit-learn, including classification models (random forest, logistic regression) and feature importance analysis for the prioritization of biologically relevant features.
- Committed to open and reproducible science: applies FAIR principles (Findable, Accessible, Interoperable, Reusable) to data and code management, ensuring long-term accessibility and reuse of research outputs.
- Proficient in version control using Git/GitLab, enabling collaborative, traceable, and reproducible research workflows.
- Familiarity with molecular visualization tools (PyMOL, VMD, ChimeraX) and Linux/HPC environments.
- Knowledge of glycobiology, post-translational modifications, or cancer biology is a plus.
- Interest in working at the interface between computational and experimental research, with willingness to interact closely with the wet-lab GLYNThERA team for experimental validation of computational predictions.
- Good command of English (spoken and written); French is not mandatory but appreciated.
- Conducts regular scientific monitoring to stay current with the latest advances in AI, bioinformatics, glycobiology, and cell biology, leveraging tools such as PubMed alerts, bioRxiv, and curated literature feeds.