New mass spectrometry (MS)-based approaches for the characterization of protein-drug-conjugates (PDCs) at intact and middle-levels

Rania Benazza

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École doctorale
Sciences chimiques | ED 222
Université de Strasbourg

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

IPHC – UMR 7178

THÈSE présentée par :

Rania BENAZZA

soutenue le : **01 décembre 2023**

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

Discipline/ Spécialité : Chimie analytique

Nouvelles approches de spectrométrie de masse pour la caractérisation de protéines thérapeutiques bioconjuguées

THÈSE dirigée par :

Dr. CIANFERANI Sarah Directrice de recherche, Université de Strasbourg

RAPPORTEURS :

Dr. VILASECA Marta Associate Professor, Barcelone IRB, Spain
Dr. GUILLARME Davy Associate Professor, University of Geneva, Switzerland

---

AUTRES MEMBRES DU JURY :

Dr. BURLET-SCHILTZ Odile Directrice de recherche, Université de Toulouse

MEMBRES INVITES

Dr. HERNANDEZ-ALBA Oscar Chargé de recherche, Université de Strasbourg
Dr. CHAUBET Guilhem Chargé de recherche, Université de Strasbourg
Dr. FRANCOIS Yannis Maître de conférence, Université de Strasbourg

À mes parents,

À mes sœurs,

À Rita,

"Knowing is better than wondering, waking is better than sleeping, and even the biggest failure, even the worst, beats the hell out of never trying"

Meredith Grey

Remerciements

Ce travail de thèse a été réalisé au sein du Laboratoire de Spectrométrie de Masse Bio-Organique (LSMBO) de l'institut Pluridisciplinaire Hubert Curien (IPHC) de Strasbourg au Département des Sciences Analytiques, UMR7178. Ce projet a reçu un financement du programme de recherche et d'innovation horizon 2020 de l'union européenne, dans le cadre de la convention de subvention d'action Marie Skłodowska Curie n° 859458.

Tout d'abord je tiens à remercier Sarah d'avoir cru en moi même à 2 316,9 km et de m'avoir accepté parmi l'équipe du LSMBO pour y réaliser ma thèse. Grâce à ton encadrement, j'ai appris comment toujours raconter une histoire avec un fil conducteur, j'ai appris à être persévérante, et à vouloir toujours faire mieux, et je te remercie pour ça! C'était un plaisir de partager avec toi ma passion pour la native et pour les mAbs, mais pas seulement… écouter du Coldplay ça c'est le vrai bonheur !

J'aimerai remercier Oscar aussi pour son encadrement de thèse, muchas gracias Chef! C'était un honneur d'être ta première étudiantE et un vrai plaisir de travailler avec toi. Maintenant le chiffre 3 ne va pas te rappeler que le match Maroc-Espagne de la coupe du monde (3-0). 3 c'est aussi le nombre d'années où tu n'as pas cessé de m'apprendre plein de choses, 3 c'est le nombre de fois que je venais te voir par jour et à chaque fois tu avais le temps pour moi, 3 est le nombre de mails/messages que je t'envoie par heure pour chercher une référence et à chaque fois tu l'as! Je te remercie plus que 3 fois pour tout ça! Et je te dis à 2030 pour la remontada ;)

Mes remerciements les plus sincères à Marta Vilaseca, Davy Guillarme, Odile Burlet-Schiltz, Yannis François et Guilhem Chaubet d'avoir accepté d'évaluer mon travail de thèse.

Je tiens à remercie les membres de Targeted Anti-Cancer Therapies (TACT) program avec qui j'ai travaillé pendant ces trois années lors des différentes collaborations et différents tact meeting! En particulier Guilhem qui nous a réuni et Graham qui m'a accueilli pour un stage chez à Almac. Merci aussi à tous les ESRs et en particulier Ilias qui m'a beaucoup aidé et qui m'a expliqué la chimie des ADC en détails, nous avons partagé plein de moments sympa tous ensemble dans tous les coins du monde. Thank you, merci, eucharisto, grazie, tessekurler, obrigado, dzieki, dhanyavaad !

Un grand merci à mes collaborateurs biologistes et biochimistes avec qui j'ai appris de nouveaux aspects de la science autre que analytiques, et pour les discussions riches et intéressantes.

Je tiens à te remercier Christine C. pour ta gentillesse et ta bienveillance, c'est grâce à notre première visio que je suis là aujourd'hui! Merci à toi Christine S., ton dynamisme et ta bonne ambiance m'ont marqué à jamais, et j'espère que l'année prochaine on fera la strasbourgeoise ensemble ;)

Merci aux deux fondateurs du club café matinal; Fabrice B. et Martine. Fabrice, bon courage pour tes prochains voyages dans les trois coins du monde, quand tu viendras au quatrième je te passerai des échantillons de dromadaires promis! Martine, merci pour les bonbons et pour les bouquets de menthe, tu as su comment rendre une marocaine vraiment heureuse ! Je sais que j'étais pénible pour les missions, mais heureusement que j'ai arrêté les congrès et les voyages TACT avant le passage vers Notilus…

Un merci à Laurence qui m'a accueilli parmi ses étudiants ECPM pendant quelques jours, c'est toujours intéressant de revenir à la théorie. Je remercie également Magali, Agnès, François, Véronique, Raymond et Delphine pour leur gentillesse et leur aide.

Merci aux deux informaticiens qui m'ont sauvé plusieurs fois, Fabrice V. et Alex, grâce à vous et après 3 ans j'ai compris qu'il faut juste redémarrer la session maintenant, et que MassLynx ne fonctionne pas mieux si je clique 5 fois.

JMS, merci à toi qui a toujours trouvé des solutions à n'importe quel problème technique, merci de m'avoir appris comment toucher aux machines même si des fois quand j'y touche c'est encore pire… Merci également de nous avoir bien nourri aux madeleines et à la charlotte (alcoolisé à la gélatine de porc MIAMMM). Tu n'as pas réussi à me faire pleurer une seule fois, mais je te souhaite bon courage avec les nouveaux ! C'est sûr qu'un jour tu réaliseras ce rêve…

Merci Hélène pour ta gentillesse et ta sympathie, et surtout merci d'avoir fait mes analyses peptide mapping et de m'avoir appris à le faire, grâce à toi je rajoute un point en plus de mon CV !

Un grand merci à toi Stéph, toi qui m'a appris énormément de choses pendant ces trois années, du LCT jusqu'au BioAccord. Une grande partie de cette thèse a été réussite grâce à ton aide et à tes conseils! Merci aussi pour les belles discussions autour d'un café ou autour d'un mail d'un collaborateur qu'on ne comprends pas… Maintenant je laisse mon fils à toi, c'est sûr qu'il n'y a pas mieux que toi pour prendre soin de mon petit BioAcc <3

Un merci à mes collègues du R2, surtout le bureau du fond (bdf), désolée de pas avoir été discrète pendant mes sessions karaoké devant le BioAccord! Bastien, j'espère que ma carte du ru sera toujours utile même avec un tarif de 9euro/repas, Valériane BRAVO pour le poste et bon courage pour la suite, Charline bon courage pour ta nouvelle aventure, profite à fond et n'oublie pas de prendre des vacances!!! Nathoun, c'était un plaisir de t'avoir parmi nous et d'avoir partagé le même avis sur avatar (c'était bien la sieste au ciné). Adrien bravo t'as enfin réalisé ton rêve de rejoindre le bdf, fais honneur à mes initiales et travaille bien sur le BR! Pauline, merci pour ta bonne humeur et pour la surprise que tu m'as fait avec Marie G., vous êtes sans doute mes swifties préférées ! Profite de ton concert Like we were in Pariiiiis!

Un merci à Noelia et Jeewan qui m'ont toujours accueilli avec plaisir dans le bureau du gouter, des fois pendant des heures et des heures, vous devez être fiers de vous car vous êtes des bons parents pour les poissons, je le sais car chaque weekend quand je viens ils sont déjà nourris haha.

Pour mes voisins du R5 au bureau de la porte, Reiko, Sarahi, Vittoria, Turkan et Zahraa, merci pour votre gentillesse. Merci aussi de m'avoir supporté à chaque fois que je venais discuter dans le bureau, et de ne pas m'avoir viré!

Merci aux gens que j'ai connu à travers Compomics même si je n'ai pas vraiment travaillé avec eux. Merci particulièrement à Arthur et Jasper, de m'avoir appris le Bivouac même si tout ce que je voulais apprendre était le AI…

Biensur merci à ma famille Supra <3 ces trois années étaient magnifiques à vos côtés, nous avons toujours su transférer les connaissances en MS structural ainsi que les bonnes habitues (le Uno, les croissants du

petit déj, les restaus supra…). Evo, nos discussions nocturnes et nos balades vers le bus G vont me manquer. Depuis que t'es partie chez les hollandais ya plus personne pour m'accompagner au mcdo, jpp de ces supra healthy… Marie L., merci de m'avoir cédé la meilleure place au bureau, une place qui vient avec une bouillotte en plus! Bon courage pour tes nouvelles aventures à Paris. Sarah N. et Léa bienvenu au club du top-down, avec l'aide de Oscar c'est sûr que vous formerez un nouveau trio de choc ! Sarah D., merci de m'avoir présenté à la vraie cuisine française même si t'es à moitié américaine, le boeuf bourguignon halal était un pur délice! Bon courage pour ta thèse et à la prochaine à Marseille ;). Merci à Oumayma et Christelle qui ont laissé une bonne trace au bureau supra!

A mes protéomistes préférées, Marie G., ma miss Alsaciana! Merci pour ta bonne ambiance et d'avoir mis des paillettes dans ma vie, j'espère qu'Aria et toi garderons de bons souvenirs de moi, jusqu'à notre prochaine aventure à Lyon for the Eras tour babyyyyy! Chachou, merci pour ces discussions profondes après 18h, tu es une personne incroyable et tu mérites que du bonheur! Merci de m'avoir accueilli à PasParis avec Titou, c'est sûr que je reviendrai…

Et à ma Dream Team: Alex, Hugo, Corentin et Jérôme mes gars sûrs ! Sans vous, cette thèse n'aurait pas été la même. Merci pour ces bons repas, ces bonnes discussions autour d'un mojito, d'un bon repas, ou d'un bon café/thé au labo en jours de semaine comme en weekend, aux horaires de travail comme pendant les nocturnes au labo! En particulier merci à mes deux tiktokeurs et mes deux chauffeurs préférés: Alex, courage pour la suite de la thèse, ton travail dur finira par bien payer j'en suis sûre ☺ par contre arrête le tiktok, t'abuses nan? Hugo, merci d'avoir apprécié le rap marocain tektek comme toto, t'as le vrai oreille musicale tmtc ;) En plus du rap, t'es un excellent guitariste et j'ai adoré t'accompagner en chantant, même si tu veux pas qu'on fasse le duo de la soirée de Noel ☹ Toi et Mamé vous avez été des voisins excellents, j'ai bien kiffé la kageneck life à coté de vous!

Corentin et Jérôme, on est arrivé le même jour mais aujourd'hui et après 3 ans nos chemins se séparent et je vous souhaite que du courage pour vos prochaines aventures successives! Coco, même si on était toujours en train de se taquiner, tu sais que qui aime bien châtie bien ;) Merci d'avoir été toujours à mes côtés, et de m'avoir appris beaucoup de choses en top-down, en course, sur Excel et jusqu'à l'impression de cette thèse, toujours serviable et jamais râleur (mhmm) ! Dans deux jour tu es Docteur, et je suis fière de t'avoir accompagné jusque-là! ☺. GG, toi qui m'a supporté tout au long de ces trois années, chaque jour littéralement. La thèse était beaucoup plus facile à tes côtés, tu m'as toujours encouragé et entouré de ta positivité. Tu as toujours su quoi dire et quand le dire, et grâce à ça j'ai survécu beaucoup de moments difficiles. Je te souhaite un très bon courage pour la fin de cette thèse, ça y'est dans quelques mois c'est toi qu'on fête, et je suis déjà très fière de toi ☺.

Et enfin, à ma petite Hirsh! Un grand merci à toi Aurélie, qui a été une des meilleures connaissances pendant ces années de thèse, tu m'as bourré de tendresse et de sympathie. Tu m'as aidé à surmonter beaucoup de difficultés tout au long de cette épreuve pas facile. Avec tes petites attentions de chocolat chaud, bouquets de fleurs, Ragusa et encore plus, tu as toujours réussi à me redonner le sourire. Hâte de fêter ton VAE ou ton mariage… mais dépêche toi t'as 1 an faut que je fête ça avant la fin de mon visa…

Merci également à mes amis qui m'ont supporté de loin comme de proches et particulièrement à Leila, merci de m'avoir transmis ta passion pour la chimie analytique et merci surtout de m'avoir supporté pendant plusieurs années, j'aurai aimé te voir ici avec Amine, mais on fêtera ça bien à Casablanca <3

Merci à Imane, ma copine qui m'a soutenu à distance, surtout merci pour tes appels réconfortant et de m'avoir entendu pendant des heures et des heures! Merci à Soha, qui m'a soutenue depuis l'enfance, bravo pour ton doctorat et hâte qu'on fête toute les deux ensembles. Merci à Justyna qui m'a soutenue du master jusqu'à aujourd'hui, tu as été ma source de soutien direct en France, hâte pour nos prochains voyages après la fin de cette thèse ;)

Et enfin merci aux personnes qui grâce à qui j'ai réussi à en arriver jusqu'ici aujourd'hui, mes parents et mes sœurs qui m'ont soutenu même malgré la distance. Baba, Mama, merci d'avoir toujours cru en moi, de m'avoir permis de vivre pleinement mon rêve, de m'avoir entouré d'amour et d'encouragement pour que je puisse arriver jusque-là. Merci à vous ! Merci également à mes sœurs; merci Sousou qui m'a toujours encouragé, merci également à mon beau-frère Faouzi pour son soutien et sa positivité, merci à vous deux de m'avoir fait le plus beau cadeau pendant ma thèse, la naissance de Rita qui m'a boosté encore plus pour donner le meilleur de moi, et une raison de sourire tous les jours comme une débile devant mon téléphone. Chay, merci pour tes séances d'aides psychologiques, tu as toujours su me surmonter le moral et m'aider à voir la grande image, bon courage pour ta thèse c'est pas facile, mais regarde la prévue, tout fini par prendre forme ;). Hibouch, merci de m'avoir donné le sourire, d'avoir été ma source de Bonheur depuis 1998 <3 La thèse loin de vous avait encore une dimension de difficulté en plus, mais j'ai senti toujours votre présence malgré la distance <3

Résumé

Nouvelles approches de spectrométrie de masse pour la caractérisation de protéines thérapeutiques bioconjuguées

Thèse soutenue par: Rania BENAZZA

Dirigée par: Dr. Sarah CIANFÉRANI

Introduction

Les développements dans le domaine de l'oncologie ont permis l'émergence de nouvelles classes de protéines thérapeutiques. Les anticorps monoclonaux (mAb), les anticorps à domaine unique (VHH), les mAb multi-spécifiques ou les anticorps immunoconjugués (ADC pour antibody-drug conjugates)[1] en sont la classe la plus représentative. Ces derniers sont conçus à partir du couplage covalent des mAb à des molécules cytotoxiques par le biais d'un bras espaceur (le linker)[2], et se distinguent des mAb par leur efficacité thérapeutique et leur spécificité accrues. La caractérisation des ADC/PDC (PDC pour protein drug conjugates) obtenus représente un véritable défi analytique en raison de leur haut poids moléculaire, de leur variabilité intrinsèque (modifications post-traductionnelles, PTM) et de l'hétérogénéité additionnelle résultant du processus de bioconjugaison[3-4].

La spectrométrie de masse (MS) couplée à la chromatographie liquide (LC) a le potentiel de surmonter ces difficultés et de faire face à la complexité structurelle de ces protéines afin d'évaluer les différents attributs de qualité critiques (CQA) exigés par les organismes de réglementation[5]. Il existe donc plusieurs approches et méthodologies pour caractériser ces protéines à plusieurs niveaux (intact, middle ou au niveau des peptides). La MS native (nMS) est une technique basée sur l'ionisation par electrospray (ESI), qui permet de transférer dans la phase gazeuse du spectromètre de masse la protéine solubilisée dans un tampon volatil et non dénaturant (typiquement de l'acétate d'ammonium, $AcONH_4$), ce qui permet de maintenir les interactions non-covalentes[6]. Pour cela, un échange de tampon est nécessaire. Il peut se faire manuellement ou de manière plus automatisée par couplage en ligne à la chromatographie d'exclusion stérique (SEC-nMS)[7]. Il s'agit d'une étape clé afin de gagner en temps d'analyse et d'apporter une autre dimension de caractérisation grâce à la séparation des espèces en fonction de leurs volumes hydrodynamiques. Les informations obtenues par SEC-nMS permettent de renseigner sur: i) la quantité de mAb non conjugué (D0), ii) le nombre moyen de drogues conjuguées par anticorps (avDAR, average drug-to-antibody ratio), iii) la distribution des drogues (DLD, drug load distribution) et iv) les variants de taille[8]. La nMS peut également être couplée à la

chromatographie échangeuse d'ions (IEX[9]) afin de séparer et caractériser les mAb en fonction de leur différence de charge nette. Plus récemment, des approches de séquençage des protéines intactes (sans digestion), appelées top down MS (TD-MS)[10], ont été décrites, permettant en plus d'une mesure de masse précise, de fragmenter les protéines directement au niveau intact (top, TD-MS) ou après une digestion contrôlée pour obtenir des fragments de 30-50 kDa (middle, MD-MS)[11]. Ainsi, la caractérisation des variants de charges (ex: PTM) et l'identification de la position des drogues sont accessibles directement, sans préparation d'échantillon extensive.

C'est dans ce contexte de développements de méthodologies basées sur des techniques de MS à la pointe pour améliorer la caractérisation des protéines conjuguées (à partir de mAb ou d'autres protéines) que se situe mon projet de thèse qui est articulé autour de deux axes principaux:

- L'amélioration des techniques de couplage LC-MS pour la caractérisation rapide et automatisée des protéines conjugués; ADC/PDC.
- Le développement de nouvelles approches de TD/MD-MS pour la caractérisation des protéines thérapeutiques.

Partie I : Introduction aux anticorps monoclonaux (mAb), aux formats basés sur les mAb, ainsi qu'à leurs stratégies de caractérisation

Cette première partie présente une introduction bibliographique aux mAb et aux formats basés sur les mAb, qui constituent la classe la plus importante des immunoglobulines G (IgG) à des fins thérapeutiques [12]. Une attention particulière est portée sur les ADC/PDC utilisés pour la thérapie anti-cancer (**Figure 1**). J'ai aussi introduit les différentes stratégies analytiques existantes pour caractériser ces biothérapeutiques avec une focalisation spécifique sur la nMS et TD/MD-MS qui font l'objet de cette thèse.

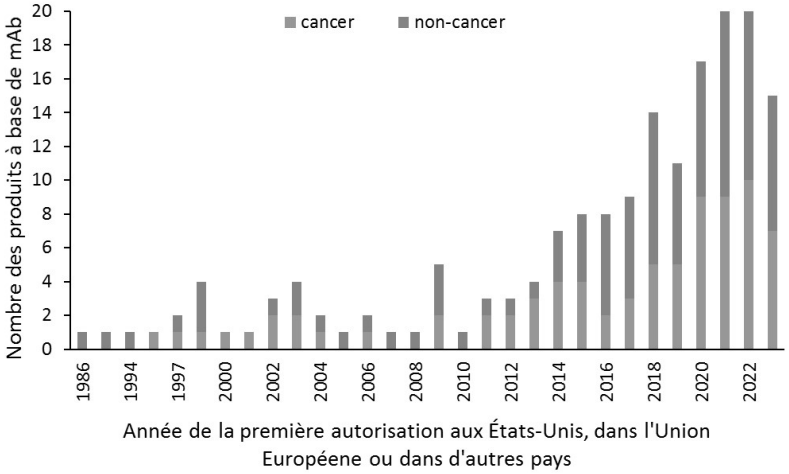


Figure 1: Les produits à base de mAb, approuvés aux Etas-Unis, dans l'Union Européene ou dans d'autres pays pour le traitement de diverses maladies, y compris le cancer. Chiffre basé sur des données publiquement disponibles au 04 octobre 2023. www.antibodysociety.org/antibody-therapeutics-product-data/

Partie II : Nouveaux couplages LC-MS pour la caractérisation des ADC/PDC dans des conditions natives et dénaturantes

La deuxième partie de mon manuscrit de thèse est axée sur l'automatisation du couplage de la SEC à la MS en conditions natives pour la caractérisation des protéines thérapeutiques conjuguées, ainsi que l'adaptation de ce couplage en conditions dénaturantes pour des systèmes plus spécifiques.

1. Evaluation d'une nouvelle plateforme LC-MS pour la caractérisation des biothérapeutiques

Ce premier chapitre concerne l'évaluation d'une plateforme intégrée LC-MS (BioAccord, Waters) pour l'analyse en routine de protéines thérapeutiques. Cet instrument est composé d'un système binaire de chromatographie liquide couplé à un détecteur à temps de vol (ToF). Le but de ces travaux était de développer des méthodes d'analyse en conditions natives basées sur le couplage de la nMS à différents types de chromatographie tels que la SEC ou l'IEX (SEC-nMS ou IEX-MS), ainsi que de systématiser le traitement des données.

En utilisant le trastuzumab comme mAb de référence et après optimisation des paramètres LC (choix de la colonne, choix de la phase mobile et du gradient), je me suis intéressée aux paramètres de MS (voltage de cône Vc, température de désolvatation T et pression en source Pi) afin d'évaluer leur influence pour mieux transmettre le mAb en conditions natives ou dénaturantes, au niveau intact et au niveau middle. De plus, le choix de la colonne a été particulièrement discuté dans le but de comparer une colonne SEC de dernière génération bio-inerte (Maxpeak BEH 250 Å 2.5 µm, 4.6 x 150 mm, Waters) et une colonne SEC classique (BEH SEC 200 Å 2.5 µm, 4.6 x 150 mm, Waters) (**Figure 2**). J'ai pu aussi évaluer la résolution, la sensibilité (injection jusqu'à 500 ng) et la précision de mesure de masse (≤27 ppm pour les protéines de ~150 kDa) du spectromètre utilisé. Les optimisations en conditions natives m'ont permis de trouver un compromis pour obtenir une désolvatation et une transmission efficaces des ions tout en préservant l'intégralité de la structure du mAb. L'ensemble de ces optimisations a été comparé aux instrumentations déjà présentes au sein du laboratoire (ex : ESI-Q-TOF Synapt G2, Waters), mettant en évidence des différences de paramétrage reliées à la possibilité d'ajuster (Synapt G2) ou pas (BioAccord) la pression dans la région de l'interface (1er zone de pompage après la source ESI). J'ai ensuite évalué le couplage entre la chromatographie d'échange de cations (CEX) et la nMS (CEX-nMS) sur le BioAccord afin de séparer et identifier les différents variants de charge des mAb à l'échelle intact ou middle. Pour cela, le choix du gradient de pH et de sels volatils a été une étape cruciale afin d'arriver à des conditions d'analyses optimales. Finalement, après ces différentes optimisations j'ai pu mettre en place des méthodologies versatiles de traitement des données en utilisant le logiciel UNIFI via l'interface WatersConnect, pour analyser les ADC/PDC nouvellement développés. Cette automatisation m'a permis d'effectuer et d'interpréter entre 10 et 15 analyses en

moins de deux heures. En conclusion, ce premier chapitre m'a permis de mettre en place des méthodes LC-MS automatisées en conditions natives pour la caractérisation rapide des protéines thérapeutiques, démontrant l'adéquation du BioAccord avec les spécifications d'un environnement industriel, pour la caractérisation de plusieurs types d'ADC/PDC produits par mes collègues dans le cadre du réseau européen ITN TACT (MSCA-ITN-2019), notamment : i) les ADC obtenus après des stratégies de conjugaison sélectives sur des mAb (BFC, Université de Strasbourg, France), ii) les PDC de l'anticaline[13] obtenus par réaction de Ugi[14] (BFC, Université de Strasbourg, France et Université technique de Munich, Allemagne) et iii) à un site les protéines bi-spécifiques assemblées après une bioconjugaison spécifique (Almac Discovery, Royaume Uni).

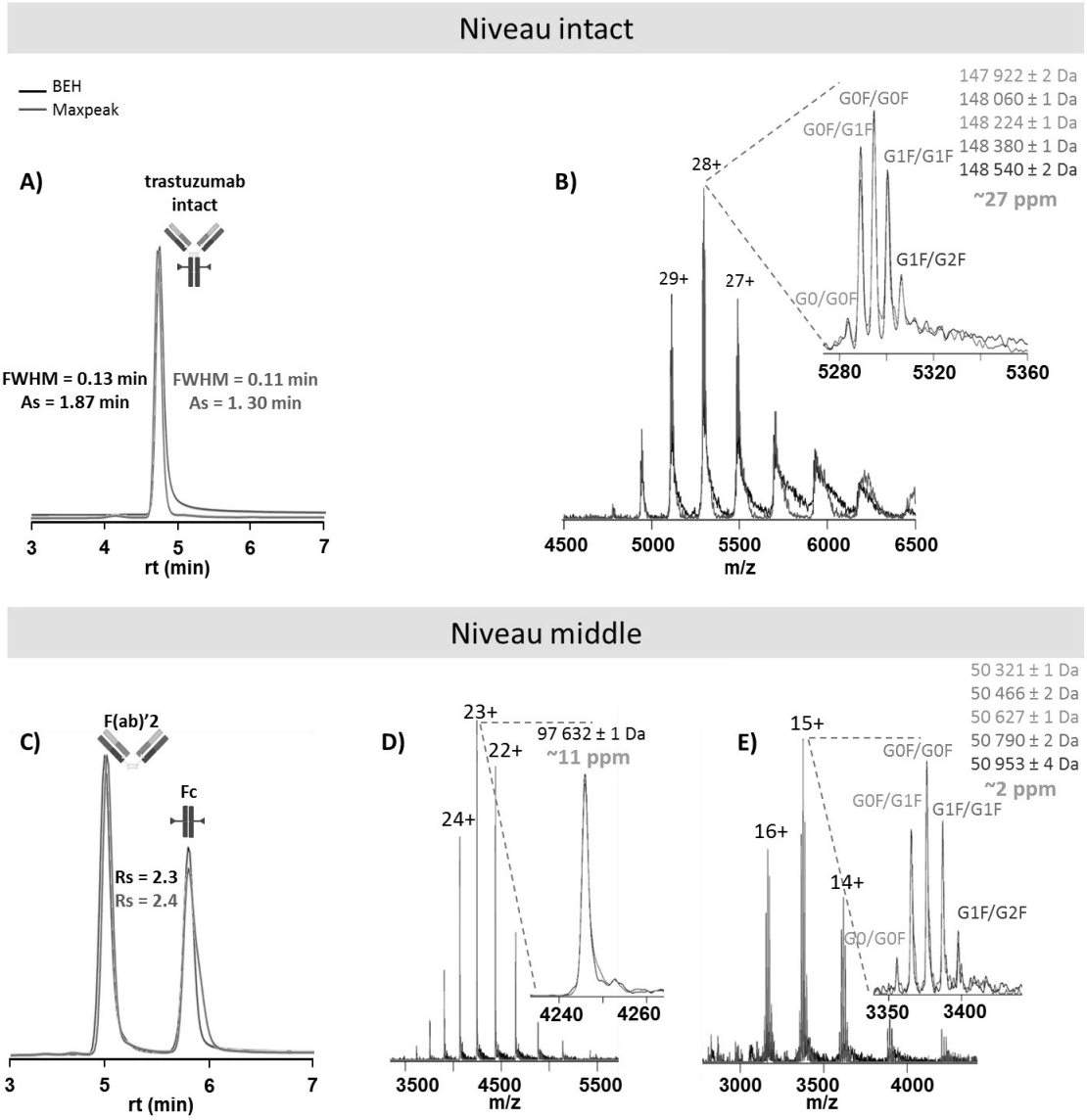


Figure 2: Analyse SEC-nMS du trastuzumab au niveau intact (haut) et middle (bas) en comparant deux différentes colonnes SEC; la colonne BEH SEC (noir) *versus* la colonne Maxpeak SEC (bleu). Chromatogrammes SEC-UV du trastuzumab intact (**A**) et du trastuzumab digéré par IdeS (**C**) montrant les différentes espèces séparées. Les spectres nMS révèlent l'identité de chaque espèce, à savoir (**B**) le trastuzumab intact, (**D**) le fragment F(ab)'2 et (**E**) la sous-unité Fc, avec un zoom sur les états de charge les plus intenses soulignant la présence de glycoformes pour le mAb intact et au niveau de la sous-unité Fc.

2. Application du couplage SEC-nMS pour la caractérisation des ADC/PDC issus de la conjugaison Ugi

Dans un second chapitre, j'ai souhaité illustrer l'apport de la méthode automatisée de SEC-nMS pour la caractérisation d'ADC et de PDC issus d'une conjugaison spécifique Ugi. La réaction Ugi est une réaction en deux étapes: **1**) la première étape consiste à conjuguer simultanément de manière covalente deux résidus d'acides aminés, l'amine de la chaîne latérale et les groupes carboxylates de deux lysines et aspartates/glutamates voisins avec un aldéhyde contenant de l'azide et un isocyanure. **2**) la deuxième étape est une dérivatisation de la drogue-linker par cycloaddition azide-alkyne (SPAAC) [51-53]. Dans notre cas d'étude, les mAb ont été conjugués à la BCN-iminobiotine, ce qui a permis l'incorporation d'une charge utile d'une masse moyenne de 769 Da (**Figure 3**). Afin d'avoir une conjugaison spécifique, il faut donc cribler plusieurs paramètres de réactions (ex : temps de réaction, pH de la réaction… etc.) ce qui nécessite énormément de temps pour caractériser les mAb conjugués à différentes conditions.

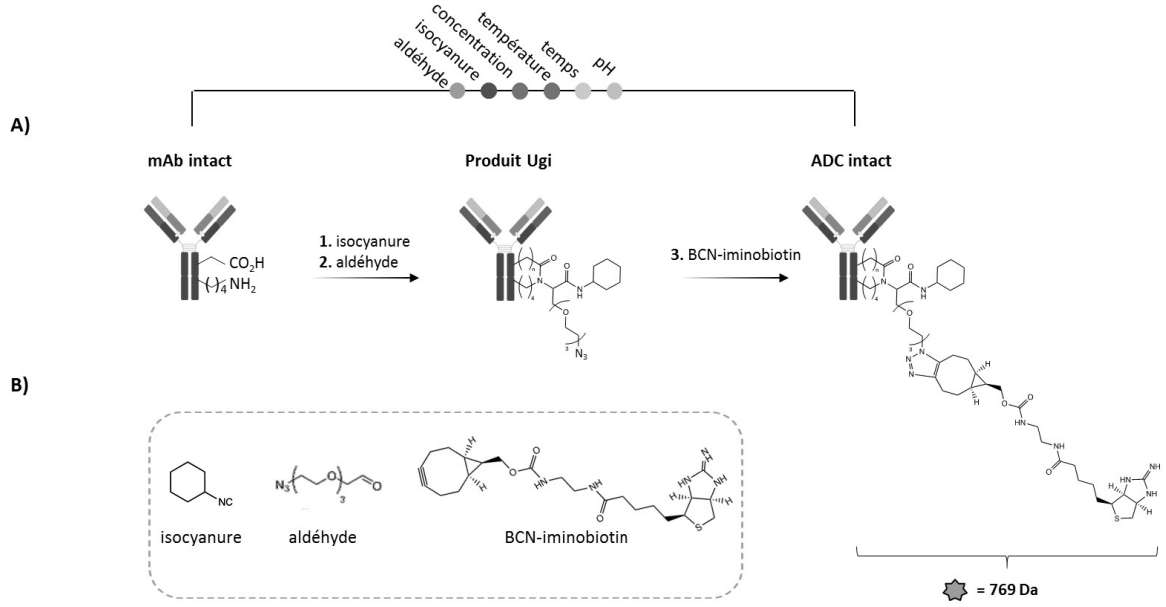


Figure 3: (**A**) Représentation schématique de la conjugaison de mAb intacts à l'aide de la réaction d'Ugi. (**B**) Les réactifs utilisés au cours des différentes étapes de la réaction d'Ugi.

La SEC-nMS a été un outil puissant dans ce cas, pour donner une mesure précise de la masse et une évaluation rapide des différents CQA à savoir, le D0, DLD et avDAR ce qui a permis de sélectionner les meilleurs conditions pour une conjugaison spécifique (**Figure 4**). La SEC-nMS a permis aussi d'évaluer l'applicabilité de la réaction Ugi sur des systèmes différents tel que les anticalines conjuguées (AcDC). Les résultats obtenus dans cette partie de ma thèse, suggèrent fortement l'intégration de la plateforme BioAccord dans les laboratoires du domaine biopharmaceutiques, afin d'accélérer le flux d'analyses.

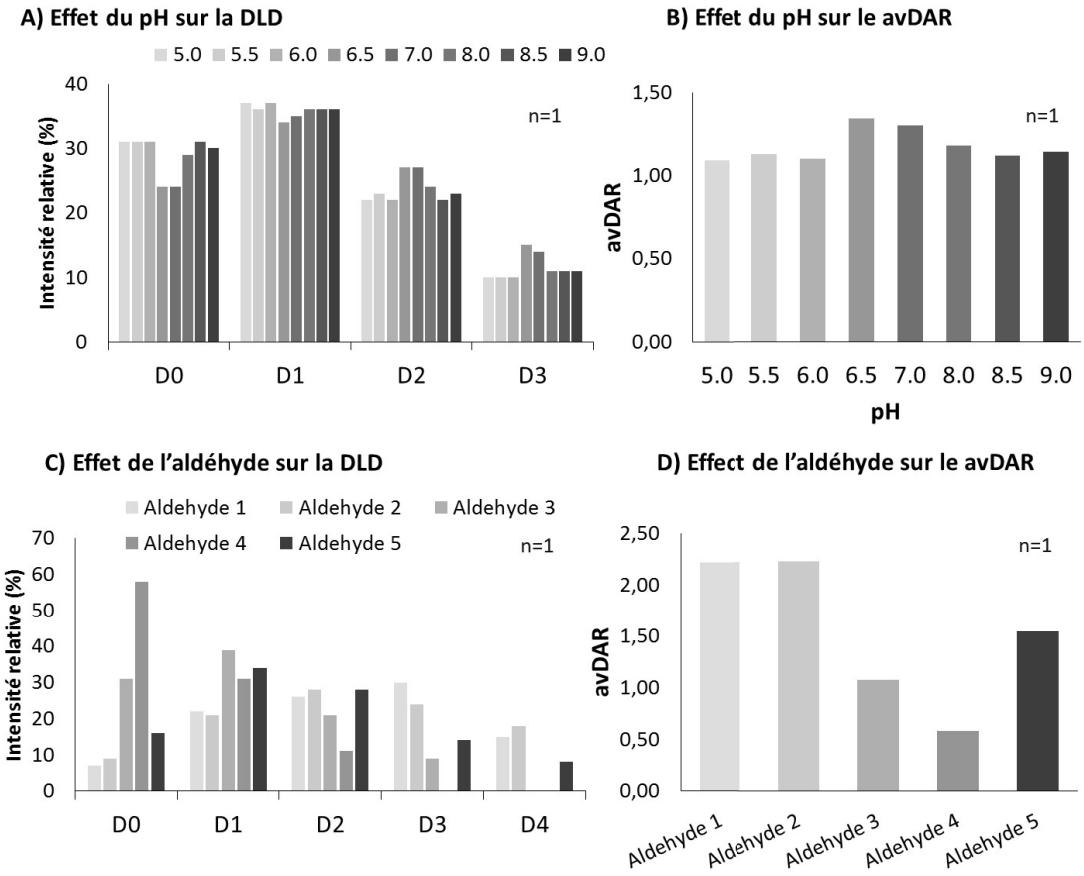


Figure 4: Analyses SEC-nMS des mAb conjugués issues de la réaction Ugi en utilisant différents pH et différents aldéhydes. L'effet du pH sur la DLD (**A**) et sur le avDAR (**B**). Effet de l'aldéhyde sur la DLD (**C**) et le avDAR (**D**).

3. Avantages de la SEC-nMS pour la caractérisation des anticorps bispécifiques (bsAb)

Le but de ce chapitre était d'évaluer la capacité de la plateforme LC-MS de BioAccord à caractériser des protéines plus complexes et de poids moléculaires plus élevé, à savoir les bsAb. En effet, j'ai démontré dans cette partie que la gamme de masse limitée (jusqu'à 7000 *m/z*) du BioAccord ne permet pas d'observer les formats de bsAb qui ont une masse de plus de ~200 kDa. J'ai alors caractérisé ces produits sur un instrument ToF similaire (Synapt G2 HDMS, Waters) qui a une gamme de m/z plus élevée (jusqu'à 10000 m/z) en couplage avec la SEC. Dans un premier lieu, la SEC m'a permis de quantifier relativement les différentes espèces issues de la production des bsAb, ce qui a donné une information sur l'efficacité de la réaction. Ensuite, la nMS a permis de révéler l'identité de chaque espèce grâce à la mesure de masse précise. De plus, j'ai évalué ici une technique complémentaire à la nMS, à savoir la masse photométrie (MP) qui est basée sur la microscopie à réflexion interférentielle. Cette technique permet l'analyse des échantillons à faibles quantité (100 pM – 100 nM) dans leur solution de stockage, ainsi que la caractérisation des espèces dans une large gamme de masse (jusqu'à 5 MDa). Les résultats obtenus avec la MP ont pu être corrélés avec les résultats de nMS, ce qui met en évidence la complémentarité de ces deux techniques pour la caractérisation des bsAb nouvellement développés (**Figure 5**).

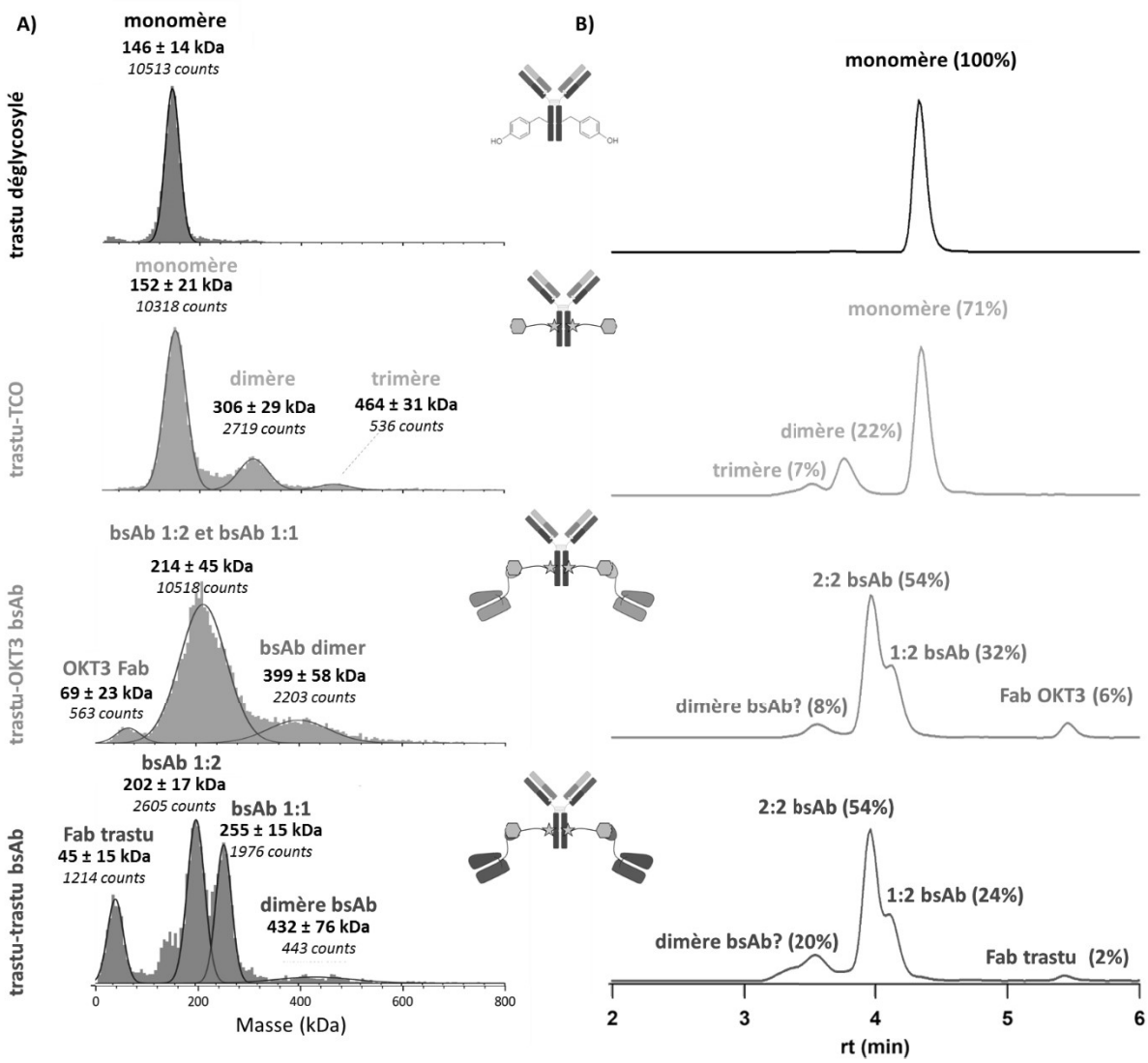


Figure 5: Analyse des échantillons obtenus lors d'une réaction de chimie click basée sur la tyrosine visant à former un bsAb trastuzumab-OKT3 1:2 et un bsAb trastuzumab-trastuzumab 1:2 à savoir : le trastuzumab déglycosylé, le trastuzumab fonctionnalisé avec un bras TCO, le bsAb trastuzumab-OKT3 et le bsAb trastuzumab-trastuzumab (**de haut en bas**). (**A**) Résultats obtenue avec la MP et (**B**) le signal UV obtenue après l'analyse SEC-nMS. La quantification relative en MP est basée sur le nombre de compte à base de mesure de molécules individuelles (counts) et celle en SEC-nMS est basée sur l'intégration des pics chromatographiques.

4. Evaluation d'une nouvelle plateforme LC-MS pour la caractérisation des biothérapeutiques

Dans le dernier chapitre, je décris le développement d'un nouveau couplage LC-MS associant la SEC en conditions dénaturantes et la MS pour la caractérisation détaillée d'un type particulier d'ADC, appelés « ADC rebridgés »[15], qui résultent d'une reformation des ponts disulfures (rebridging) avec incorporation simultanée d'une molécule cytotoxique. Cette méthode de conjugaison est utilisée en biochimie afin de modifier spécifiquement un mAb (**Figure 6**) ou son fragment de liaison à l'antigène (Fab) et conduit à la formation de nombreux produits de réactions non souhaités en plus des produits finaux attendus, tels que les sous unités libres (demi-mAb, chaine lourde (HC) et chaine légère (Lc) pour le mAb intact, ou le Fd et Lc dans le cas du Fab).

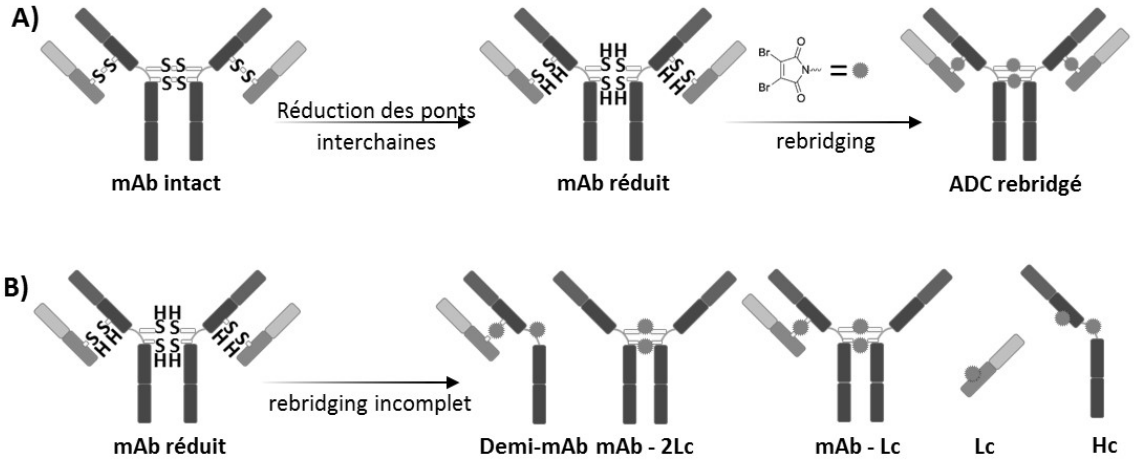


Figure 6: Représentation schématique de la formation d'un ADC rebridgé. (**A**) Un rebridging complet conduisant à l'incorporation de quatre charges utiles entre les chaînes lourdes et légères de l'anticorps. (**B**) Conjugaison incomplète d'un mAb intact conduisant à un mélange de sous-produits tels qu'un demi-mAb, un mAb sans 2 Lc, un mAb sans une Lc et des sous-unités Lc et Hc. La charge utile est représentée par une étoile rouge.

Dans ce contexte, ni l'approche classique de chromatographie liquide de phase inverse couplée à la MS (rpLC-MS) ni la SEC-nMS n'ont permis une caractérisation optimale et facile de l'ensemble des produits de réaction. J'ai ainsi développé une méthode LC-MS originale combinant les avantages de la séparation en taille de la SEC réalisée en phase mobile dénaturante à la précision de la mesure de masse en conditions dénaturantes, appelée dSEC-MS. Cette approche permet l'analyse simultanée des produits de réaction covalents attendus, ainsi que celles des produits secondaires de réaction, *via* leur séparation en fonction de leurs tailles. J'ai d'abord optimisé les paramètres dSEC (choix de la colonne, débit et phase mobile) et MS (Vc, T°) en utilisant un trastuzumab intact, digéré à la papaïne (Fab et Fc), digéré à l'IdeS (Fc/2, LC, Fd) et réduit au TCEP (LC et Hc), ce qui m'a permis de proposer une nouvelle méthode analytique rapide (15 min) grâce à l'utilisation de colonnes chromatographiques SEC bio-inertes et de dernière génération (MaxPeak Premier Protein SEC 250Å, 1.7 µm, 4.6 x 150 mm, Waters, Manchester, UK). Une étude similaire a déjà été rapportée dans la littérature mais celle-ci nécessitait un temps d'analyse très long (80 min) en utilisant des phases stationnaires induisant des réactions secondaires avec les analytes[16]. Finalement, en utilisant une colonne bio-inerte, j'ai pu obtenir des séparations chromatographiques efficaces des différentes espèces (Rs = 5.8) suivies d'une mesure de masse précise, ainsi que des pics chromatographiques plus symétriques (As = 1.74 pour le pic majeur). Ensuite, j'ai appliqué la méthode dSEC-MS optimisée à deux Fab et un mAb intact rebridgés dans le but de déterminer l'efficacité de la réaction ainsi que les CQA du produit conjugué (**Figure 1C**). En résumé, dans ce dernier chapitre, j'ai mis en place une méthode dSEC-MS qui permet d'évaluer l'efficacité des réactions de rebridging de mAb (~150kDa) ou de Fab (~50 kDa), en caractérisant l'ensemble des produits de réactions, du produit majoritaire ciblé aux produits issus de réactions incomplètes. Ces résultats de cette étude font l'objet d'un article soumis à *Talanta*.

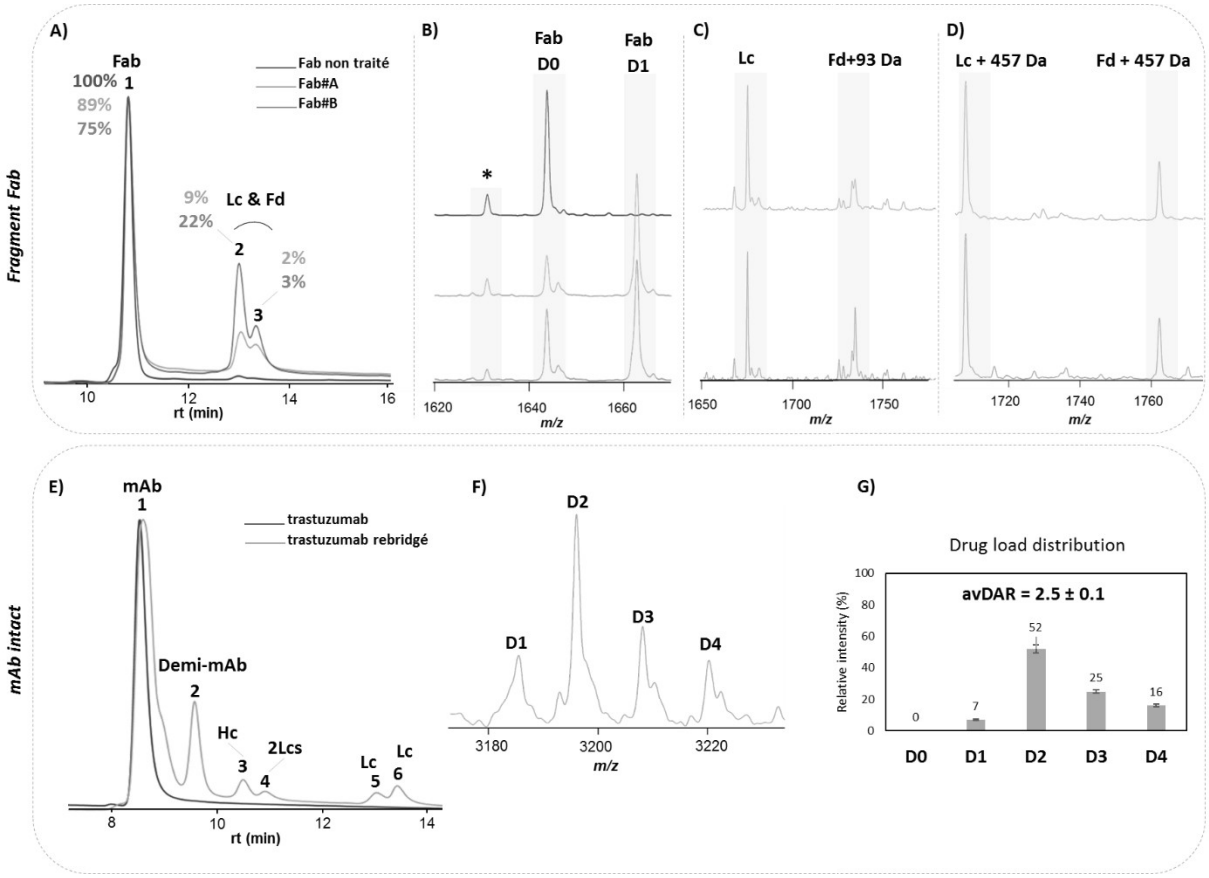


Figure 7: Analyse dSEC-MS de Fab et de mAb intact rebridgés. (**A**) Superposition des signaux UV-SEC des trois échantillons de Fab à savoir le Fab de référence (**bleu**), Fab#A (**rose**) et Fab#B (**vert**). La partie droite de la figure représente les spectres de masse correspondants de chaque échantillon avec (**B**) le zoom sur l'état de charge 30+ pour le pic Fab (~10,8 min), (**C**) le zoom sur le 14+ pour le deuxième pic (~12,9 min) et le (**D**) le zoom sur le 14+ pour le dernier pic (~13,3 min). (**E**) Profils SEC-UV dans des conditions dénaturantes du trastuzumab intact (**bleu**) et du trastuzumab rebridgé (**rose**). Les pics de 1 à 6 représentent les différentes espèces observées avec leurs masses correspondantes détectées dans la MS. (**F**) Profils MS de l'espèce de mAb rebridgé intact zoomé sur le 48+ et (**G**) représente la distribution de la charge utile de mAb rebridgé intact.

Les analyses des formats basés sur les mAb au niveau intact avec la MS native nous permettent de confirmer si la conjugaison a eu lieu ou non grâce à l'évaluation de l'avDAR et de la DLD. Cependant, pour déterminer le site spécifique de conjugaison, il serait nécessaire d'aller plus loin et de caractériser la structure primaire des ADC/PDC par la caractérisation des fragments.

Partie III : Développement de nouvelles approches MS de type top- et middle-down pour la caractérisation des ADC/PDC

Le second domaine de recherche que j'ai abordé pendant ma thèse porte sur le développement d'approches TD/MD-MS pour la caractérisation de la structure primaire des protéines thérapeutiques, en utilisant un spectromètre de dernière génération (Orbitrap Eclipse Tribrid, Thermo Fisher Scientific) équipé de plusieurs modes de fragmentations. Habituellement la caractérisation de la structure primaire de ce type de protéines est effectuée par des approches bottom-up[17] utilisées en

protéomique, qui consistent à séquencer les peptides obtenus après une digestion enzymatique par nanoLC-MS/MS. Ces approches de peptide mapping peuvent être chronophages et des artéfacts peuvent être induits soit au niveau de la préparation d'échantillon (conditions de digestions) ou de l'interprétation des données (grand nombre d'ions fragments). Des approches permettant le séquençage des protéines intactes (sans digestion préalable) sont donc intéressantes pour limiter les artéfacts et aboutir directement à des informations de séquence primaire, incluant l'identification et la localisation des PTM et des drogues (dans le cas d'ADC/PDC).

1. Implémentation d'approches TD- et MD-MS pour la caractérisation des protéines et des mAb de référence

Ce premier chapitre est dédié à la description des optimisations réalisées en LC, MS/MS et traitement des données pour l'analyse TD-MS de protéines de références (Myoglobine ~17 kDa et carbonique anhydrase ~29 kDa) et MD-MS d'un mAb de référence (trastuzumab ~150 kDa). L'analyse TD-MS de la myoglobine a permis de mettre en évidence l'importance de l'optimisation de plusieurs paramètres surtout le choix de l'ion précurseur et l'énergie/temps de dissociation, afin d'avoir un maximum de couverture de séquence. Ensuite, j'ai pu montrer la complémentarité de plusieurs techniques de fragmentation (HCD, ETD et UVPD) qui a permis d'atteindre 99% de couverture de séquence en combinant les fragments des trois fragmentations. En effet, obtenir une couverture de séquence presque complète n'est pas un défi quand il s'agit de la fragmentation de protéine de faible poids moléculaire et qui ne contiennent pas de PTM. J'ai pu confirmer ensuite cette hypothèse, avec l'analyse de la carbonique anhydrase où j'ai obtenu seulement 64% de couverture de séquence totale dans les mêmes conditions que ceux de la fragmentation de la myoglobine, mais avec l'identification de la PTM (N-acetylation en position N-ter). L'analyse TD-MS de protéines de plus de ~30-50kDa restant encore un véritable défi[18], nous avons réduit la taille de notre mAb par digestion enzymatique suivie d'une réduction des ponts disulfures, pour travailler sur des sous unités de ~25-30kDa (approche MD-MS) (**Figure 8**).

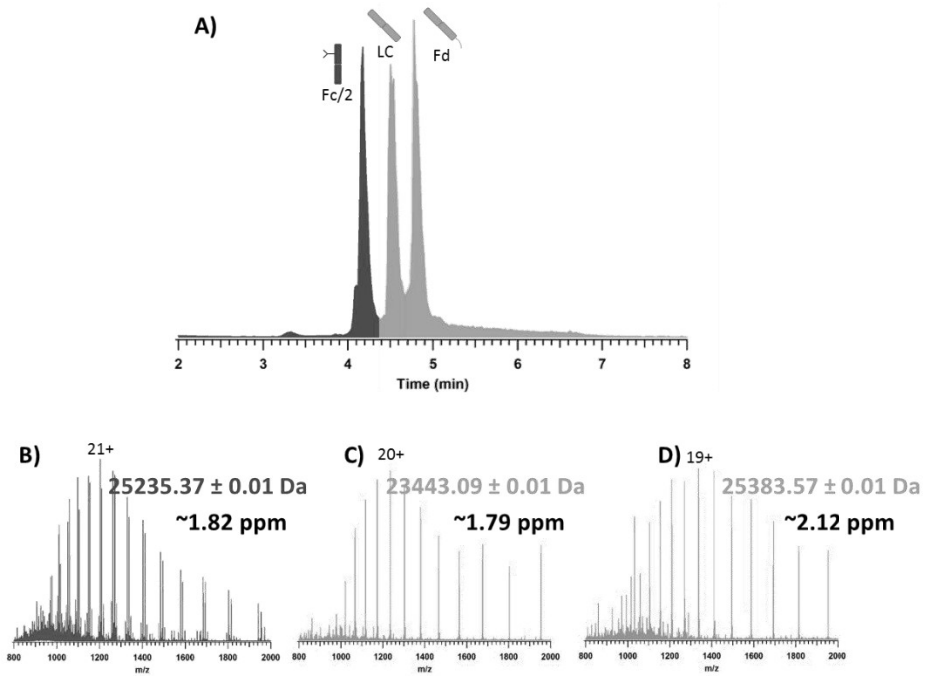


Figure 8: Analyse rpLC-MS des différentes sous-unités obtenues après une digestion contrôlée du trastuzumab à savoir ; le fragment Fc/2, Lc et Fd. (**A**) Signal UV des différentes sous-unités et (**B**) leur spectres MS avec leur masses précises.

Dans le cadre de ce travail, j'ai évalué plusieurs modes de fragmentation : la dissociation induite par collisions à haute énergie (HCD)[19], la dissociation par transfert d'électrons (ETD)[20] et la photodissociation dans l'ultraviolet (UVPD)[21], chacune de ces activations permettant d'obtenir un type différent de fragments (*b/y* dans le cas de l'HCD, *c/z* pour l'ETD et *a/x*, *b/y* et *c/z* pour l'UVPD). Dans le cadre de ma thèse, j'ai optimisé l'ensemble des étapes analytiques de l'approches LC-TD/MD-MS, à savoir:

Optimisations en chromatographie : L'optimisation du choix de la colonne chromatographique ainsi que de la phase mobile et du débit d'analyse est une étape primordiale afin de séparer les trois sous-unités (Lc, Fd et le fragment cristallisable Fc/2), résultant d'une digestion du trastuzumab à l'IdeS suivi d'une réduction au DTT, avant de les analyser en MD-MS.

Optimisation des signaux MS1 et choix des ions précurseurs : La première étape de l'analyse MD-MS consiste en une acquisition MS1 qui permet la mesure précise de chaque sous-unité. Afin d'obtenir une couverture de séquence maximale pour chaque sous-unité du mAb, il était d'abord nécessaire d'améliorer le signal des ions en optimisant plusieurs paramètres MS1 tels que la fragmentation en source (SID), la lentille de radiofréquence (RF lens) et la fenêtre d'isolement pour mieux transmettre les ions précurseurs de chaque sous-unité et ainsi pouvoir obtenir un meilleur rapport S/N pour les ions fragments. Ensuite, j'ai évalué l'influence du choix de l'ion précurseur (PI) sur l'efficacité de fragmentation. Ces premières optimisations m'ont permis de conclure que i) le choix du PI a un rôle très important pour améliorer la couverture de séquence, ii) que l'ion le plus chargé est le mieux

fragmenté surtout dans le cas de l'ETD, et iii) que la sélection simultanée de plusieurs ions engendre une baisse de la couverture de séquence comme déjà indiqué dans la littérature[10].

Optimisation des paramètres de fragmentation spécifiques aux modes d'activation étudiés (HCD, ETD, UVPD) : j'ai ainsi évalué l'influence du temps d'activation pour l'ETD, de l'énergie de fragmentation en l'HCD et du temps d'irradiation de l'UVPD sur le recouvrement de séquences. J'ai ainsi pu constater que : i) des recouvrements de séquence de 15-26% en HCD, ii) de 35% à 44% en ETD ; et iii) en UVPD de 16% à 23% ont été obtenus. En combinant les résultats obtenus à partir de ces trois techniques de fragmentation, un recouvrement de 63 et 75% des séquences des différentes sous-unités a été obtenu, avec une localisation du site de glycosylation sur la Fc/2, ce qui met en évidence la complémentarité des différentes techniques d'activation utilisées (**Erreur ! Source du renvoi introuvable.**).

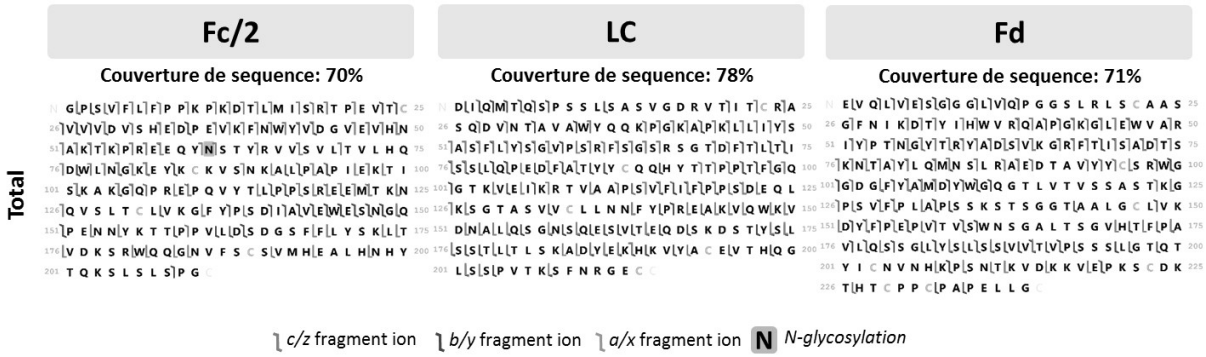


Figure 9: Couverture de séquence obtenue pour chaque sous-unités après analyse MD-MS en combinant les fragments issus des trois techniques de fragmentations; HCD (ions fragments b et y en bleu), ETD (ions fragments c et z en rouge) et UVPD (ions fragments a et x en vert, en plus des ions b, c, y et z)

Dans l'espoir de couvrir les parties de la séquence encore inaccessibles, j'ai évalué l'utilisation de la réduction de charge par transfert de protons (PTCR)[22] . Cette technique consiste à faire réagir un cation à charge multiple de notre protéine avec un anion à charge unique issue d'une molécule perfluorée, afin de réduire la charge des ions fragments de la protéine et mieux séparer les ions [23]. L'utilisation de la PTCR m'a permis de décomplexifier les spectres et d'accéder à de nouvelles régions à fragmenter. De plus j'ai eu recours à un logiciel (ClipsMS)[24] qui permet d'assigner les fragments internes qui ne sont pas reconnus par les logiciels classiques de TD-MS. Cet algorithme m'a aidé à identifier plus de fragments et à maximiser la couverture de séquence de toutes les sous-unités. Ces optimisations ont permis d'atteindre jusqu'à 80% de la couverture de séquence de chaque sous-unité.

Cependant, l'inclusion de ces fragments internes peut induire l'assignement de faux-positives, et donc il faut l'utiliser avec une grande précaution (**Figure 10**).

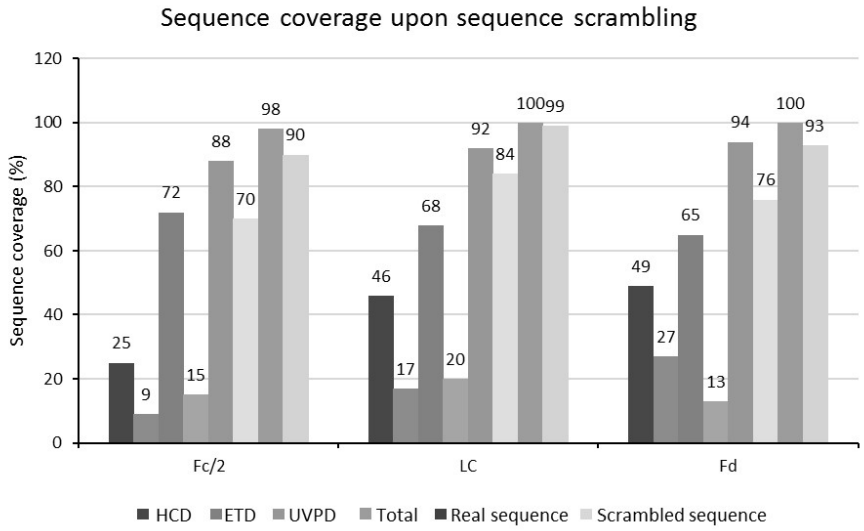


Figure 10: La couverture de séquence des différentes sous-unités (Fc/2, Lc et Fd) en utilisant différentes techniques de fragmentations à savoir ; HCD (bleu), ETD (rouge) et UVPD (vert), en prenant en considération les fragments internes. Les résultats obtenus avec les séquences réelles sont en couleurs sombres, et ceux obtenues avec la séquence mélangée sont en couleurs plus claires.

En résumé, l'utilisation combinée des trois techniques de fragmentation implémentées sur l'Orbitrap Eclipse (ETD, HCD et UVPD) montre une homogénéité et une complémentarité de ces techniques pour la fragmentation des sous-unités du trastuzumab, elle peut être maximisé de 10% en utilisant de la PTCR et ClipsMS pour le traitement des données. Cette méthodologie est prête à être utilisée pour la fragmentation des larges protéines.

2. Caractérisation complète d'un nanobody conjugué (sdADC) par TD-MS

Le second chapitre concerne l'application de l'approche TD-MS développée précédemment pour caractériser un anticorps à domaine unique (sdAb pour single domaine antibody, encore appelé nanobody)[25]. Il s'agit d'une protéine thérapeutique de faible poids moléculaire (~15 kDa) qui est formée par un domaine variable de la chaîne lourde d'un anticorps (VHH pour variable domain of heavy-chain only antibody). Ces protéines se distinguent des mAb par leur petite taille, leur structure simple, leur spécificité, leur grande affinité pour les antigènes et leur stabilité remarquable[26], ce qui en fait une classe de biomolécules particulièrement intéressante pour des applications médicales en diagnostic et en thérapie[27]. Comme pour les mAb intacts, ces VHH peuvent être conjugués à des drogues cytotoxiques (sdADC pour single domain antibody drug-conjugate) afin d'améliorer leur spécificité et leur potentiel thérapeutique.

Durant un stage réalisé dans le cadre du consortium européen TACT, j'ai pu produire et conjuguer un VHH au sein de la société Almac Discovery (Edimbourg, Royaume Uni) afin de l'utiliser pour le développement et la validation de la méthode TD-MS. Il s'agit d'une protéine ciblée contre le récepteur du facteur de croissance épidermique (EGFR) surexprimé dans un grand nombre de cellules cancéreuses[28]. En ciblant l'EGFR, la protéine anti-EGFR vise à bloquer les signaux de croissance qui

favorisent la prolifération des cellules tumorales, contribuant ainsi à freiner la progression de la maladie et à améliorer les résultats du traitement chez les patients atteints de cancer[29-30].

La première étape de ce projet a consisté en l'expression de la protéine anti-EGFR dans *Escherichia coli*, suivie d'une purification par chromatographie d'affinité en utilisant une intéine et un domaine de liaison à la chitine (CBD)[31]. Ensuite, j'ai utilisé la cystéamine pour réaliser le clivage de la protéine. La cystéamine est un agent réducteur et donc son utilisation permet de réduire les ponts disulfures présents dans la structure de l'anti-EGFR, ce qui permet de rompre les liaisons intermoléculaires et libérer la protéine sdAb sous forme monomérique. Enfin, la dernière étape est la conjugaison du sdAb à une molécule fluorescente (Alexa Fluor, AF488, comme mimique de drogue non cytotoxique) grâce à une réaction maléimide avec la cystéamine, permettant d'éviter la réduction préalable du pont disulfure tout en ciblant une conjugaison spécifique sur le résidu C-terminal (**Figure 11**). A partir des absorbances UV, d'analyses par gel d'électrophorèse SDS-PAGE et par caractérisation MS de la protéine conjuguée et non conjuguée, environ 300 µg des protéines sdAb et sdADC ont été obtenus avec une pureté de ~99%.

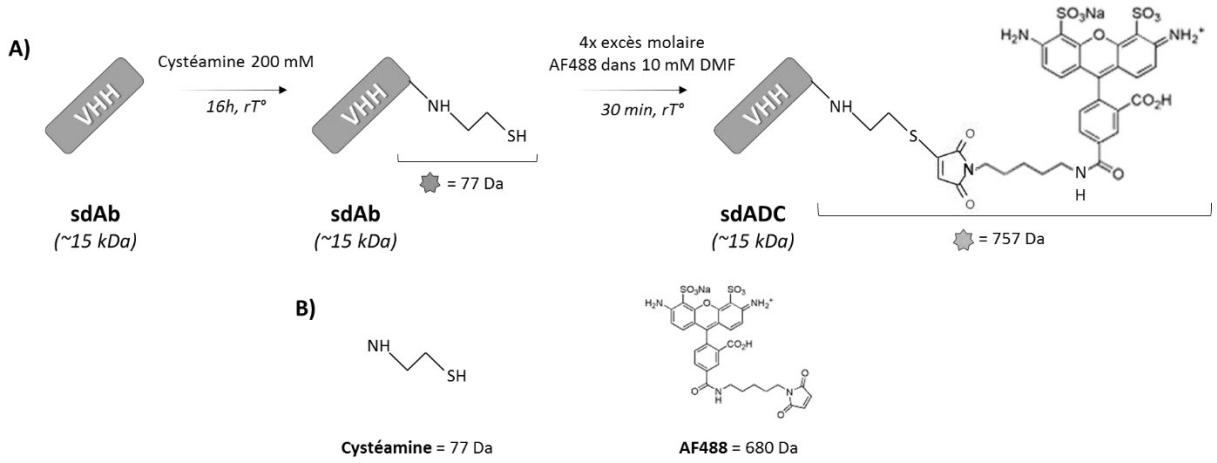


Figure 11: (**A**) Représentation du processus de production d'anticorps à chaine unique conjuguées (sdADC) qui consiste en i) clivage du nanobody avec une cysteamine qui se positionne en partie C-ter du nanobody avec une addition d'une masse de 77 Da représenté par une étoile bleu. ii) la deuxiième étape consiste à conjuguer spécifiquement la cystéamine ajoutée en ajoutant une molécule fluorescente (AF488) qui induit l'addition d'une masse de 757 Da, représentée par une étoile orange. (**B**) Les structures de la cystéamine et AF488 utilisés lors de la réaction de conjugaison.

La seconde partie de mon travail a été réalisée à nouveau au LSMBO et concerne la caractérisation analytique du sdAb et de son ADC (sdADC). J'ai d'abord utilisé la SEC-nMS, afin de déterminer la DLD et le DAR moyen du sdADC, ce qui m'a permis de confirmer la conjugaison homogène d'une seule molécule de AF488 sur le sdAb (average DAR = 1, pas de D0 détecté), comme attendu. Afin de localiser le site spécifique de conjugaison, j'ai appliqué la méthode TD-MS préalablement optimisée en l'adaptant aux deux systèmes, sdAb et sdADC. En dissociant avec l'HCD seulement, j'ai pu confirmer la localisation du site de la drogue qui est bien en position C-terminale, puisque les fragments obtenus

sont spécifiques de la modification et de la position, et ceci malgré un faible recouvrement de séquence (31%). Pour confirmer cette information et maximiser la couverture de séquence, j'ai utilisé l'ETD et l'UVPD, pour lesquels après optimisation des temps de dissociation, et en combinant les résultats obtenus des trois techniques de fragmentation j'ai pu atteindre jusqu'à 88% de couverture de la séquence du sdADC (**Figure 12**). Ces résultats montrent clairement la complémentarité des trois méthodes d'activation et confirment le site de modification sur le résidu C-terminal. De plus, cet exemple TD-MS du sdAb/sdADC m'a également permis de mettre en évidence les apports d'outils en libre accès pour le traitement des données de TD-MS, avec notamment la prise en compte des fragments internes avec le logiciel ClipsMS[24], ce qui a permis d'augmenter de plus de 10% la couverture de séquence pour arriver avec 99%.

Finalement j'ai voulu valider la présence des ponts disulfures dans mes échantillons de sdAb/sdADC en réalisant des expériences comparatives entre le sdADC réduit et non réduit. J'ai pu identifier les fragments qui contiennent les cystéines réduites/oxydés et donc confirmer la position du pont disulfure en utilisant particulièrement la fragmentation UVPD. Ceci est en accord avec une étude qui a rapporté que l'UVPD a la capacité de cliver les liaisons S-S en induisant des fragmentations à la fois homolytiques et par transfert d'hydrogène[32].

En résumé, dans cette partie, j'ai illustré l'apport de l'approche TD/MD-MS développée dans le cadre de ma thèse pour la caractérisation des protéines thérapeutiques, sans préparation d'échantillon laborieuse, en utilisant un spectromètre de masse de dernière génération disposant de plusieurs modes de fragmentation complémentaires, le tout combiné à l'utilisation de logiciels performants pour le traitement de données (article en cours de rédaction). Cette méthode permettra à l'avenir d'accélérer le développement des nouvelles molécules biothérapeutiques, en complétant les informations obtenues grâce aux autres techniques de la MS.

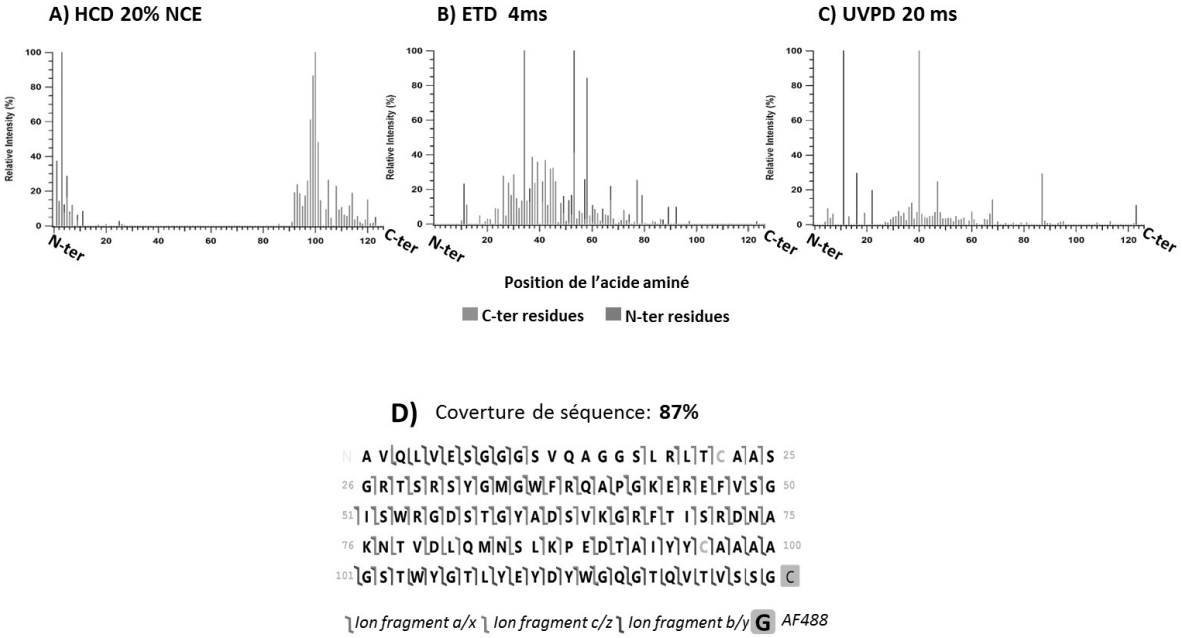


Figure 12: Localisation des différents fragments C-ter et N-ter obtenues lors de la fragmentation du sdADC par (**A**) HCD en utilisant 20% d'énergie normalisée (NCE), (**B**) ETD pendans 4 ms et (**C**) UVPD pendant 20 ms. (**D**) Carte de fragmentation montrant une couverture de séquence de 87% avec la localisation du site de la conjugaison (représentée en carré orange) en combinant les résultats de HCD, ETD et UVPD.

Conclusion générale et perspectives

Mon projet de thèse s'articule autour du développement d'approches de pointe en MS pour la caractérisation des biothérapeutiques, visant à résoudre deux limitations importantes rencontrées au début de ma thèse : i) le manque d'automatisation des méthodes LC-MS pour la caractérisation des PDC/ADC, et ii) le recours à des approches chronophages et qui peuvent générer des artéfacts pour la caractérisation de leur structure primaire des ADC/PDC par empreinte peptidique.

J'ai tout d'abord développé des méthodes de MS couplées à la chromatographie d'exclusion stérique en conditions natives (SEC-nMS) et en conditions dénaturantes (dSEC-MS) spécifiquement adaptées aux PDC/ADC. Ces méthodes, de l'analyse jusqu'au traitement des données, permettront de disposer de nouveaux outils performants et adaptés à la caractérisation des formats de PDC/ADC toujours plus complexes développées par l'industrie biopharmaceutique. Bien que la méthode dSEC-MS nécessite des améliorations supplémentaires pour augmenter la résolution de la séparation des espèces ayant le même ordre de grandeur en volume hydrodynamique, diverses optimisations peuvent être envisagées. Parmi celles-ci figure l'utilisation de différentes phases mobiles et de colonnes bio-inertes (de 30 cm de longueur par exemple). L'émergence de ces couplages versatiles pourrait également les rendre utiles en tant que méthodes de routine dans les laboratoires biopharmaceutiques pour caractériser différents types de biothérapeutiques, comme : i) les bsAb pour étudier leur structure, afin de mieux comprendre leur impact sur l'activité biologique, ou ii) les biosimilaires pour évaluer

d’éventuelles différences cliniquement significatives en effectuant des études comparatives avec leurs mAb de référence.

Dans un second temps, j'ai développé une approche TD/MD-MS pour la caractérisation d’ADC. Cette approche m’a permis de caractériser en détail la structure primaire du trastuzumab (PTM incluses) et d’un VHH anti-EGFR bioconjugué en cours de développement, en identifiant les types et les positions des PTM et/ou sites de conjugaison. En localisant précisément ces « hot spots » ou « points de greffage », ces approches m’ont permis une meilleure compréhension de la structure et de ces biothérapeutiques. Cependant, l’application de l’approche TD-MS aux mAb d’une taille d’environ 150 kDa reste un défi, ouvrant ainsi de nouvelles perspectives : i) l'évaluation de nouvelles plateforme combinant différentes techniques de fragmentations à des niveaux multiples tel que la nouvelle cellule trap d’ions Omnitrap (Fasmatech, Grèce), ou ii) l’adaptation l’utilisation potentielle de nouveaux spectromètre de masses qui se distinguent d’une haute sensibilité et d’une haute résolution grâce à une meilleure focalisation et moins pertes d’ions, tel que le SELECT SERIES MRT récemment développé (Waters) [29] ou l'Orbitrap Astral MS (Thermo Fisher) [30].

L’amélioration des approches TD/MD-MS pourrait également passer par le développement de nouvelles séquences de fragmentation (type MS3) afin de pousser plus loin les capacités de ces techniques d’activation, et aussi confirmer les sites et positions des PTM préalablement déterminés en MS2. D'autre part, l'association des approches TD/MD-MS à d'autres techniques chromatographiques (telles que la SEC ou la CEX) permettrait d’obtenir une séparation des différentes variants de taille ou de charge suivi de leur fragmentation en ligne, ce qui pourrait être utile dans l'industrie biopharmaceutique pour une caractérisation plus complète et plus rapide de la séquence primaire et des différentes PTM. De plus, l’implémentation des fragmentations des protéines en top-down ou middle-down pourraient être combinées aussi avec la mobilité ionique d’un côté pour séparer les protéines en fonction de leurs sections efficaces (CCS pour collision cross section) et associer leurs spectres de fragmentation avec leurs conformations, ou pour utiliser cette dimension pour mieux séparer les ions fragments et ainsi donner un recouvrement de séquence plus complète. La communauté top-down pourrait aussi en bénéficier de logiciels plus sophistiqués et plus adaptés pour un assignement de fragments terminaux et internes avec plus de confiance.

Références

1. Carter, P. J.; Lazar, G. A., Next generation antibody drugs: pursuit of the 'high-hanging fruit'. *Nat Rev Drug Discov* **2018,** *17* (3), 197-223.
2. Kostova, V.; Desos, P.; Starck, J. B.; Kotschy, A., The Chemistry Behind ADCs. *Pharmaceuticals (Basel)* **2021,** *14* (5), 442-449.
3. Fu, Z.; Li, S.; Han, S.; Shi, C.; Zhang, Y., Antibody drug conjugate: the "biological missile" for targeted cancer therapy. *Signal Transduct Target Ther* **2022,** *7* (1), 93.
4. Kaplon, H.; Crescioli, S.; Chenoweth, A.; Visweswaraiah, J.; Reichert, J. M., Antibodies to watch in 2023. *MAbs* **2023,** *15* (1), 2153410.
5. Wakankar, A.; Chen, Y.; Gokarn, Y.; Jacobson, F. S., Analytical methods for physicochemical characterization of antibody drug conjugates. *MAbs* **2011,** *3* (2), 161-72.
6. Leney, A. C.; Heck, A. J., Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom* **2017,** *28* (1), 5-13.
7. Botzanowski, T.; Erb, S.; Hernandez-Alba, O.; Ehkirch, A.; Colas, O.; Wagner-Rousset, E.; Rabuka, D.; Beck, A.; Drake, P. M.; Cianferani, S., Insights from native mass spectrometry approaches for top- and middle- level characterization of site-specific antibody-drug conjugates. *MAbs* **2017,** *9* (5), 801-811.
8. Ehkirch, A.; Hernandez-Alba, O.; Colas, O.; Beck, A.; Guillarme, D.; Cianferani, S., Hyphenation of size exclusion chromatography to native ion mobility mass spectrometry for the analytical characterization of therapeutic antibodies and related products. *J Chromatogr B Analyt Technol Biomed Life Sci* **2018,** *1086*, 176-183.
9. Fekete, S.; Beck, A.; Veuthey, J. L.; Guillarme, D., Ion-exchange chromatography for the characterization of biopharmaceuticals. *J Pharm Biomed Anal* **2015,** *113*, 43-55.
10. Hernandez-Alba, O.; Houel, S.; Hessman, S.; Erb, S.; Rabuka, D.; Huguet, R.; Josephs, J. L.; Beck, A.; Drake, P. M.; Cianferani, S., A case study to identify the drug conjugation site of a site-specific antibody-drug-conjugate using middle-down mass spectrometry. *JASMS* **2019,** *30* (11), 2419–2429.
11. Fornelli, L.; Srzentic, K.; Huguet, R.; Mullen, C.; Sharma, S.; Zabrouskov, V.; Fellers, R. T.; Durbin, K. R.; Compton, P. D.; Kelleher, N. L., Accurate Sequence Analysis of a Monoclonal Antibody by Top-Down and Middle-Down Orbitrap Mass Spectrometry Applying Multiple Ion Activation Techniques. *Anal Chem* **2018,** *90* (14), 8421-8429.
12. Muhammed, Y., The Best IgG Subclass for the Development of Therapeutic Monoclonal Antibody Drugs and their Commercial Production: A Review. *Immunome Research* **2020,** *16* (1).
13. Rothe, C.; Skerra, A., Anticalin((R)) Proteins as Therapeutic Agents in Human Diseases. *BioDrugs* **2018,** *32* (3), 233-243.
14. Sornay, C.; Hessmann, S.; Erb, S.; Dovgan, I.; Ehkirch, A.; Botzanowski, T.; Cianferani, S.; Wagner, A.; Chaubet, G., Investigating Ugi/Passerini Multicomponent Reactions for the Site-Selective Conjugation of Native Trastuzumab*. *Chemistry* **2020,** *26* (61), 13797-13805.
15. Marquard, A. N.; Carlson, J. C. T.; Weissleder, R., Expanding the Scope of Antibody Rebridging with New Pyridazinedione-TCO Constructs. *Bioconjug Chem* **2020,** *31* (6), 1616-1623.
16. Liu, H.; Gaza-Bulseco, G.; Chumsae, C., Analysis of reduced monoclonal antibodies using size exclusion chromatography coupled with mass spectrometry. *J Am Soc Mass Spectrom* **2009,** *20* (12), 2258-64.
17. Boeuf, A.; Debaene, F.; Ayoub, D.; Diemer, H.; Ehkirch, A.; Wagner-Rousset, E.; Van Dorsselaer, A.; Cianferani, S.; Beck, A., Mass Spectrometry-Based Strategies for Therapeutic Antibodies Extensive Characterization and Optimization (OptimAbs). *Structural Biology in Drug Discovery: Methods, Techniques, and Practices* **2020,** *21*, 503.

18. Fornelli, L.; Parra, J.; Hartmer, R.; Stoermer, C.; Lubeck, M.; Tsybin, Y. O., Top-down analysis of 30–80 kDa proteins by electron transfer dissociation time-of-flight mass spectrometry. *Anal Bioanal Chem* **2013,** *405*, 8505–8514.
19. Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M., Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* **2007,** *4* (9), 709-12.
20. Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* **2004,** *101* (26), 9528-33.
21. Brodbelt, J. S.; Morrison, L. J.; Santos, I., Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chem Rev* **2020,** *120* (7), 3328-3380.
22. W. J. Herron; D. E. Goeringer; McLuckey, S. A., Gas-Phase Electron Transfer Reactions from Multiply-Charged Anions to Rare Gas Cations. *J. Am. Chem. Soc.* **1995,** *117* (46), 11555-11562.
23. Huguet, R.; Mullen, C.; Srzentic, K.; Greer, J. B.; Fellers, R. T.; Zabrouskov, V.; Syka, J. E. P.; Kelleher, N. L.; Fornelli, L., Proton Transfer Charge Reduction Enables High-Throughput Top-Down Analysis of Large Proteoforms. *Anal Chem* **2019,** *91* (24), 15732-15739.
24. Lantz, C.; Zenaidee, M. A.; Wei, B.; Hemminger, Z.; Ogorzalek Loo, R. R.; Loo, J. A., ClipsMS: An Algorithm for Analyzing Internal Fragments Resulting from Top-Down Mass Spectrometry. *J Proteome Res* **2021,** *20* (4), 1928-1935.
25. Pronk, S. D.; Schooten, E.; Heinen, J.; Helfrich, E.; Oliveira, S.; van Bergen En Henegouwen, P. M. P., Single Domain Antibodies as Carriers for Intracellular Drug Delivery: A Proof of Principle Study. *Biomolecules* **2021,** *11* (7), 927.
26. Hassanzadeh-Ghassabeh, G.; Devoogdt, N.; De Pauw, P.; Vincke, C.; Muyldermans, S., Nanobodies and their potential applications. *Perspective* **2013,** *8* (6), 1013–1026.
27. Pillay, T. S.; Muyldermans, S., Application of Single-Domain Antibodies ("Nanobodies") to Laboratory Diagnosis. *Annals of Laboratory Medicine* **2021,** *41* (6), 549-558.
28. Sigismund, S.; Avanzato, D.; Lanzetti, L., Emerging functions of the EGFR in cancer. *Mol Oncol* **2018,** *12* (1), 3-20.
29. Ho, E. C. H.; Qiu, R.; Miller, E.; Bilotta, M. T.; FitzGerald, D.; Antignani, A., Antibody drug conjugates, targeting cancer-expressed EGFR, exhibit potent and specific antitumor activity. *Biomed Pharmacother* **2023,** *157*, 114047.
30. Gibson, T. J.; Smyth, P.; McDaid, W. J.; Lavery, D.; Thom, J.; Cotton, G.; Scott, C. J.; Themistou, E., Single-Domain Antibody-Functionalized pH-Responsive Amphiphilic Block Copolymer Nanoparticles for Epidermal Growth Factor Receptor Targeted Cancer Therapy. *ACS Macro Lett* **2018,** *7* (8), 1010-1015.
31. Ta, D. T.; Redeker, E. S.; Billen, B.; Reekmans, G.; Sikulu, J.; Noben, J. P.; Guedens, W.; Adriaensens, P., An efficient protocol towards site-specifically clickable nanobodies in high yield: cytoplasmic expression in Escherichia coli combined with intein-mediated protein ligation. *Protein Eng Des Sel* **2015,** *28* (10), 351-63.
32. Macias, L. A.; Brodbelt, J. S., Investigation of Product Ions Generated by 193 nm Ultraviolet Photodissociation of Peptides and Proteins Containing Disulfide Bonds. *J Am Soc Mass Spectrom* **2022,** *33* (7), 1315-1324.

Scientific communications

List of publications:

Accepted

1. Koutsopetras I., A. K. Mishra, Benazza R., Hernandez-Alba O., Cianférani S., Chaubet G., Nicolai S. and Waser J.. Cysteine-**Cysteine Cross-Conjugation of both Peptides and Proteins with a Bifunctional Hypervalent Iodine-Electrophilic Reagent**. *ChemEurJ*, **2023**.
2. Sebastiani M., Behrens C., Dörr S., Gerber H., Benazza R., Hernandez-Alba O., Cianférani S., Klebe G., Heine A. and Reuter K. **Structural and biochemical investigation of the heterodimeric murine tRNA-guanine transglycosylase**. *ACS chem. Biology,* **2022**.

Submitted/In preparation

1. Benazza R., Hernandez-Alba O. and Cianférani S. **SEC-MS in Denaturing Conditions (dSEC-MS) for Rapid In-Depth Analysis of Rebridged Monoclonal Antibody-Based Formats**. *Submitted to Talanta*, **2023**.
2. Benazza R., Papadakos G., Diemer H., Cotton G., Cianférani S. and Hernandez-Alba O. **Characterization of anti-EGFR nanobodies using top-down mass spectrometry (TD-MS).** *In preparation*.
3. Koutsopetras I., Vaur V., Benazza R., Diemer H., Sornay C., Ersoy Y., Rochet L., Longo C., Hernandez-Alba O., Erb S., Detappe A., Skerra A., Wagner A., Cianférani S. and Chaubet G. **Site-Selective Protein Conjugation by a Multicomponent Ugi Reaction.** *Submitted to ChemEurJ*, **2023**.
4. Shajan I., Rochet L., Tracey S. R., Jackowska B., Benazza R., Hernandez-Alba O., Cianférani S., Scott C. J., Chudasama V., van Delft F. L. and Albada B. **Rapid Access to Potent Bispecific T-cell Engagers with Biogenic Tyrosine Click Chemistry**. *Submitted to Bioconjugate Chem*, **2023**.
5. Gervason S., Want K., Dukewitch R., Benazza R., Hernandez-Alba O., Cianférani S., Guigliarelli B., Burlat B. and D'Autréaux B.. **Fe-S cluster assembly by fusion of two [1Fe1S] intermediates into a bridging [2Fe2S] cluster mediated by IscU dimerization**. *Submitted to Nature Chem. Bio.,* **2023**.

List of presentations:

Oral communications

1. Benazza R., Koutsoupetras I., Chaubet G., Hernandez-Alba O. and Cianférani S. **SEC-MS in denaturing conditions (dSEC-MS) for in-depth analysis of rebriged mAbs**. *Journées GDR ChemBio.* June **2023**, Strasbourg, France.
2. Benazza R., Papadakos G., Diemer H., Cotton G., Cianférani S. and Hernandez-Alba O. **Full characterization of single-domain antibody-drug-conjugates (sdADC) with assessment of modification and conjugation sites using top-down mass spectrometry (TD-MS)**. *Journées Club Jeune de la FPS*. Avril **2023**, Strasbourg, France.
3. Benazza R., Hernandez-Alba O. and Cianférani S. **New MS-based approaches for the characterization of PDCs**. *Journées Pratiques Françaises des Sciences Analytiques (JPFSA).* April **2022**, Marrakech, Morocco.

Poster communications

1. Benazza R., Papadakos G., Cotton G., Hernandez-Alba O. and Cianférani S. **TD-MS approaches for straightforward characterization of an anti-EGFR single-domain antibody-drug conjugates**. .*JFSM*. September **2023**, Nantes, France.
2. Benazza R., Papadakos G., Cotton G., Hernandez-Alba O. and Cianférani S. **Full characterization of sdAb conjugates using top-down mass spectrometry (TD-MS)**. Analytics. June **2022**, Nantes, France.
3. Benazza R., Beaumal C., Hernandez-Alba O. and Cianférani S. **Optimization of MD-MS analysis and data treatment to improve the characterization of mAbs.** I*NSERM*. January **2022**, Online presentation.

Table of contents

Main abbreviations

AAV	Adeno Associated Virus
AcDC	Anticaline-Drug Conjugate
ADC	Antibody-Drug Conjugate
ADCC	Antibody-Dependent Cellular Cytotoxicity
ADCP	Antibody-Dependent Cellular Phagocytosis
AEX	Anion Exchange Chromatography
AF488	Alexa Fluor 488 dye
AGC	Automatic Gain Control
AI-ECD	Activated Ion Electron Capture Dissociation
AI-ETD	Activated Ion Electron Transfer Dissociation
AI-UVPD	Activated Ion Photodissociation
AMP	Ampicillin
ANC	Antibody-Nanoparticles Conjugate
Anti-EGFR	anti-Epithelial Growth Factor Receptor
ATD	Accurate Drift Time
avDAR	average Drug-to-Antibody Ratio
BBS	Borate Buffered Salin
BEH	Bridged Ethylene Hybrid
BGE	Background Electrolytes
BiTE	Bi-specific T-cell Engager
bsAb	Bispecific Antibody
BSB	Binding Site Barrier
BSM	Binary Solvent Manager
BU-MS	Bottom-Up Mass Spectrometry
CBD	Chitin Binding Domain
CCS	Collison Cross Section
CDC	Complement Dependent Cytotoxicity
CDMS	Charge Detection Mass Spectrometry
CDR	Complementarity-Determining Region
CE	Capillary Electrophoresis
CEX	Cation Exchange chromatography
CID	Collision Induced Dissociation
CIU	Collision Induced Unfolding
CL	Constant domain of the light chain
CQA	Critical Quality Attribute
CTDP	Consortium of Top Down Proteomics
CV	Compensation Voltage
CZE	Capillary Zone Electrophoresis
DLD	Drug Load Distribution
DMF	Dimethylformamide
dSEC	denaturing Size Exclusion Chromatography
DTT	Dithiothreitol
E.coli	Escherichia coli
EBZ	Ethynylbenziodazolone reagent
ECD	Electron Capture Dissociation
ECnoD	Electron Capture no Dissociation
ECuvPD	Electron-Capture Dissociation with Photodissociation
EMA	European Medicine Agency
EPL	Expressed Protein Ligation Technology
ESI	Electrospray Ionization
ETciD	Electron-Transfer with Collision Induced Dissociation
ETD	Electron-Transfer Dissociation
EThcD	Electron-Transfer Higher-Energy Collision Dissociation
ETnoD	Electron-Transfer no Dissociation
F(ab')2	two Fab fragments linked together
Fab	Fragment Antigen-Binding region
Fc	Fragment Crystallizable
FcR	Fc receptor
Fd	Fragment with variable and constant domains of HC
FDA	Food and Drug Administration
FPOP	Fast Photochemical Oxidation
FWHM	Full Width at Half Maximum
GlcNAc	N-Acetylglucosamine
GSH	Glutathione
HC	Heavy Chain
HCAb	Heavy Chain only Antibody
HCD	Higher-energy Collision Dissociation
HDX	Hydrogen/Deuterium Exchange
H-ESI	Heated Electrospray Ionization
HIC	Hydrophobic Interaction Chromatography
HMWS	Higher Molecular Weight Species
HOS	Higher Oder Structure
ICH	International Council for Harmonization
IEX	Ion Exchange Chromatography
IgG	Immunoglobulin G
IM-MS	Ion Mobility Mass Spectrometry
INN	International Non-proprietary Names
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRM	Ion Routing Multipole
IRMPD	Infrared Multiple Photo-Dissociation

iSCAT	Interferometric Scattering Microscopy
ISD	In Source Decay
iSD	In Source Dissociation
IT	Ion Trap
IW	Isolation Window
LB	Luria Bertani media
LC	Liquid Chromatography
Lc	Light chain
LiP-MS	Limited Proteolysis Mass Spectrometry
LMWS	Low Molecular Weight Species
mAb	Monoclonal Antibody
MAC	Membrane Attack Complex
MALDI	Matrix-Assisted Laser Desorption Ionization
MAM	Multi Attribute Monitoring
MaxEnt1	Maximum Entropy 1
MD-MS	Middle Down Mass Spectrometry
MMAE	Monomethyl Auristatin E
MMAF	Monomethyl Auristatin F
MP	Mass Photometry
MS	Mass Spectrometry
MW	Molecular Weight
NCE	Normalized Collision Energy
NGM	Next-Generation Maleimides
NK	Natural Killer cells
nMS	Native Mass Spectrometry
nTDMS	Native Top-Down Mass Spectrometry
OT	Orbitrap
PBS	Phosphate Buffered Saline
PD	Pyridazinedione
PDB	Pyrrolobenzodiazepines
PDC	Protein-Drug Conjugate
PEO	Polyethylene Oxide
pI	Isoelectric Point
ppm	Part Per Million
PPP	Perfluoroperhydrophenanthrene
PTCR	Proton-Transfer Charge Reduction
PTM	Post-Translational Modification
QC	Quality Control
Q-ToF	Quadrupole-Time-of-Flight
rAb	Recombinant Antibody
RF	Radio Frequency
Rs	Resolution
scFv	Single-Chain Variable Fragment
sdAb	Single Domain Antibody
sdADC	Single Domain Antibody-Drug Conjugate
SEC	Size Exclusion Chromatography
TACT	Targeted Anti-Cancer Therapies
TCEP	Tris(2-carboxyethyl)phosphine
TCO	Trans-Cyclooctene
TD-MS	Top Down Mass Spectrometry
TIC	Total Ion Chromatogram
ToF	Time-of-Flight
TsAb	Trispecific Antibody
UV	Ultra Violet
UVPD	Ultra Violet Photodissociation
Vc	Variable Chain
VH	Variable domain on the heavy chain
VHH	Variable Heavy Chain Antibody
VL	Variable domain on the light chain
WHO	World Health Organization
XL-MS	Cross-Linking Mass Spectrometry

General introduction

General introduction

Developments in cancer therapy have led to the emergence of new classes of therapeutic proteins [1]. Monoclonal antibodies (mAbs) [2], single-domain antibodies (VHHs) [3], multi-specific mAbs [4] or immunoconjugated antibodies (ADCs for antibody-drug conjugates) [5-6] are the most representative class. The latter are based on the covalent coupling of mAbs to cytotoxic molecules via a spacer arm (the linker) [6-7] and are distinguished from mAbs by their enhanced therapeutic efficacy and specificity. Characterizing the resulting ADCs/PDCs (PDCs for protein drug conjugates) represents a real analytical challenge, due to their high molecular weight, intrinsic variability (post-translational modifications, PTMs) and the additional heterogeneity resulting from the bioconjugation process [6, 8].

Mass spectrometry (MS) coupled with liquid chromatography (LC) has the potential to overcome these challenges and cope with the structural complexity of these proteins in order to assess the various critical quality attributes (CQAs) required by regulatory agencies [9-13]. There are therefore several approaches and methodologies for the characterization of these proteins at several levels (intact, middle or peptide level). Native MS (nMS) is a technique based on electrospray ionization (ESI), which transfers the solubilized protein into the gas phase of the mass spectrometer in a volatile, non-denaturing buffer (typically ammonium acetate), thereby maintaining non-covalent interactions [14]. This requires buffer exchange that can be carried out manually, or in a more automated way by on-line coupling to size exclusion chromatography (SEC-nMS) [15-16]. This is a key step in saving analysis time and adding another dimension to characterization by separating species according to their hydrodynamic volumes. The information obtained by SEC-nMS provides insights on: i) the quantity of unconjugated mAb (D0), ii) the average number of conjugated drugs per antibody (avDAR, average drug-to-antibody ratio), iii) drug distribution (DLD) and iv) size variants [17]. nMS can also be coupled with ion exchange chromatography (IEX) to separate and characterize mAbs according to their net charge difference [18-19].

More recently, approaches for sequencing intact proteins (without digestion), known as top down MS (TD-MS)[20-22], have been reported which, in addition to accurate mass measurement, enable proteins to be fragmented directly at the intact level [23-27] or after controlled digestion to obtain 30-50 kDa subunits (middle, MD-MS) [25, 28-30]. Thus, these approaches make it possible to characterize charge variants (e.g. PTMs) and identify specific drug positions directly, without extensive sample preparation. It is in this context of developing methodologies based on state-of-the-art MS techniques to improve the characterization of conjugated proteins (from mAbs or other proteins) that my thesis project is situated, where my main **PhD objectives** are:

i) **Improving LC-MS coupling techniques for rapid, automated characterization of ADCs/PDCs**

ii) **Developing new TD/MD-MS approaches for the characterization of the primary structure of these bioconjugates.**

The manuscript outlining this work is organized in three forthcoming parts:

- The **first part** offers a brief overview of immunoglobulins G (IgGs) [31] including their structure, their subclasses and their different mechanisms of actions. Another section is devoted to mAb therapeutics with a focus on mAb-based formats dedicated to cancer therapy with their market development [1]. A separate chapter is focused on the different analytical strategies used for mAb-based formats characterization.
- The **second part** focuses on evaluating a new benchtop LC-MS platform (the BioAccord LC-MS system, Waters) for the characterization of different mAb therapeutics. The first chapter discusses the optimization and implementation of automated SEC-nMS and CEX-nMS approaches using reference mAbs and ADCs. The second chapter presents an application of the automated SEC-nMS workflow for the characterization of site-specific ADCs and PDCs derived from Ugi conjugation [32]. The third chapter illustrates the limitations of the BioAccord instrument for the characterization of bsAbs, with a more comprehensive characterization on another MS instrument (Synapt G2 HDMS, Waters). Finally, the last chapter describes the development of a SEC-MS method in denaturing conditions (dSEC-MS) specifically adapted for the characterization of rebriged mAb-based conjugates.
- The **third part** is centered on the set-up of TD- and MD-MS approaches on a state-of-the-art MS instrument (Orbitrap Eclipse™ Tribrid™ MS, Thermo Fisher Scientific) comprising different activation techniques. The first chapter is dedicated to the optimizations of a MD-MS approach for the characterization of a reference mAb sample (trastuzumab) with assessment of the N-glycosylation PTM. The second chapter is focused on a detailed study for the characterization of a nanobody conjugate (anti-EGFR nanobody), produced during a my secondment at Almac Discovery (Edinburgh, UK), using native MS and TD-MS. nMS was used to assess homogeneity, DLD and avDAR. TD-MS using a combination of complementary fragmentation techniques (HCD [33], ETD [34] and UVPD [35]) allowed validation of the primary structure of the nanobody, including precise localization of the conjugation site and assessment of disulfide bridge pairing.

Part I: Introduction to monoclonal antibodies (mAbs), mAb-based products and their analytical strategies

Chapter 1: Introduction to monoclonal antibodies and mAb-based formats

1. Immunoglobulins G (IgGs)

In the late 19th century, immunoglobulins (Igs) or antibodies were introduced by von Behring and Kitasato to describe proteins generated by the immune system as a response to a pathogen's antigen. Those glycoproteins secreted by B- lymphocytes plasma cells, play an important role in recognizing the B cells on specific antigens, which allow them to bind to these entities in order to neutralize or eliminate them [31].

There are five classes of immunoglobulins identified (IgG, IgD, IgE, IgA and IgM), which in addition to their varied molecular size, they differ by the type of heavy chains they possess namely: γ-chains, δ-chains, ε-chains, α-chains and μ-chains for IgG, IgD, IgE, IgA and IgM, respectively [36-37]. This difference in heavy chains induces considerable variabilities in the function and immune response of those Ig types. Among these isotypes, IgGs are the most abundant in the body representing up to 80% of the serum, therefore it is the most extensively studied class [36] (**Figure 1**).

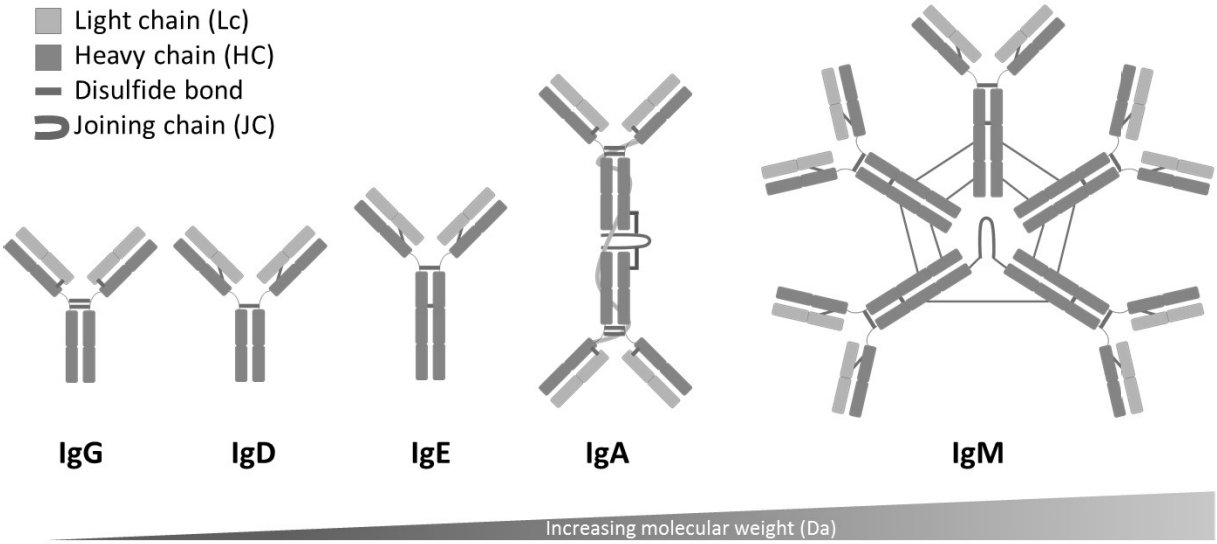


Figure 1: The different IgG isotypes including IgG, IgD, IgE, IgA and IgM depicted from the lowest (left) to the highest (right) molecular weight Ig. Adapter from Arnold *et al.* (2007) [38]

1.1 IgGs structure

IgGs are large monomeric proteins composed of four polypeptide chains that form a flexible Y shape. Two polypeptide chains are identical light chains (Lc, 25 kDa) and the two others are identical heavy chains (HC, 50 kDa). The light chain is composed of one variable (VL) and one constant (CL) domain, whereas the heavy chain contains one variable (VH) and three constant (CH1, CH2 and CH3) regions. Both polypeptide chains pairs are linked by a certain number of inter-chain disulfide bonds, depending on the IgG subclass, that form a rigid segment of the hinge region.

Typically, the region containing one variable (VH) and one constant (CH1) domains on the heavy chain, is referred to as the Fd fragment (25 kDa) due to its diversity and variability. The combination of this fragment and the Lc, composes the fragment antigen-binding region (Fab, 50 kDa), which is the region that binds to antigens through the arm tips of the variable domains. Each variable domain contains three complementarity-determining regions (CDRs: CDR1, CDR2 and CDR3) randomly arranged on the amino acid sequence, which constitutes a paratope. On a F(ab')2 fragment, which is the term that designates two Fab fragments linked together through the hinge region's disulfide bridges, six CDRs are identified. These hypervariable regions play a crucial role in the binding affinity enhancement of antibody-antigen interaction [39].

The remaining fragment that contains constant domains of the heavy chain (CH2 and CH3), represents the fragment crystallizable region (Fc, 25 kDa). The Fc fragment plays a crucial role in ensuring the mediation of effector functions by either binding to Fc receptors (FcR) on effector cells [40], or activating other immune mediators. The effector functions are modulated by the presence of glycans associated to the Fc fragment (**Figure 2**).

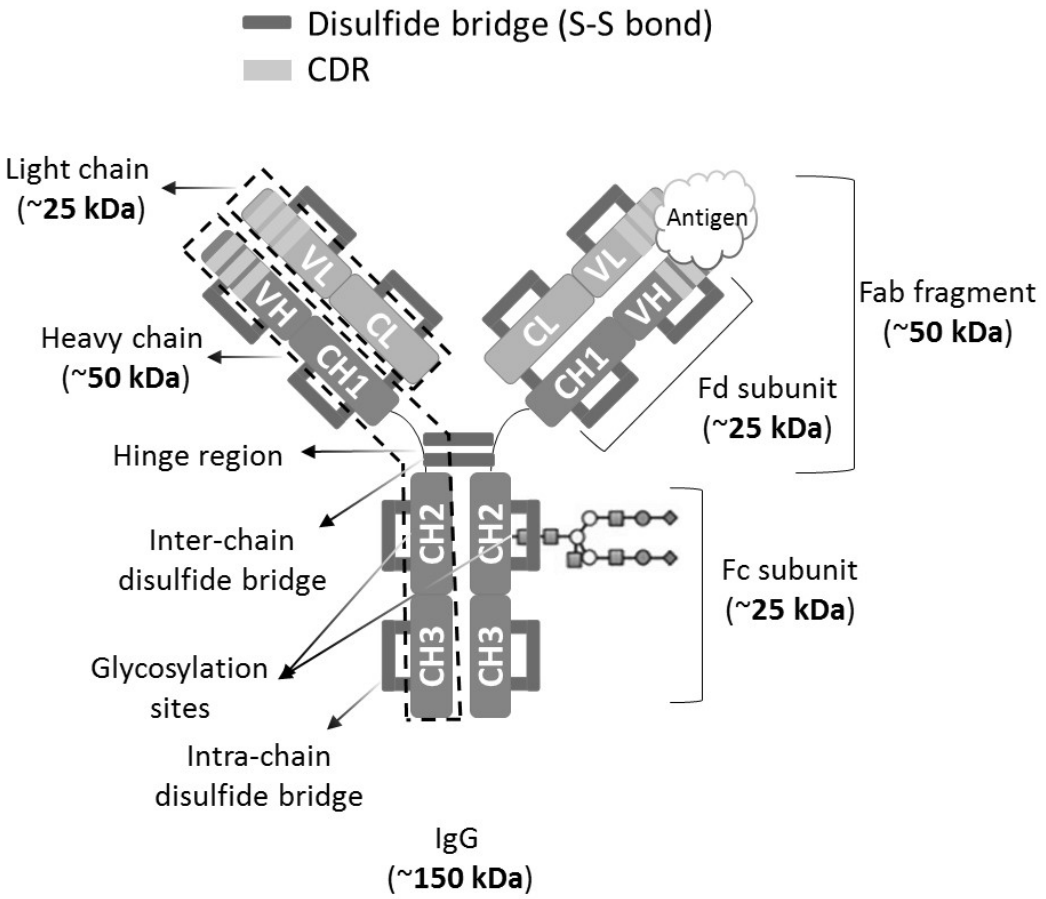


Figure 2: Structure of an IgG illustrated with an IgG1. The heavy chain containing constant domains (CH1, CH2 and CH3) with a variable domain VH is depicted in green. The light chain including a variable domain VL and a constant domain CL is depicted in orange. The glycoforms are located on the Fc part of the heavy chain. Disulfide bonds are represented in blue lines and CDRs are represented in pink rectangles. Adapted from Loureiro *et al.* (2015) [41]

1.2 IgGs subclasses

IgGs can be divided into several subclasses based on their effector functions and binding properties [42]. These variations are due many structural differences, mainly the number and the position disulfide bonds; while IgG3 contains twelve inter-chains disulfide bridges that stabilizes their Fc regions, the number of those S-S linkages is reduced to only three for IgG2 and two for IgG1 and IgG4. The disulfide bridges linking both heavy chains on IgG4 are particularly flexible, which induces a phenomenon called Fab-arm exchange, leading to asymmetrical bispecific antibodies generation [43]. Another dissimilarity lies in the position of the heavy chain's CH1 cysteine that is responsible for the intra-chain linkage with the carboxy-terminal light chain. This cysteine is located in position 220 for IgG1, where for IgG2, IgG3 and IgG4, it is in the 131st position. Moreover, IgG2 subclass exists in three isoforms namely IgG2A, IgG2B and IgG2A/B, depending on their disulfide bridges type and position (**Figure 3**). Owing to their high abundance in human serum (about 65% of the total IgG), their high stability, their less aggregation and their shorter serum half-life, IgG1 are the most frequently used for drug development [36, 40].

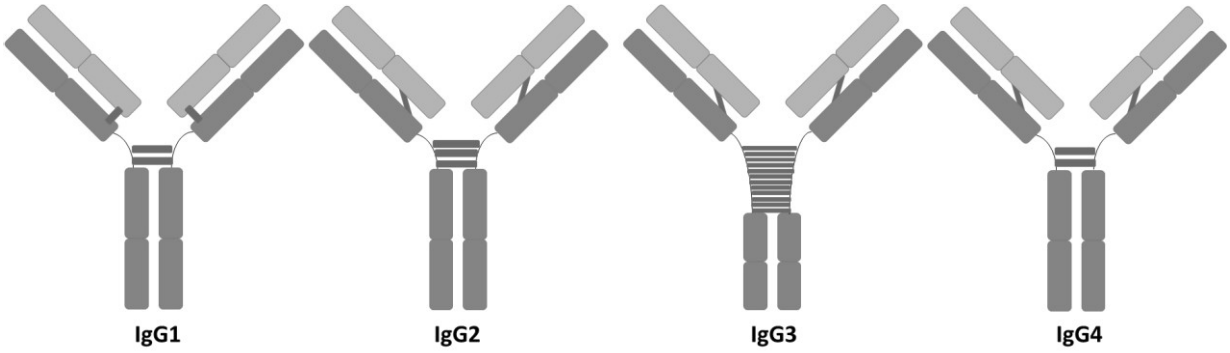


Figure 3: IgGs different subclasses including IgG1, IgG2, IgG3 and IgG4. The main difference between the four subclasses is the number and position of disulfide bonds depicted in blue. Light and heavy chains are depicted in orange and green, respectively. Adapted from Muhammad *et al.* (2020) [36]

1.3 IgGs glycosylation

All IgGs contain N-linked oligosaccharides (N-glycans) attached to their Asn297 residue at the Fc domain, hence the name glycoproteins [44]. The heptasaccharide core of glycans composed from two N-Acetylglucosamine (GlcNAc) and three mannose, could be attached to 0, 1 or 2 galactose moieties forming GO, G1 and G2 glycoforms, respectively [45]. When fucosylated at the asparagine-attached galactose, they are defined as G0F, G1F and G2F. Sialylated glycoforms, which represents minor species in usual IgGs, are named G2S1, G2S2, G2S1F and G2S2F (**Figure 4**). The role of glycosylation relies in tuning the immunological properties of antibodies, and therefore affecting their therapeutic efficacy [38]. For instance, an increase of fucosylation decreases the binding affinity of antibodies to FcγIIIa receptor, which drastically reduces the antibody-dependent cellular cytoxicity (ADCC) [46]. Therefore,

many recent advances focus on modulating glycans in antibodies, which is known as glycoengineering of antibodies [45, 47-49]. Traditional glycoengineering strategies are based on genetic approaches that aim either to modify the structure of glycans through the knockout of a specific enzyme to improve their properties, or to overexpress a key enzyme to increase the antibodies defense mechanisms.

Another promising strategy is the development of glycan remodeling [45, 49], this is achieved by chemistry based technologies that lead to provide specific glycosylation profiles with desired properties [50]. Moreover, glycoengineering strategies have been developed to specifically conjugate antibodies by addition of unnatural sugars or exploiting the existing N-glycosylation site [45].

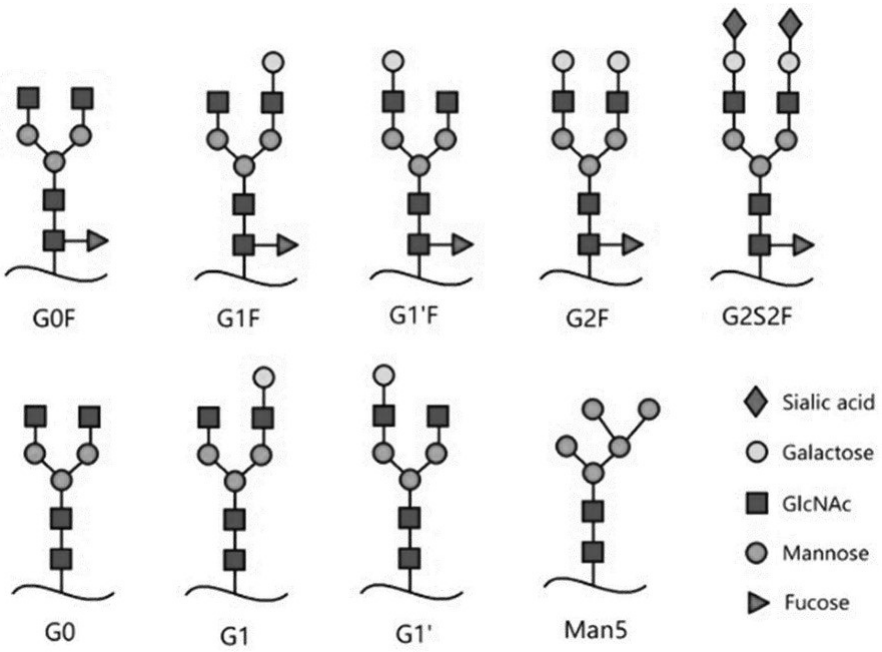


Figure 4: Major N-glycoforms types on mAb therapeutics. Adapted from Xu *et al.* (2021) [51]

1.4 IgGs disulfide bonds

Disulfide bonds are covalent linkages between sulfur of two cysteine residues (S-S linkage) [52]. For antibodies, they play an important role in fixing and stabilizing the tertiary structure and can be divided in two types: disulfide bridges within the same chain called intra-chain, and disulfide bridges linking two polypeptide chains referred to as inter-chain disulfide bonds [52]. Each IgG contains a total number of twelve intra-chain disulfide bonds, while the number of inter-chain ones differs from one subclass to another (**Figure 3**). Those two types presents different level of solvent exposure [52]; inter-chain are highly exposed making the involved cysteines more reactive, thus easily accessible for cysteine site-specific conjugations. Consequently, the degradation of the disulfide bonds result in sulfhydryls (R-SH) releasing [53-54], that are mainly detected in variable domains where their level is higher than in constant domains. The presence of these sulfhydryls affects negatively the stability of IgGs, but no structural

changes were highlighted in literature. Disulfide bonds could be also decomposed under basic conditions, through the β-elimination mechanism which transforms one of the involved cysteines into a persulfide and a dehydroalanine that forms a non-reducible thioether upon Michael-like addition [55]. Finally, a rare type of linkages is trisulfide bonds found in A, B and A/B forms of IgG2 (**Figure 5**). They are formed through reaction of an intact disulfide bond with hydrogen sulfide (H_2S) obtained from degradation of cysteines. Although these major PTMs do not affect the stability or the antigen binding, they are carefully monitored during the development of antibody conjugates.

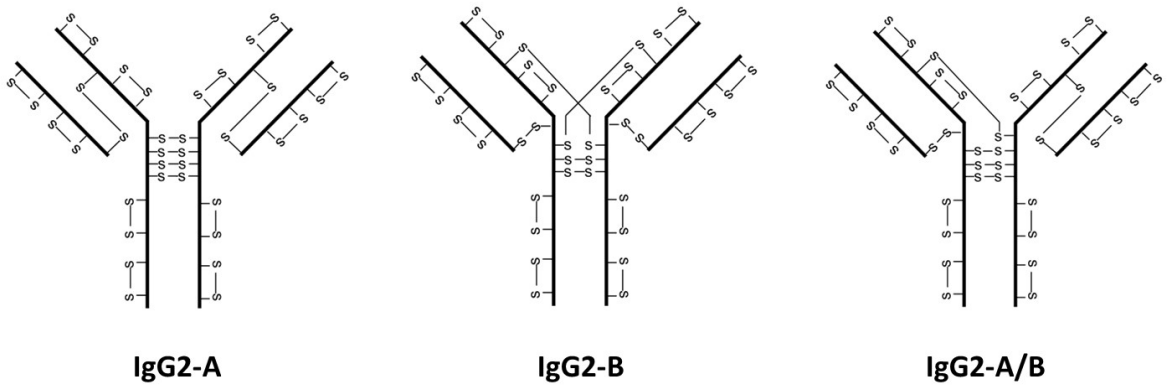


Figure 5: Non-classical disulfide bonds including the different IgG2 sub-isotypes namely; IgG2-A, IgG2-B and IgG2-A/B. The main difference between these types is the disulfide bond position. Adapted from Liu *et al.* (2012) [52]

1.5 Antibodies mechanism of action

Antibodies play an important role by protecting the body against infections through different mechanisms. The key functions include: i) Direct effects upon the recognition and binding to specific antigens, which leads to either neutralizing the latter molecules or blocking their ability to enter host cells, ii) Indirect effects that activate complement mechanisms including antibody-dependent cellular cytotoxicity (ADCC), antibody-dependent cellular phagocytosis (ADCP) and complement-dependent cytotoxicity (CDC), which lead to destruction of antigens [56]. In fact, each complement mechanism is mediated differently. ADCC is mediated by natural killer (NK) *via* their (FCγRIII) receptor that binds to the Fc region and induces cytotoxic granules release that kill the target cell [57]. Unlike ADCC, ADCP induces macrophage upon target cell phagocytosis [58]. CDC targets cell lysis also *via* antibodies binding to C1q complement protein, which induces formation of membrane attack complex (MAC) that consists of complement proteins C5b-C9 that generate membrane in the pores and lead to cell destruction [59]. Finally, one of the main functions is iii) Drug delivery of multiple components, where the antibody acts as a drug carrier (**Figure 7**).

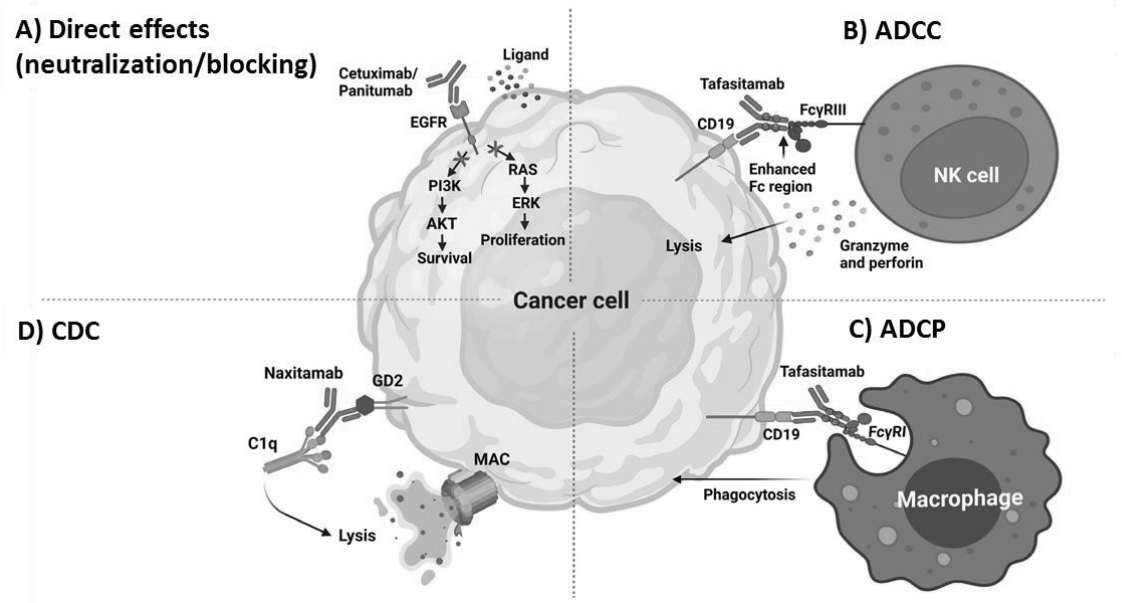


Figure 6: Antibodies different mechanisms of actions in cancer therapy resumed in four actions; (**A**) Direct effects that include neutralization or blocking effects, (**B**) ADCC mediation, (**C**) ADCP mechanism and (**D**) CDC mediation. Adapted from Nava *et al.* (2023). [60]

2. Monoclonal antibodies (mAbs): a better option for therapeutic purposes

In contrast to polyclonal antibodies, which are produced by several different immune cells and have affinity for the same antigen on different epitopes, monoclonal antibodies (mAbs) are considered as a better therapeutic option for their high specificity to a single epitope [36]. Since the development of the first mAb (Muromonab CD3, a murine mAb) used in kidney transplant rejection prevention, several groups focused on the development of numerous mAbs from different sources [36, 61]. Murine mAbs development was limited by the risk of immunogenicity they presented [62], and lead to new technologies generation of predominantly or fully humanized mAbs [63]. While chimeric antibodies are mAbs composed of a constant region from human sequences [64], humanized mAbs contain all human sequences except the antigen binding CDRs which are derived from the mouse. Finally, fully human antibodies are composed of whole human sequences and have significantly reduced immunogenicity [65](**Figure 7**).

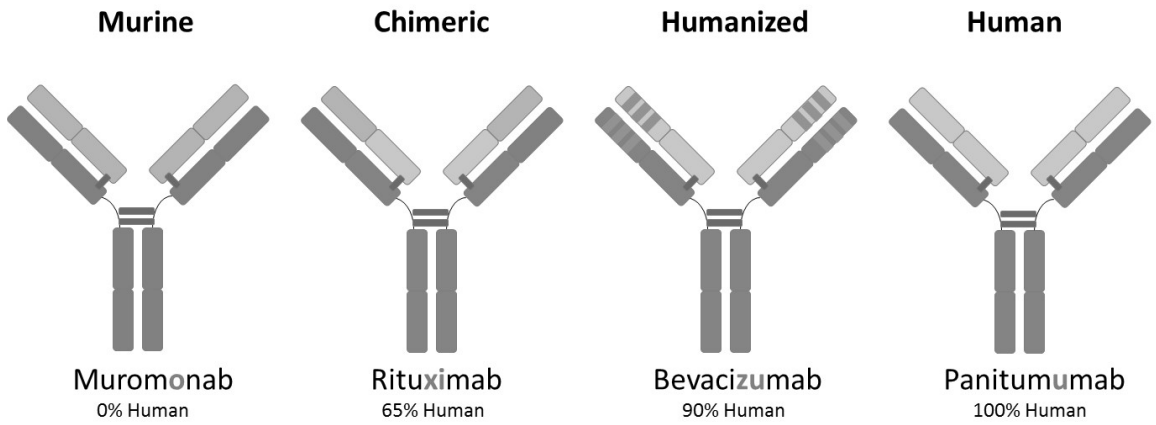


Figure 7: Different sources of mAbs leading to different mAbs formats; murine, chimeric, humanized and fully human mAbs. The immunogenicity in humans decreases with the humanization of mAbs. Murine regions are depicted in orange and human region are in green. Adapted from Tan *et al.* (2016) [66]

2.1 MAbs nomenclature

The appropriate nomenclature for antibody-based therapeutics is important for clinical development. The International Nonproprietary Names (INN) program of the World Health Organization (WHO), implemented a nomenclature for mAbs which employed at the beginning a prefix, a stem and a suffix [2]. The stem that was used for all monoclonal antibodies containing an immunoglobulin variable domain designed against a specific target was "mab", placed as a suffix. In 2021, the INN replaced the suffix –mab with four new stems to differentiate the multiple types of therapeutics, employing a suffix of –tug for unmodified IgGs, -bart for artificial IgGs, -ment for IgG fragments and –mig for multi-specific IgGs [67].(**Table 1**).

Table 1: Nomenclature of mAbs based on the International Nonproprietary Names (INN) adopter in October 2021. Adapted from Guimaraes *et al.* (2022) [67]

Prefix	Sub-stem	Infix for target class	Suffix
Random	-ami-	Serum amyloid protein (SAP)/amyloidosis	-tug -bart -mig -ment
	-ba-	Bacterial	
	-ci-	Cardiovascular	
	-de-	Metabolic or endocrine pathways	
	-eni-	Enzyme inhibition	
	-fung-	Fungal	
	-gro-	Growth factor and growth factor receptor	
	-ki-	Cytokine and cytokine receptor	
	-ler-	Allergen	
	-sto-	Immunostimulatory	
	-pru-	Immunosuppressive	
	-ne-	Neural	
	-os-	Bone	
	-ta-	Tumor	
	-toxa-	Toxin	
	-vet-	Veterinary use	
	-vi-	viral	

2.2 MAb-based formats

The emergency of mAbs have revolutionized the field of therapeutics, contributing in the treatment of various cancers, autoimmune disorders, infectious, cardiovascular and neurological diseases. Considering the continuous interest in developing improved and highly specific mAb treatments [68], recent advances focused on engineering new mAb-based formats termed "biobetters" [5, 69]. This plethora of mAb-based therapeutics includes biosimilars, antibody-drug conjugates (ADCs), bispecific antibodies (bsAbs), recombinant antibodies (rAbs), and heavy chain-only antibodies (HCAbs). Smaller mAb-based fragments are also isolated such as antigen binding fragments (Fab or F(ab')2), single-chain variable fragments (scFv) or even smaller single-domain antibodies (sdAb) [68].

a) Antibody-drug conjugates (ADCs)

In the early 20th century, Paul Ehlich first introduced antibody-drug conjugates (ADCs) as "magic bullets" that can directly access desired targets and kill cancer cells without harming normal cells. Consequently, these alternative cancer treatments have been the fastest growing class in oncology with 14 ADCs approved in the market by regulatory agencies [6-7, 70].

i. Structure

Typical ADCs are formed from three key components; a mAb attached covalently to a cytotoxic drug (payload) *via* a spacer arm called a "linker" [7]. Upon the choice of the target antigen expressed on tumor cells, the ideal ADC relies on each building block choice, which can affect the final efficacy and safety of the ADC in question:

- **MAb**: this moiety is the one that is responsible for the specific binding between antigens and the target cell. The high antibody-antigen affinity results in faster internalization. However, this is challenging for solid tumors due to existence of binding site barrier (BSB), where the strong binding of the antibody and the antigen results in trapping the ADC near the blood vessels that are away from tumor cells [6]. Consequently, a reasonable affinity is often optimized prior to conjugations. Moreover, IgG1 are the most commonly used for ADCs production as they are the most abundant in serum and have higher half-life (21 days *versus* only 7 for IgG3, for example). Finally, the size of used antibodies is highly critical as large molecular weight mAbs (~150 kDa) are often challenging for blood penetration, thus many groups focus on developing miniaturized antibodies for further conjugation.
- **Cytotoxic payload**: this part is the cargo that releases toxicity into cancer cells after ADCs internalization. Taking into consideration that only 2% of the ADC is able to reach tumor sites, payloads should be highly effective [71]. Moreover, a cytotoxic drug should be stable and should have available functional groups react with the antibody *via* the linker. The majority of cytotoxic payload derives from two major families: tubulin inhibitors and DNA damaging agents. Tubulin inhibitors include: i) auristatin derivatives that act as tubulin promoters and perturb microtubule growth (the component responsible

for cell division and rapid proliferation of tumor cells), among which monomethyl auristatin E and F (MMAE and MMAF, respectively.) are the main examples and ii) maytansinoid derivatives that block the formation of mature microtubules, which includes DM1 and DM4 drugs [72-73]. DNA damaging agents are more effective as they reach picomolar levels and include: i) calicheamicin which breaks the DNA double strand (*e.g.* calicheamicin γ1 used in Gemtuzumab ozogamicin), ii) duocarmycin that alkylates DNA (used in anti-RON ADCs), iii) topoisomerase I inhibitors that intercalate DNA (*e.g.* DXd used in trastuzumab Deruxtecan) and iv) pyrrolobenzodiazepines (PDB) that are used for DNA crosslink (used only in Loncastuximab tesirine) [74].

- **Linkers**: a linker is the important piece that attaches ADC to its payload, and should be able to release the drug at the target cancer cell without premature release in plasma and without ADC aggregation [6]. Among the two main linker types that are employed in ADC production, cleavable linkers are environment-sensitive, and include chemical cleavage linkers that release the payload through hydrolysis of the ADC at acidic pH (4.8-6.0), or through disulfide bonds reduction by intracellular glutathione (GSH). In the same cleavable linkers' family, enzyme sensitive linkers such as glucuronide or peptide based ones, are the most used in various ADCs (*e.g.* valine-citrulline linker used for Brentuximab vedotin) [75]. On the other hand, non-cleavable linkers are less dependent on the in-vivo environment as they directly release the payload after ADC internalization and lysosomal processing, along with the last amino-acid residue (*e.g.* thioether linkers used in ado-trastuzumab emtansine, T-DM1) [76] (**Figure 8**).

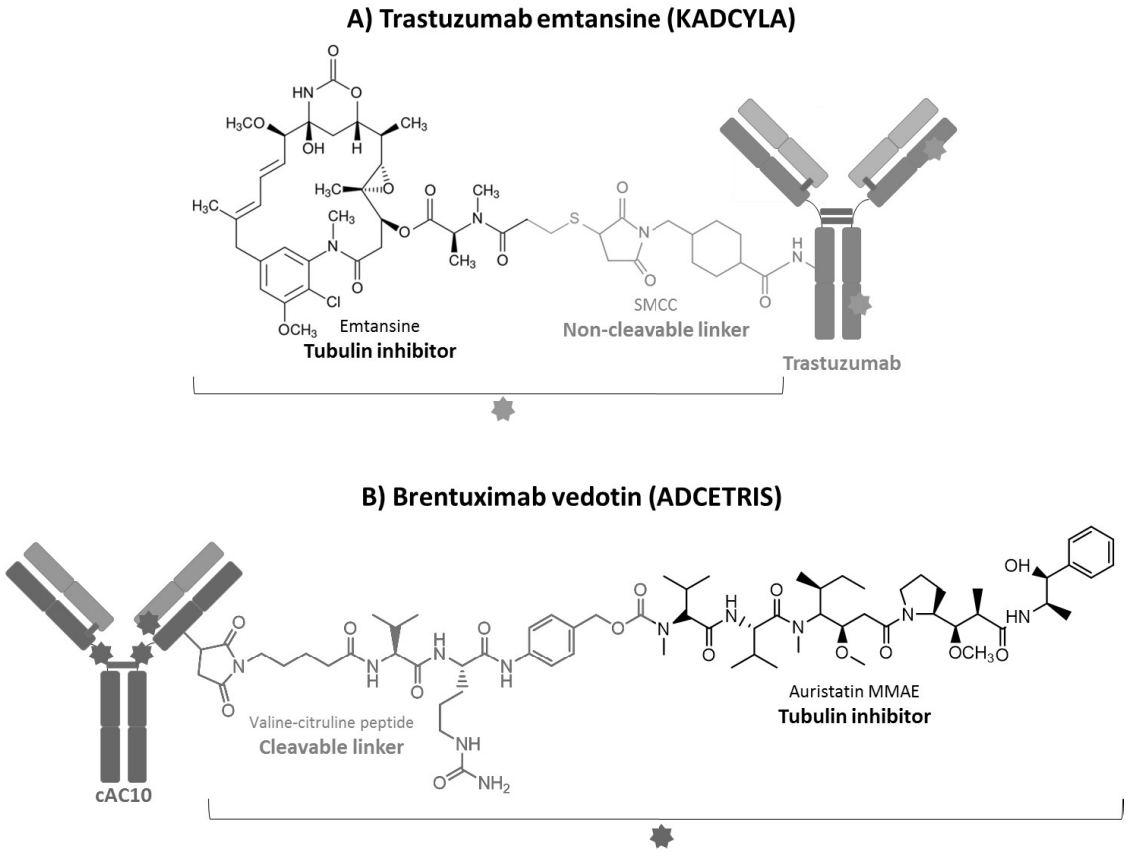


Figure 8: ADCs structure depicted with the example of trastuzumab emtansine (Kadcyla), a conjugated trastuzumab to a emtansine tubulin inhibitor through a non-cleavable linker (SMCC), and Brentuximab vedotin (ADCETRIS) which is a cAC10 mAb conjugated to an auristatin MMAE tubulin inhibitor through a cleavable linker (valine-citrulline peptide). Adapted from Beck *et al.* (2019) [10]

ii. Mechanism of action

ADCs act through three different mechanisms in the body; a first mechanism includes the internalization of the ADC upon binding to the cancer cell target. The ADC then releases its payload in the lysosome through endosome fusion, which induces cell apoptosis. Another mechanism involves the binding of the Fab fragment of the ADC to the antigen epitope, while the Fc region is bound to FcR on NK cells, which mediates direct cell killing effects namely; ADCC, ADCP and CDC effects. Finally, an ADC could inhibit the downstream signal transduction (cell signaling) of an antigen's receptor, by specifically binding to the same antigen's epitope, which induces cell apoptosis [6] (**Figure 9**).

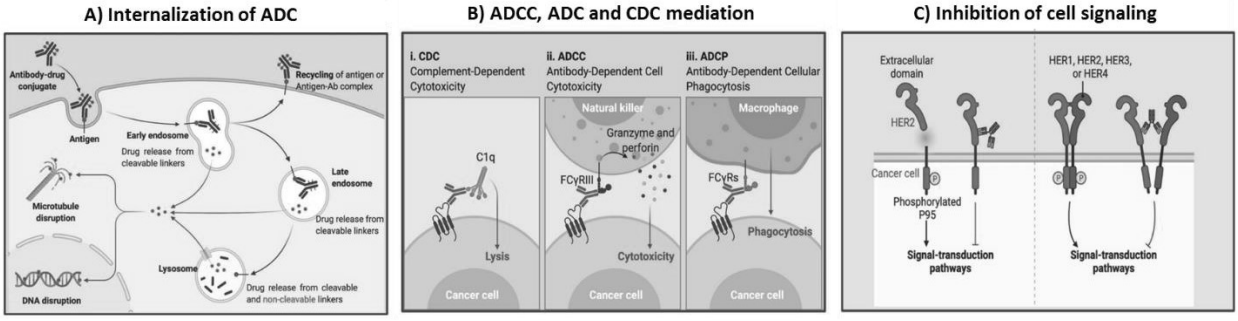


Figure 9: ADCs mechanisms of actions through three main functions: (**A**) internalization of the ADC, (**B**) Mediation of ADCC, ADC and CDC mechanisms and (**C**) Inhibition of cell signaling. Adapted from Fu *et al.* (2022) [6]

iii. Conjugation approaches

Beyond the choice of the antibody moiety, the linker and the payload, the conjugation strategy to attach these three components is highly important for an ideal ADC formation [77]. The conjugation method is the one that determines the number of linker/drug attached to the antibody, quantified by the average drug-to-antibody ratio (DAR), which directly affects the homogeneity, toxicity and efficacy of an ADC [10]. Higher DAR values (*e.g.* D8) lead to heterogeneous products with high cytotoxicity which results in solubility issues and high clearance [78]. In contrast, low DAR values (Unconjugated mAb D0 or D1, D2) lead to competitive internalization with higher DAR species, and thus reduce their therapeutic effect. Two major types of conjugations are being developed currently:

- **Stochastic conjugations**: Typically, stochastic conjugations are performed on pre-existing lysine and cysteine residues *via* suitable reactions, which are the most frequently used. Lysines can be reacted with activated carboxylic acid which results in an amide bond between the mAb and the payload (*e.g.* TDM1) [76], however due to the presence of many lysine residues on a usual mAb (~90 lysine), this strategy could lead to highly heterogeneous DAR8 ADCs. On the other hand, cysteine conjugation are based on reacting a prior reduced inter-chain cysteine residue with the linker/payload and results in more homogeneous DAR species (*e.g.* Brentuximab vedotin) [6, 75].

- **Site-specific conjugations**: many groups focused on site-specific ADCs conjugation as stochastic ones face many challenges including stability and heterogeneity issues [79]. A common approach is the introduction of engineered reactive cysteines that lead to production of DAR2 site-specific ADCs (*e.g.* anti-MUC16 ADC) [6]. A more attracting strategy is disulfide re-bridging conjugation [80-83], which aims at incorporating payloads between reduced disulfide bridges using cysteine-selective cross-linking reagents such as bissulfone reagents, next-generation maleimides (NGMs) and pyridazinediones (PDs) [82, 84-85]. Site-specific conjugation could be also reached upon introduction of unnatural amino acids that contain functional groups that are able to react with the liner/antibody moiety. As mentioned in the glycosylation paragraph, glycan remodeling and glycoconjugation are also now used to conjugate the N-glycan on the Fc region to payloads [45, 47, 49]. Finally, pClick technology was recently developed and uses a crosslinker that binds a chemically modified peptide to a specific antibody, which enable payload attachment [6] (**Figure 10**).

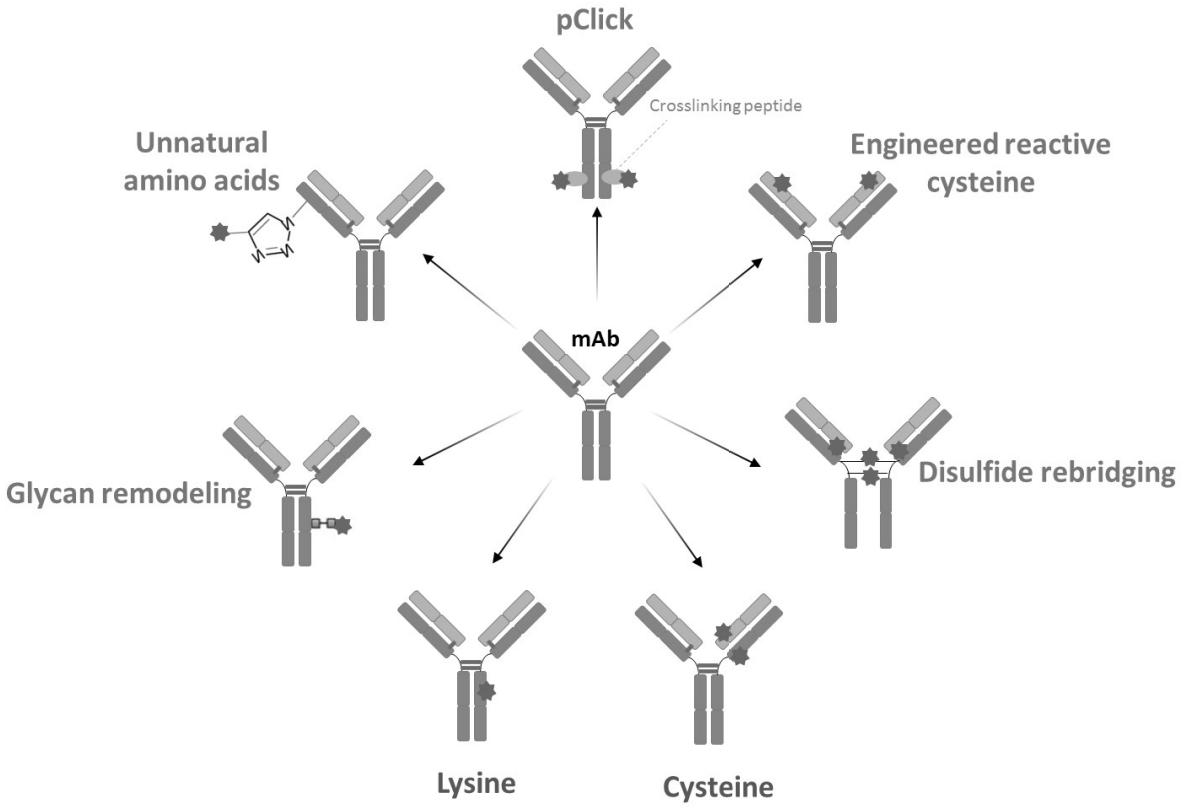


Figure 10: The various conjugation strategies for stochastic mAb conjugation (red) and site-specific conjugations (green). The cytotoxic drug is depicted in a red star. Adapted from Fu *et al.* (2022) [6]

Recent advances in drug discovery focus also on the development of Protein-Drug Conjugates (PDCs) using the same conjugation strategies [86]. Depending on the therapeutic goal, PDCs can provide several unique advantages over ADCs, particularly the smaller size that facilitates their penetration into solid tumors.

b) Bispecific antibodies (bsAbs)

Bispecific antibody (bsAb) term was first mentioned in 1960 by Nisonoff et al, to appoint to a man-designed antibody-based molecule that targets two specific antigens or epitopes [87]. BsAbs are an emerging class of therapeutics in cancer therapy and other diseases, with two bsAbs currently approved in the market: blinatumomab (for cancer therapy) and emicizumab (for bleeding disorder), and over 85 bsAbs in clinical phase [4]. They come in various formats designed for specific goals are can be distinguished in two main formats: IgG-like and non-IgG-like bsAbs [4, 87-88].

IgG-like bsAbs are derived from classical mAbs but with increased specificity thanks to the addition of a second antigen-binding region (a single-chain variable fragment, scFv) to the Fab or Fc region. Various formats exist based on the production technology namely: heterodimeric IgG, scFv-IgG and DVD-Ig bsAbs… etc. Each of these formats exhibit unique advantages, but the main advantage of these formats

stem mainly from their structural similarity to classical mAbs where the effector functions (ADCC and CDC) remain ensured thanks to the Fc-region mediation.

Non-IgG-like bsAbs are made of two scFvs that can bind two different specific antigens. This class encompasses a variety of formats such as; bispecific T cell engagers (BiTEs) that bring together T cells and the target cell, TandAbs allowing the simultaneous binding of two different antigens, or nanobody-based bsAbs that fuse two nanobodies with different antigen targets. Although these small bsAb-formats lack the Fc region that mediates the effector functions and have relatively lower serum half-life [89], they have multiple advantages compared to the intact bsAb forms. In particular, their enhanced tissue penetration and their low immunogenicity made them an exciting area in the cancer field research (**Figure 11**). Moreover, a trispecific antibody (TsAb) format is developed to engage three different epitopes by one single antibody, which provides improved targeting and increased potency compared to bsAbs [90].

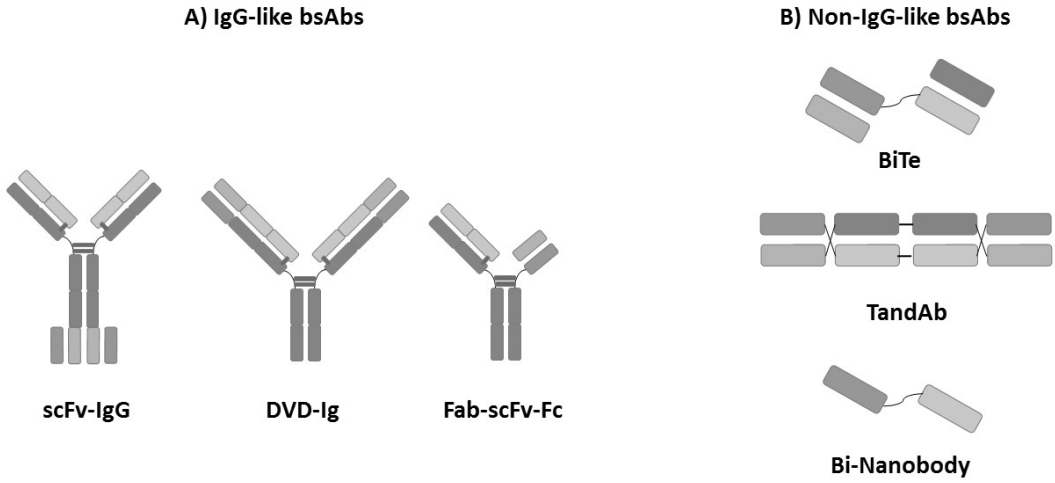


Figure 11: The two types of bispecifics family namely (**A**) IgG-like bsAbs and (**B**) non-IgG-like bsAbs. Adapted from Kang *et al.* (2022) [91]

c) Heavy chain only antibodies (HCAbs)

Heavy chain only antibodies (HCAbs) are derived from camelids and contain one single variable domain (VHH) and two constant domains (CH2 and CH3), unlike conventional antibodies that have the entire light chain and the CH1 domain [92]. Several studies showed the similar therapeutic characteristics of HCAbs compared to intact mAbs, moreover, their smaller size can enhance their ability to target epitopes that are challenging for conventional mAbs to access. However, one of the main challenges of HCAbs production, is the potential immunogenicity from non-human, which requires humanization prior to clinical development and thus extends the drug discovery process [63].

d) Single domain antibodies (sdAbs, VHH or nanobodies)

Nanobodies are the smallest fragments derived from HCAbs that represent ~1/10th its full size while preserving the characteristics of specificity and binding affinity. Their small size (~15 kDa), stability and solubility made them a more promising format than their mAb counterparts, in many fields such as diagnostics, imaging and therapeutics [3, 93-96]. The first nanobody-based therapeutic; caplacizumab for the treatment rare blood clotting disorders was approved in 2018 by the EMA, and paved the way to the development of various other nanobody drugs especially for cancer therapy [3, 96-97]. Similarly to antibodies, sdAbs can be conjugated to increase their potency and therapeutic efficiency. Unlike other mAb fragments (scFvs or Fabs), they have significantly higher stability and they are easily produced, which makes them interesting candidate for targeted therapies [3, 95-100].

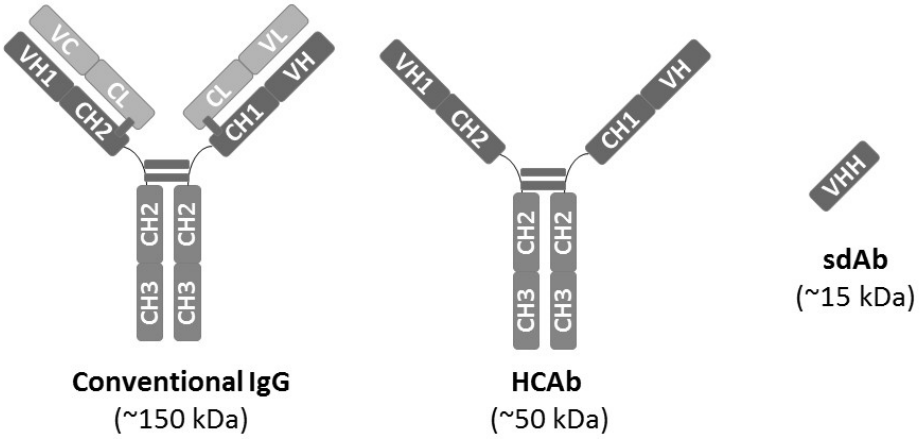


Figure 12: Comparison of conventional IgG (~150 kDa), HCAbs (~50 kDa) and sdAb (~15 kDa) or nanobodies derived from HCAbs. Adapted from Jin *et al.* (2023) [98]

e) Other small mAb-based formats

The current trend in biopharmaceutical field shifts towards the development of smaller mAb-based fragments, owing to their particular ability to retain full antigen-binding [89]. Among these products, we can distinguish Fc-fusion proteins that are composed of the Fc region of a particular antibody and a desired linked protein; they can bind to the neonatal Fc receptor (FcRn) which prevents the IgG degradation. Currently there are 13 Fc-fusion products approved in the European Union and United States with three biosimilar versions of etanercept. Fab or F(ab)2 formats are also one of the successful biotherapeutics, with currently three FDA-approved Fabs. Although the latter molecules are not the smallest mAb-based formats (~50 kDa and ~100 kDa), they allow easier production and good stability overall. The native structure they provide, does not require any engineering resource for an ideal linker thus they provide less immunogenicity [89] (**Figure 13**).

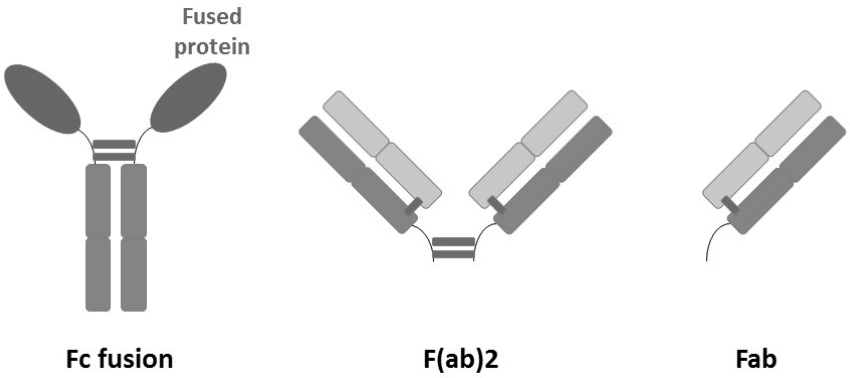


Figure 13: Other mAb-based formats including Fc fusion and smaller formats: Fab and F(ab)2. Adapted from Bates *et al.* (2019). [89]

f) Biosimilars

The term biosimilar designates the concept of "generic" mAbs, and was introduced in 2013 after the approval of the first infliximab (anti-TNFα) biosimilar by the European Medicine Agency (EMA) [101]. In fact, biosimilars are supposed to have no structure, function or clinical safety differences compared to their originator (reference), therefore they can only be considered better in term of affordability and market accessibility [11, 101]. With many mAb products reaching their patent expiry, up to 249 mAb biosimilars were approved by regulatory agencies [101]. This approval requires a compromise of quality and affordability, which is achieved by different biosimilarity studies [102].

Some biosimilars were found to be slightly different than their originator in term of glycans level, which is believed to have direct effect on the clinical efficacy as it is highly dependent on Fc-region glycosylation. However, health authorities were aware of these variations as the level of glycosylation found was not prone to high clinical relevance [103].

3. mAb Market

Following the initial approval of the first therapeutic mAb, muronomab-CD3 (Orthoclone OKT3) approved in 1986 by the US FDA, mAb-based formats have been emerging as the fastest growing therapeutic class. Designed for several therapeutic areas; neurological disorders, cardiovascular diseases [5]… etc., and particularly cancer [4, 6, 91, 104-108]. The record has been reached in 2021 and 2020 with 20 products first approved in Europe and in the US, as well as in other countries [106]. The market of mAb products have been significantly increasing for the last decade (**Figure 14)** with a global estimation of 162.47 billion USD in 2021, expected to reach up to 390.58 billion USD by 2030. The table showing the trend of mAb products growth in the market excludes the data related to biosimilars approval, which reached until today, 43 US FDA-approved biosimilars [1]. The latter format has been exceedingly growing in recent years with a big market estimated at 21.8 billion USD in 2022 and expected to reach up to 27.20 USD later this year (2023) (**Figure 14**).

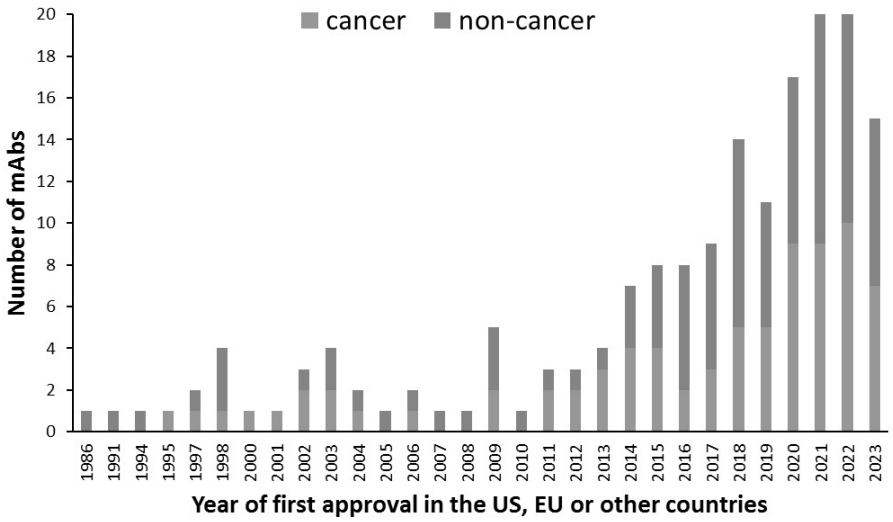


Figure 14: mAb products approved in the US, EU or other countries market for the treatment of various diseases including cancer. Cancer data are represented in orange. Figure based on data publicly available as of October 04, 2023. A searchable table of the figure data is available at .www.antibodysociety.org/antibody-therapeutics-product-data/

In addition to the mAb products approved by regulatory authorities, nearly 800 mAb therapeutics are in clinical development in oncology, which represents over 60% of the clinical pipeline [106]. This particular interest in mAb-products for cancer treatment is owed to the precise targeting of cancer cells thanks to the different mAb-formats prior discussed in this chapter. **Figure 15** highlights the increasing trend of bsAbs and ADCs in particular, in addition to full length mAbs for cancer therapy [1, 91, 106].

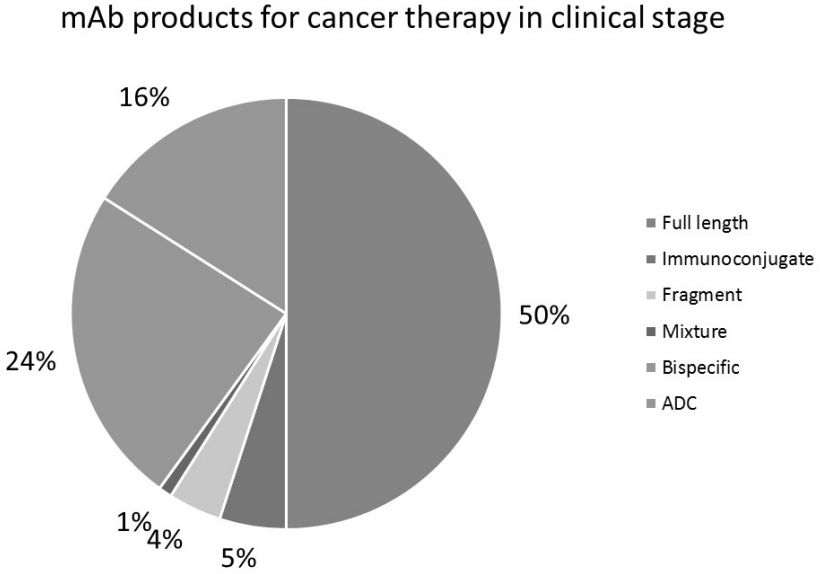


Figure 15: mAb products dedicated for cancer therapy, that are in clinical stage of trials. Many mAb-based formats are represented in different colors, including ADCs (pink), bsAbs (orange) and full length antibodies (green) that are significantly more compared to the other forms. Figure based on data publicly available as of October 04, 2023. A searchable table of the figure data is available at .www.antibodysociety.org/antibody-therapeutics-product-data/

In view of their considerable potency and major specificity, ADCs have been especially expanding in oncological field. Since Mylotarg (Gemtuzumab ozogamicin) was first approved in 2000 for the treatment of acute myeloid leukemia (AML) [109], it marked the beginning of ADC era in cancer-targeted therapy, with 14 ADCs approved worldwide and over 100 in the clinical stage (**Figure 16**) [1, 106, 110].

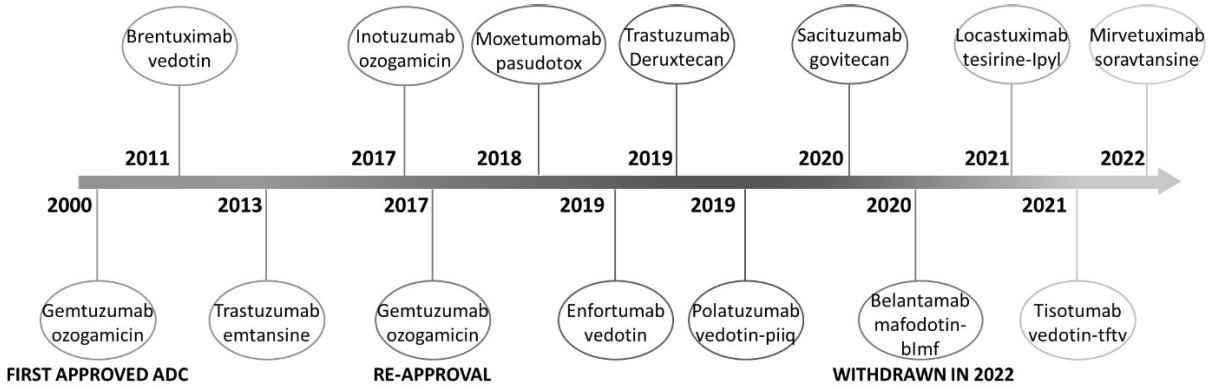


Figure 16: Timeline depicting the development and approval of ADCs over the past century since the approval of the first ADC: Mylotarg. Adapted from Fu *et al.* (2022) [6]

Due to the latest advances in biotechnology and the global healthcare needs especially in the cancer field, the development of mAb-based products is significantly growing. The competition race of making efficient and affordable mAb therapeutics constantly needs robust and fast analytical strategies to monitor and characterize their entities. Therefore, an increasing need of fast, robust and versatile on state-of-the-art instruments is also highlighted.

4. Targeted anti-cancer therapies (TACT) program

In our TACT program consortium, we are particularly focused on the next generation targeted-anti-cancer therapies. This European consortium brings together academic and industrial institutions with various background: chemistry, biochemistry and analytical chemistry. TACT includes 11 early state researchers (ESRs) in the three mentioned domains, with expert personal investigators (PIs). Within our group (Laboratoire de Spectrométrie de Masse Bio-Organique, LSMBO, university of Strasbourg), we are particularly focused on developing specific analytical techniques for the characterization of the developed molecules inside the consortium (**Figure 17**).

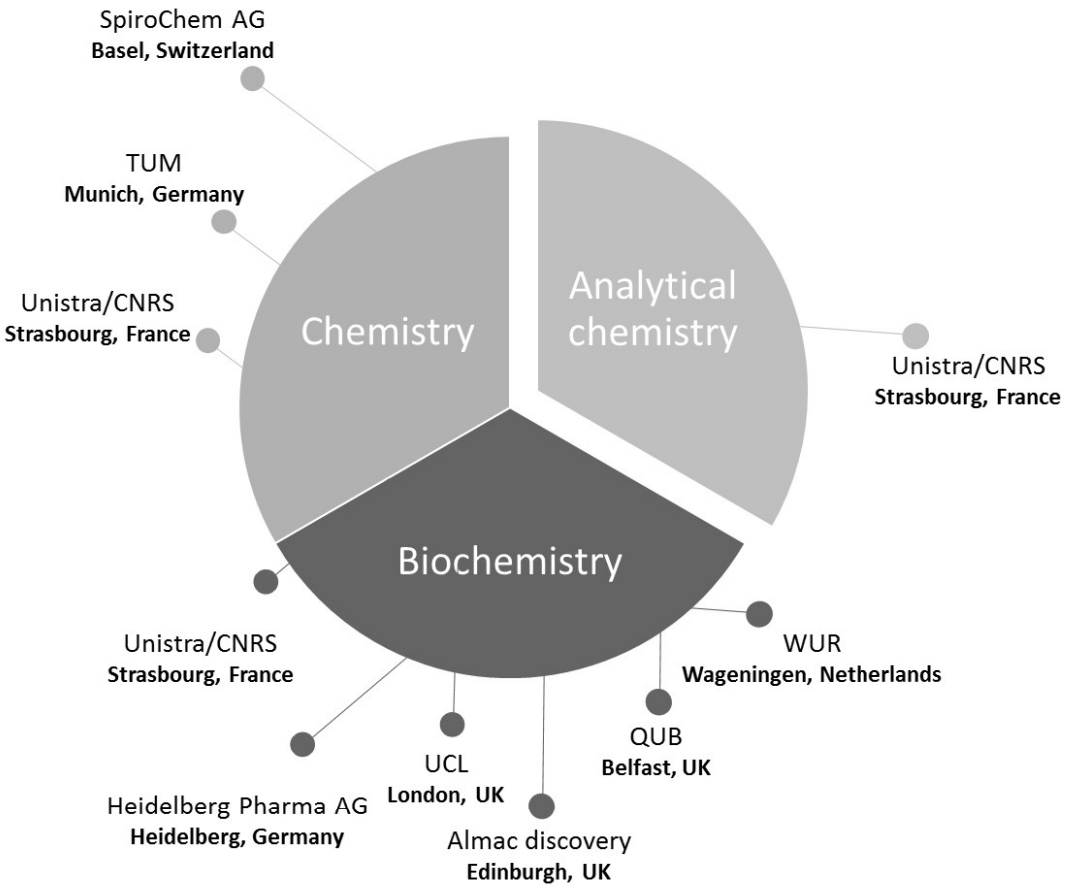


Figure 17: Graph depicting the different fields of the TACT program aiming at developing next-generation antibodies, among which our LSMBO group (Unistra/CNRS) is focused on the development of the analytical strategies dedicated for the characterization of the newly developed anti-cancer therapies.

Chapter 2: Analytical strategies for mAb-based products characterization

1. Monitoring mAbs CQAs is essential

The increase development of multiple mAb-based products requires a constant monitoring of these moieties, due their different complicated molecular structures and various manufacturing processes. Indeed, for a comprehensive evaluation of this plenty of therapeutics and in order to support their market authorization, a combination of physico-chemical properties and isoform patterns that have been defined by the International Conference on Harmonization (ICH) Q6B specifications [111], is usually required. Those specifications are defined as a list of critical quality attributes (CQAs) that are specific properties to ensure the safety, efficacy, purity and the overall quality of the final product (**Table 2**) [112], and therefore need to be assessed during the development process. Identifying CQAs at an early stage of biotherapeutics development, is a valuable step, as it can allow to improve the product quality and establish robust control strategy [113]. For mAb conjugates in particular, it is necessary to quantify the amount of unconjugated mAb (D0), as well as to determine the drug-load distribution (DLD) and the average drug-to-antibody ratio (avDAR). Moreover, the identification and quantification of size variants (HMWS and LMWS) and charge variants (acidic and basic species) is at utmost interest for ADCs/PDCs comprehensive characterization.

Table 2: List of different critical quality attributes related to molecular variants of mAbs. Adapted from Alt *et al.* (2016).[113]

Category	Quality attribute
Size-related Variants	HMWS and LMWS
Charge-related Variants (Acidic)	Deamidation in CDR and in Non-CDR
	Glycation in CDR and in Non-CDR
Charge-related Variants (Basic)	Aspartic Acid Isomerization in CDR and in Non-CDR
	Aspartic Acid Isomerization in Non-CDR
	N-Terminal Leader Sequence (may be molecule specific) or Pyroglutamic Acid
	C-Terminal Lysine, Proline (IgG1) or Leu (IgG4) Amidation
Oxidation-related Variants	Oxidation in CDR (Met, Trp) and in Non-CDR (Met, homo-variant, or hetero-variant)
Fc Glycosylation	Afucosylation, Galactosylation, High-Mannose, Sialylation (NANA, NGNA) and Non-Glycosylated Heavy Chain
Structural Variants	Cysteine Forms, Sequence Variants and Protein Structure

In addition, the development of mAb-related products is particularly complex, due to the involvement of several post-translational modifications (PTMs) during the manufacturing and storage phases. As mAbs are subject to PTMs and degradation during cell culture, purification, storage and even after administration, PTMs have direct consequences on their potency, stability and immunogenicity [8]. MAbs can undergo a dozen of PTMs (**Table 3**), among which glycosylation (N, varying with glycoforms)

is one of the well-known. The impact of glycosylation in mAbs have already been discussed in the previous chapter (**Part I, chapter 1**) [38, 44], thus the characterization of mAbs glycosylation profile represents an important CQA. Disulfide bond breakage (-SH SH-, +2 Da) or linkage (S-S, -2 Da) represent another major PTM that have been well studied in mAbs [52], due to effect they have on the structural integrity of mAbs. Moreover, deamidation (+1 Da), oxidation (+16 Da), acetylation (+42 Da), phosphorylation (+80 Da) as well as amino acids clipping (*e.g.* K-clipping, -128 Da) and cyclization (PyroE/PyroQ, -18/17 Da) are also PTMs that are widely encountered in proteins developments and account for additional heterogeneity, hence must to be assessed carefully [8, 114]. In term of PTMs assessment, mass spectrometry is also a valuable insight and is often coupled to liquid chromatography and electrophoretic methods for a more comprehensive characterization of these variations.

Table 3: Post Translational Modifications of mAbs and their potential impact on stability, function, immunogenicity and pharmacokinetics/dynamics. Adapted from Ambrogelly *et al.* (2018). [115]

PTM	Potential impact
N-terminal • PyroGlu • Leader sequence Leader sequence	Low risk to comparability because of lack of impact on efficacy thus do not affect safety.
C-terminal • Partial removal of C-terminal lysine • Amidation • Truncation	Low risk to comparability because of their low percentage and lack of impact on efficacy thus do not affect safety.
Fc-glycosylation • Sialic acid • α-1,3 Gal • Terminal Gal • Absence of core-fucosylation • High mannose	• N-Glycolylneuraminic acid (NGNA) is immunogenic • α-1,3 Gal on Fab oligosaccharides is immunogenic • Presence of galactose enhances CDC • Absence of core-fucose enhances ADCC • mAbs with high mannose show enhanced ADCC and shorter half-life
Asn deamidation	Decrease potency in CDR
Asp isomerization	Decrease potency in CDR
Succinimide	Decrease potency in CDR
Met and Trp Oxidation	Decrease potency in CDR and could decrease FcRn binding affinity and result a in shorter half-life
Cysteine-related variants • Disulfide isoforms • Free cysteine Trisulfide bond Thioether, D-cysteine, cysteinylation	IgG2 disulfide bond isoforms may affect potency. Higher amounts of free cysteines decrease mAb thermal stability and trigger formation of covalent aggregates. Other modifications are considered low risk because of their low levels or natural presence in humans
Glycation	Decrease potency in CDR and increases mAb aggregation
Fragments	Low risk because of their low levels
Aggregates	Causes immunogenicity and loss of efficacy

In summary, the complexities and variabilities encountered during the development of biotherapeutics need to be assessed carefully through the identification and quantification of CQAs and PTMs, and mass spectrometry plays an important role in this context.

2. MS toolbox for MAb-based therapeutics characterization

The assessment of CQAs and PTMS could be performed through the characterization of mAb-based products at different levels [10](**Figure 18**): i) at the bottom level, mAb products are enzymatically digested to release peptides (700-7,000 Da), which are further analyzed by LC-MS/MS approaches (peptide mapping). The analysis of those peptides gives a detailed information on the type and location of PTMs through the amino acid primary structure characterization [116-117]. ii) At the middle level, the proteolysis is limited and controlled, which provides relatively small mAb subunits (25-100 kDa) that are further characterized with LC-MS/MS approaches [118]. Moreover, a middle-down approach is used at this stage, to investigate in-depth the PTMs and drug conjugation positions [28-29, 119]. iii) The top level provides an information on the intact mass of mAbs (150 kDa) while preserving the integrity of their structure. This can be achieved either by a classical denaturing MS approach, where only covalent linkages are maintained [120-121], or in native conditions (nMS) where even non-covalent bonds are maintained [122-123]. The latter approach is particularly interesting for covalently linked and non-covalently linked mAb conjugates, where the information on D0, DLD and DAR is obtained [121], and could be hyphenated to LC or ion mobility (IM) separation to obtain information on charge/size or conformation variants [16, 124-127]. Furthermore, at the top level, the intact mass and the primary structure information could both be correlated through a top-down MS approach either in native or in denaturing conditions [128-129]. These two approaches (Native and top down MS) are both complementary when it comes to characterizing mAb-based formats, which is going to be discussed in details in this manuscript. iv) Finally, obtaining higher order structure (HOS, >150 kDa) information upon the characterization of epitope mapping, structural conformation or aggregate analysis is also possible using the previously mentioned techniques (IM-MS, nMS, TD-MS). Additionally, labelling approaches such as; Hydrogen/Deuterium Exchange MS (HDX-MS) [102, 130], cross-linking MS (XL-MS) [102, 131], as well as mass photometry that have proved their utility in the characterization of HOS of mAbs. [132-134].

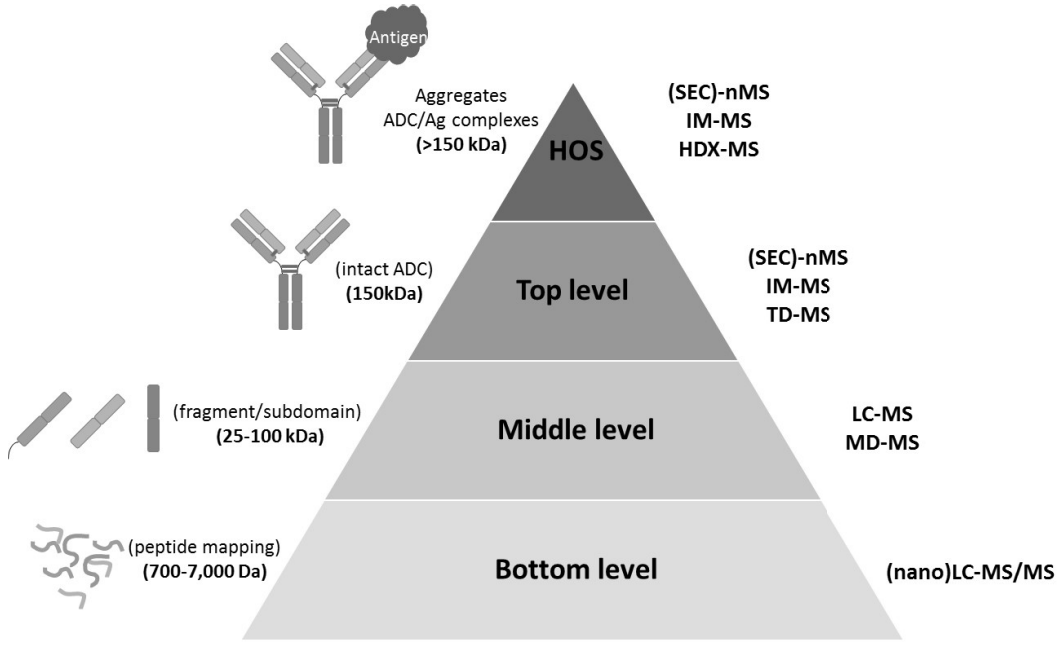


Figure 18: The different characterization level of mAb-based products using various techniques. At the bottom level peptide mapping is performed through LC-MS/MS analysis. At the middle level mAb subunits are analyzed by a simple LC-MS or following a MD-MS approach. At the top level, intact mAb is analyzed using native MS approaches, IM-MS or TD-MS. The characterization of HOS is performed using nMS, IM techniques as well as peptide centric approaches such as HDX-MS.

3. Native mass spectrometry (native MS, nMS)

Native MS (nMS) is based on electrospray ionization (ESI), where proteins solubilized in a non-denaturing buffer are transferred into the gas phase of the mass spectrometer, which maintains their non-covalent interactions [14]. In the context of mAb characterization, nMS provides valuable information through accurate mass measurements of intact mAbs and through their CQAs assessment.

3.1 Key milestones in native MS (nMS)

Firstly, until the 1980s, the field of mass spectrometry focused on the measurement of small *m/z* organic molecules. With the introduction of soft ionization techniques such as ESI [135] and Matrix-Assisted Laser Desorption Ionization (MALDI) [136], a more comprehensive characterization of larger molecules could be achieved, including macromolecule complexes. While MALDI generates singly charged ions, giving rise to high *m/z* signals, in addition to encountering problems in preserving large macromolecule complexes, its use was limited for native MS analyzes. Conversely, ESI produces multiply charged protein ions, which are more commonly detected in the available *m/z* ranges. Hence, since its introduction in 1989, it was the go-to ionization method for biomolecules analysis [135].

After the introduction of ESI, Chait and Henion groups focused both in 1991 on the analysis of noncovalent protein/ligand complexes using nMS [137-138]. In the following year, Loo's group reported on the characterization of noncovalent protein/protein interactions by nMS [139]. In parallel, nMS discoveries were supported by the development of new analyzers that widened the scope of

applications. Quadrupole analyzers were first modified to achieve a maximum acquisition range of >4,000 *m/z*, by lowering radiofrequencies, yet suffered from low resolving power [140-141]. Few years later, time-of-flight (ToF) analyzers were favored [142], which achieved higher sensitivity and higher resolution in comparison to quadrupole analyzers and thus was used to observed noncovalent interactions in 1996 [143]. In the same year, the ESI-Q-TOF was presented and became the most successful platform for nMS [144]. Few years later, Q-TOF instruments were favored for nMS applications after enhancing of their transmission and desolvation of high *m/z* ions [145-146]. More recently, after the introduction of Orbitrap analyzer in 2006, the scope of applications of nMS was widened to the characterization of large macromolecular complexes [147-148]. Furthermore, Orbitrap-based charge detection MS (CDMS) was established, to monitor charge and *m/z* of single ions and was widely used for determining the mass of intact proteins [149].

In parallel, the term of native MS (nMS) was conceived in 2004 [14, 150], although native MS can be tracked back to the 90s, to designate the analysis of intact proteins and their protein complexes maintained by non-covalent interactions, in a native-like folded state. Shortly after, nMS was coupled to size exclusion chromatography (SEC), in order to automate buffer exchange on the one hand, and to identify and quantify the size variants on the other hand. This coupling, along with the emerging of robust and resolutive instruments, paved the way to new applications such as mAb aggregates and fragments characterization [15-16, 124-125, 151-152]. The SEC-MS coupling was even more improved thanks to the discovery of sub-3µm SEC columns in 2017 [153], which allowed achieving better separation efficiencies, thus better HMWS and LMWS characterization, and quantification. Indeed, SEC-MS is the most commonly used for the characterization of mAb products. However the native MS field was also highlighted by several native LC-MS couplings, such as the first HIC-MS coupling for mAbs glyco-profiling [154] or the CEX-MS coupling for mAb charge variants analysis [19, 155-156]. Until now, SEC-MS workflow in particular, is enhanced when using bio-inert SEC columns [157] that reduce nonspecific interactions, and therefore improve the overall signal of intact mAb structures. **Figure 19** summarizes all the key milestones in nMS field. Finally, nMS offered further insights into direct characterization of proteins from the crude growth media [158], which makes it a promising approach for the characterization of newly designed biotherapeutics.

In conclusion, nMS convincingly demonstrated its suitability for mAb products characterization. The various instrumental and analytical advancements paved the way to establish nMS workflow routinely implemented methods in biopharma labs.

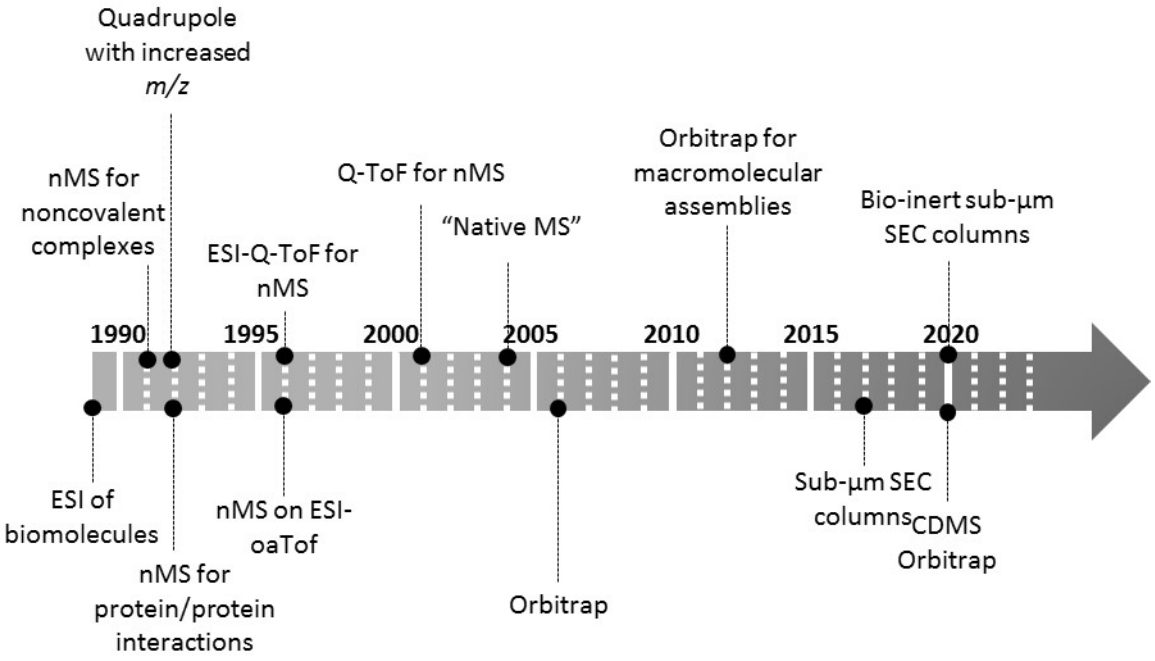


Figure 19: key milestones in nMS. Adapted from Leney *et al.* (2016) [14] and from Füssl *et al.* (2021) [123].

3.2 The role of nMS in mAb-based products characterization

The overall information that could be obtained using nMS are about the homogeneity/heterogeneity, binding affinity and stoichiometry of biomolecules. The intact mass measurement based on competition experiments are performed to obtain the information related to the specificity or affinity of partners in a complex (*e.g.* mAb-antigen complex). Moreover, the comparison of mass spectra from native and denaturing conditions, we could directly link the masses to the stoichiometry of interaction between mAb/antigen, for instance. Therefore, these intact mass measurements could easily allow identifying glycoforms and assess HOS (dimer, trimer, and tetramer). In addition, nMS is particularly well suited for ADCs for additional reasons discussed below.

a) Simplification of MS spectra of lysine-linked ADCs:

On the one hand, nMS plays a valuable role in simplifying MS spectra of heterogeneous lysine-linked ADCs [120]. Indeed, the example of T-DM1 in **Figure 20** shows that in denaturing conditions, the MS signal based on lower *m/z* range (2,000-5,000 *m/z*) exhibits a large number of high charge states (centered on the 45+), thus leading to a heterogeneous MS spectrum with challenging identification of overlapping DAR species (D0-D8). Although, deglycosylation of the ADC sample could somehow de-complexify the spectra, the overlapping species can still lead to miss-deconvolution and miss-assignment of the different DAR species. Conversely, when analyzing the same ADC in native conditions, the multi-charged profile shifts towards higher *m/z* range (4,000-7,000 *m/z*) due to charge reduction effect, leading to a better quality spectrum, which facilitates the DAR species identification and quantification.

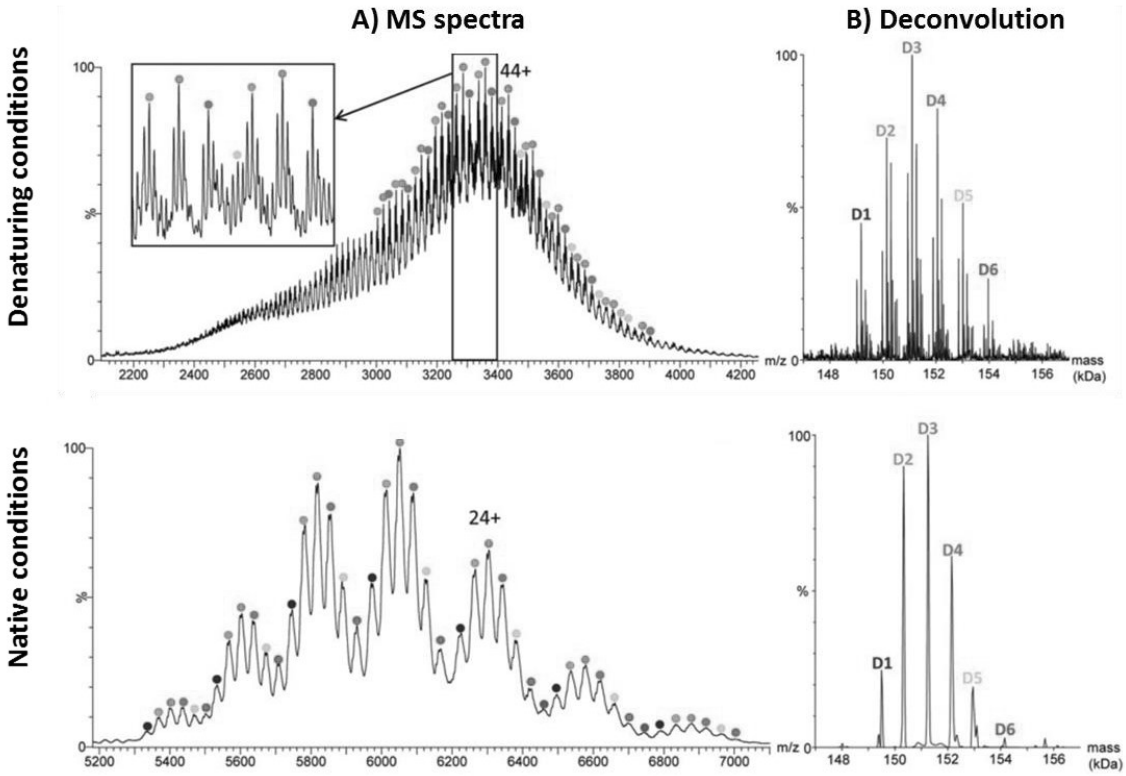


Figure 20: Trastuzumab emtansine (T-DM1) ADC characterized by mass spectrometry under denaturing conditions (**upper panel**) and under native conditions (**bottom panel**) on a Q-ToF instrument. The figure shows (**A**) a crowded mass spectrum in denaturing conditions with high charge states, *versus* a more clear mass spectrum in native conditions with reduced charge states. (**B**) Deconvolution spectra where similar DAR species (D1-D6) are observed in denaturing and native conditions. Adapted from Marcoux *et al.* (2015) [120]

b) Preservation of non-covalent interactions of cysteine-linked ADCs:

On the other hand, as the main advantage of nMS is preservation of the native-like structure of proteins and protein complexes in the gas phase, this can be beneficial for cysteine-linked ADCs [159]. The example of BV shown in **Figure 21**, depicts clearly the phenomenon of protein unfolding in denaturing conditions, which leads to the identification of a mixture of overlapping covalent and non-covalent DAR species [122]. The signal observed at lower *m/z* range (1,000-4,000 *m/z*), does not allow to conclude on the avDAR or DLD as mAb-payload fragments are observed in addition to unconjugated mAb (D0). Conversely, using nMS, the integrity of the ADC is maintained through non-covalent interactions. Thus, the obtained spectrum in higher *m/z* range (4,000-7,000 *m/z*) leading to one homogeneous distribution, allow unambiguously to determine the different 0-8 DAR species as well as to determine the avDAR, which is calculated according to equation (1).

$$1.\ avDAR = \frac{\sum_{k=0}^{8} k\ \times\ intensity\ DARk}{\sum_{k=0}^{8} intensity\ DARk}$$

In summary, nMS have demonstrated to be well adapted for ADCs, in addition to other mAb products characterization, which allowed its integration in the biopharmaceutical field.

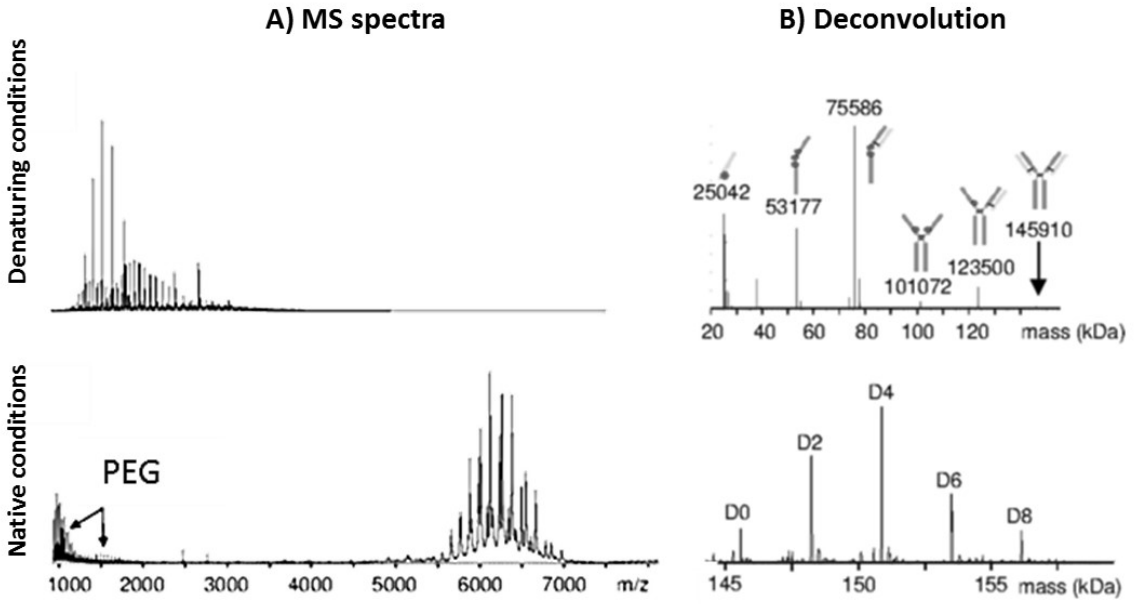


Figure 21: Brentuximab vedotin (BV) ADC characterized by mass spectrometry under denaturing conditions (**upper panel**) and under native conditions (**bottom panel**) on a Q-ToF instrument. The figure shows (**A**) a mixture of overlapping species in denaturing conditions, *versus* one homogeneous species corresponding to the intact ADC in native conditions. (**B**) Deconvolution spectra where DAR species (D0-D8) are only observed in native conditions while in denaturing conditions masses corresponding to fragment subunits are observed. Adapted from Debaene *et al.* (2014). [122]

3.3 Hyphenation of non-denaturing LC to native MS

The natural complexity of therapeutic proteins, ADCs in particular, usually requires additional separation methods for simplification means prior to MS analysis. Capillary electrophoresis (CE) and liquid chromatography (LC) techniques such as size exclusion chromatography (SEC), ion exchange chromatography (IEX) or hydrophobic interaction chromatography (HIC) are often hyphenated to native MS for intact and middle level analyses of mAb products [10, 121, 125, 160]. Additionally, several studies have been published using ion mobility spectrometry (IMS) for an additional level of separation when characterizing ADCs [127, 152].

a) Capillary zone electrophoresis (CZE)

CZE is based on the separation of proteins based on their electrophoretic mobility, which is proportional to their charge-to-size ratio [161]. Furthermore, CZE can also provide information upon protein shape and conformation changes. CZE separation is carried out in aqueous solutions present in a fused-silica capillary (50-75 µm inner diameter; 20-100 cm length) where a high voltage is applied (up to 30 kV). CZE-MS can be successfully achieved with MS-compatible background electrolytes (BGEs), such as ammonium acetate at near physiological pH. Furthermore, the native coupling could be enhanced by using neutral (uncharged) agent, such as polyacrylamide or cellulose [162]. Due to its minimal sample required (<1µg), CZE-MS is an attractive method for characterization of biotherapeutics such as ADCs [163].

b) Size exclusion chromatography (SEC)

The separation in SEC is strictly based on the hydrodynamic volume of the proteins, which is strongly related to their molecular size. The separation mechanism is governed by size-dependent protein exclusion from or diffusion into the pores of the stationary phase (entropy-driven), instead of interactions with the stationary phase (enthalpy-driven) [164]. In this manner, large proteins are rapidly excluded from the pores and thus elute first. Conversely, smaller proteins diffuse into the spherical porous particles and elute later. Thus, the pore size of silica packed columns is normally controlled, as it is a crucial parameter for the elution order. Currently, SEC available columns have an average pore size in the 125-900 Å range, among which those with pore diameters of 200-300 Å are used for separation of 10-500 kDa proteins [153, 165]. The separation is carried out in isocratic mode with mobile phase acting as a carrier solvent. Traditional solvents are composed of non-volatile salts (*e.g.* phosphate buffers at 5-50 mM), hence are not compatible with MS detection. When coupled to nMS, volatile salts in the 100-500 mM range with pH ~7.0 are usually used. Those buffers allow maintaining the ionic strength of the protein structure, and therefore maintains even its weak non-covalent interactions [15, 165].

The SEC-nMS workflow depicted in **Figure 22** is particularly useful for mAb products characterization, mainly due to the automated buffer exchange using short SEC columns (<50 mm) with relatively small pore diameters (<200 Å), which save an enormous amount of time in comparison to manual desalting (*via* gel filtration devices or ultrafiltration columns). For separation of mAb size variants, longer columns are usually required (150-300 mm) packed with sub-µm particles, to achieve fast and efficient separation of LMWS and HMWS [157, 166-167].

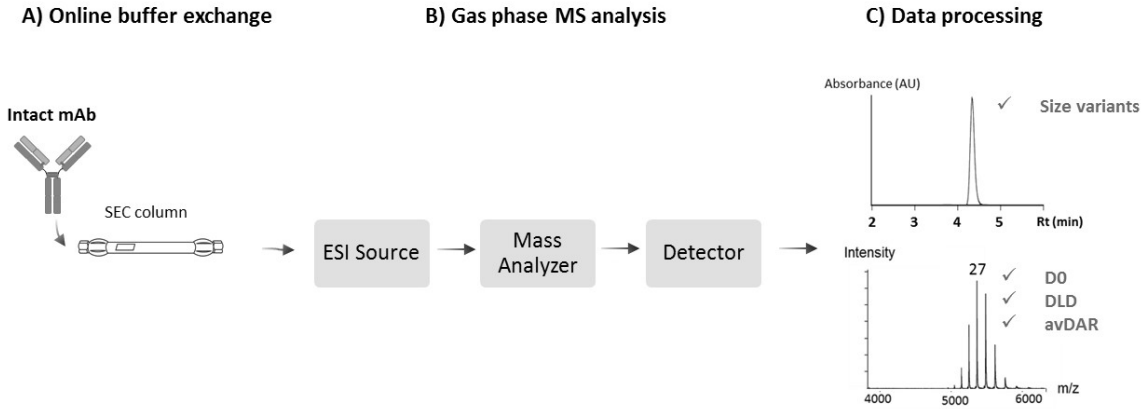


Figure 22: SEC-nMS workflow for mAb analysis resumed in three steps: (A) online-buffer exchange using SEC columns, (B) gas phase MS analysis and (C) data processing that provides information on size variants, D0 quantity, DLD and the avDAR.

c) Ion exchange chromatography (IEX)

IEX is the go-to method for the characterization of charge variants of mAbs (acidic and basic variants as a result of PTMs) [18, 155, 168-170]. The separation is based on electrostatic interactions between the net positively or negatively charged protein and protein complexes, and immobilized ionic groups such as

strong or weak acidic/basic moieties, on the stationary surface. Cation exchangers are negatively charged (Cation exchange chromatography, CEX) and are subdivided in two groups; i) weak cation exchangers that are acid linked to the resin (carboxymethyl group) and are stable for pH>6.0 and ii) strong cation exchangers that are acid attached to the resin (sulfopropyl group) remaining negatively charged for a wide range of pH. Conversely, anion exchangers are positively charged (Anion exchange chromatography, AEX) and are subdivided in two groups as well; i) weak anion exchangers like diethylaminoethane linked to the resin and ii) strong anion exchangers like quaternary amine functional group attached to the resin [18, 171]. The straightforward coupling of IEX and MS can be achieved with compatible mobile phases comprising volatile salts. The elution can be achieved by applying a salt gradient *i.e.* increasing the salt concentration at a constant pH [168, 172]. A separation based on pH gradient with low-ionic-strength mobile phases (20-50 mM) has also gained interest in IEX-MS applications [173]. As proteins elution is promoted when the pH of the mobile phase is close to the pI of proteins, CEX-MS has been particularly useful for the characterization of mAbs as they are basic proteins [155-156]. Moreover, in some cases, salt-mediated pH gradients (50-200 mM ammonium acetate with pH gradient) can enhance the separation performance, which was widely demonstrated by several groups [155, 174-175].

d) Hydrophobic interaction chromatography (HIC)

HIC is employed for native separation of proteins based on their difference in hydrophobicity, and therefore can give an information on conjugated mAbs misfolding or mAbs PTMs that directly induces hydrophobicity changes [176-177]. The stationary phase in HIC is silica- or polymer-based with moderately hydrophobic ligands (*e.g.* short n-alkyl or phenyl groups attached to its surface). Therefore, proteins elution relies on hydrophobic interactions, which are modulated by concentration variations of a salt-induced precipitation electrolyte (*e.g.* ammonium sulfate) in the mobile phase. HIC-MS has emerged as a useful coupling for analysis of ADCs characterization, in particular for cysteine-linked ADCs [10, 123, 177]. Furthermore, development of a 2D-LC method comprising HIC (using conventional non-volatile salts) and SEC for desalting protein fractions prior to nMS detection was particularly interesting for profiling complex protein samples [178].

e) Ion Mobility Spectrometry (IMS)

IMS is based on the measurement of ions' drift time under the influence of an electric field, through a region filled with a background gas [179]. The drift time is correlated to the mass, the charge and the shape/conformation of ions, where the most compact ions travel faster and thus are detected first in the MS [180]. Conversely, more charged and more extended ion conformations collide more frequently with gas molecules and are detected later resulting in higher drift times. The collision cross sections

(CCS) is obtained based on the drift time, which offers global information about the conformation of proteins in the gas phase [181]. Native IMS-MS (nIMS-MS) is therefore a valuable tool for the assessment of mAbs PTMs such as disulfide bonds, glycosylation, deamidation and more [120, 122, 127, 152, 165, 178, 182-183]. Moreover, there are many IMS methods with different instrument platforms (**Figure 23**) where each has its advantages and disadvantages, depending on the specific application. Drift tube IMS (DTIMS) is the classic IMS model that provides CCS measurements directly and is widely adopted as method in IM-MS research [181]. In travelling-wave IMS (TWIMS), CCS values cannot be directly calculated and thus require calibrations. TWIMS utilization has been increased compared to DTIMS ion mobility [184-185]. Trapped IMS (TIMS), operates reversely to DTIMS using a gas flow. [128, 186-187]. Field asymmetric IMS (FAIMS) [188] is an atmospheric pressure IMS technique and unlike other IMS techniques, it cannot provide CCS measurements [189-190]. Differential mobility analyzers (DMA) operate similarly to DTIMS and can detect very large analytes (such as mAbs and viruses) and therefore provide CCS calculations that DTIMS cannot afford [191]. One of the alternative approaches that overcome limitations of differentiating very close mAbs conformations, is collision-induced unfolding (CIU)[182]. CIU are performed by increasing the accelerating potential difference which induces ion activation before IMS separation [192]. The ion activation leads to conformational transitions through different conformational states/intermediates, related to unfolding/compaction of the proteins. These conformations are assessed by arrival time distribution (ATD) changes which are recorded at each collision voltage (CV) [193]. The subtle changes of conformational state in the gas phase, are therefore monitored through CIU fingerprints (unfolding plots) [126].

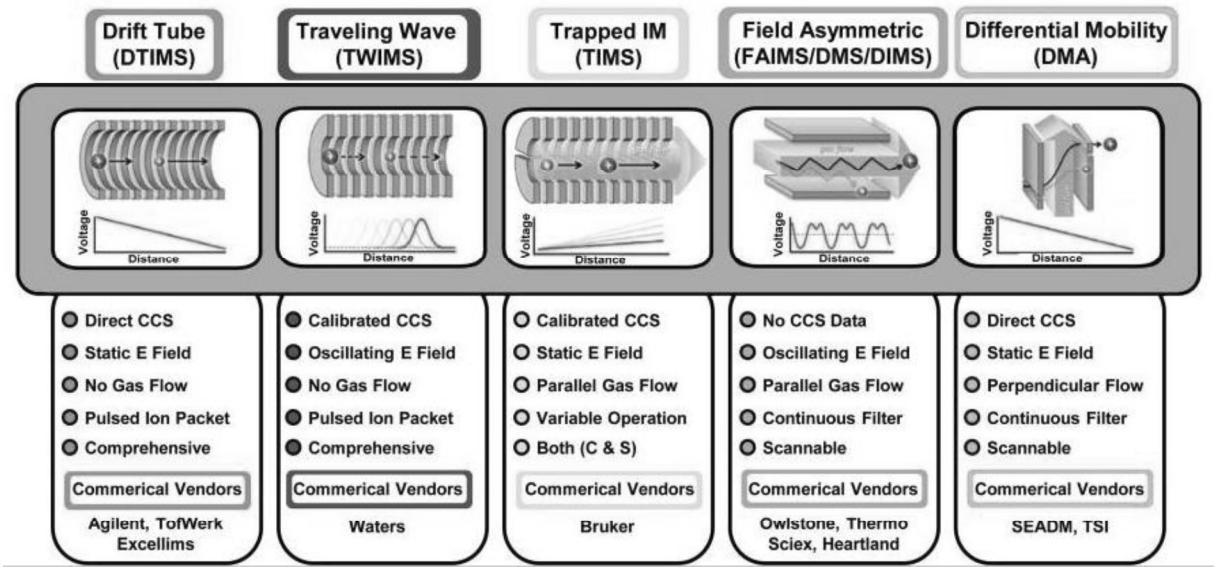


Figure 23: Different types of ion mobility devices with main differences, name of vendor and main advantages. Adapted from Dodds *et al.* (2019). [180].

3.4 Instrumental and data processing considerations for nMS

The technological advances in the context of instrumentation and processing softwares supported the emerging of different analytical strategies for mAb-based products characterization, at their intact or at their amino acid level. In the case of nMS, the ionization of large assemblies required analyzers with extended *m/z* ranges. Recent ToF analyzers have met this criterion as they can reach more than 10,000 *m/z* (*e.g.* Synapt G2 HDMS, Waters). In addition, the development of orbitrap analyzers have particularly impacted the field of nMS, as the *m/z* was extended up to 80,000 *m/z* (*e.g.* Q Exactive UHMR, Thermo Fisher) with considerably higher mass resolving power, for more precise mass analysis. With the increasing interest for nMS biotherapeutics characterization, biopharmaceutical companies aimed at coupling LC techniques to nMS for CQAs assessment [123]. Therefore, more versatile and automated platforms have been developed (*e.g.* The BioAccord LC-MS, Waters) to answer their needs.

Native MS data normally apply simple algorithm for intact mass measurement, based on the *m/z* ratio of ions distribution in the MS spectra [194]. The molecular weight determination is performed either manually using the formula (2) or through softwares using different algorithms (*e.g.* Masslynx, UNIFI, Biopharma Finder or Protein Metrics).

$$2.\ z(m1) = \frac{x\,(m2 - mp)}{m2 - m1}$$

Where z is the calculated charge for m1, m2 is the ion with x less charge (in this case x=1) and mp is the mass of the proton (1.00728 Da).

Once the z value is determined, the MW can be easily calculated following this equation (3):

$$3.\ MW = z\,(m1 - mp)$$

4. Mass photometry

The biotherapeutics characterization field have also took advantage of the recently developed technique; mass photometry (MP). MP affords direct mass measurement of intact noncovalent assemblies, based on the interaction between the scattered light of biomolecules and the reflected light by the crystal surface upon irradiation with a visible laser (523 nm). The light scattered by a particle correlates linearly with its volume and its refractive index. Thus, as the properties of proteins have negligible variations, their scattering light is directly proportional to their mass, which can be converted into molecular weight through a series of calibration using biomolecules with known masses [195-196].

The benefits of MP rely mainly on the direct analysis of proteins in their original buffer, without prior sample preparation [132-133, 195, 197-198]. Additionally, MP uses very small sample amount (100 pM – 100 nM concentration range), compared to nMS that requires > 5 µM for intact proteins analysis.

Moreover, MP can accurately measure molecular masses in the 40 kDa – 5 MDa range, which represents a valuable approach for the assessment of mAb/antigen complexes (~300 kDa), proteasomes (~2 MDa) as well as AAVs (~5 MDa) [199].

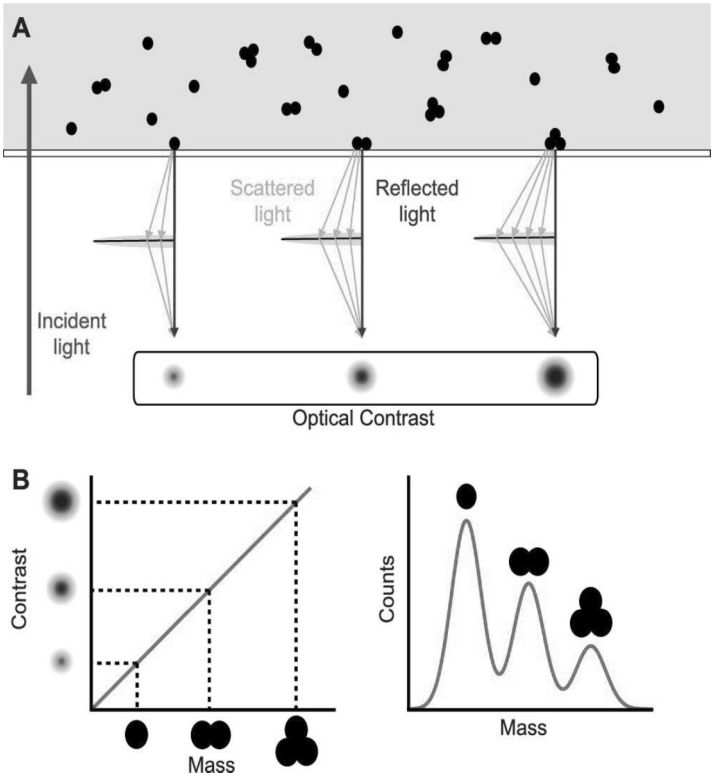


Figure 24: Schematic of MP principle. (**A**) Binding of protein molecules to a surface scatter light. Interference with the reflected ligjt produces optical contrast varying with the size of the detected protein. (**B**) Linear relationship between protein mass and optical contrast which is used to determine the mass distributions. Adapted from Paul *et al.* (2022). [195]

5. Top down MS (TD-MS)

Top down mass spectrometry (TD-MS) consists in the direct ionization and MS/MS fragmentation of intact proteins in the gas phase using different activation techniques [20]. Thus, it is based on protein backbone cleavage and can be performed either in denaturing conditions or in native conditions (native top-down MS, nTDMS), unlike other intact mass approaches that do not induce backbone cleavage (**Figure 25**) [200-201]. The information obtained on the primary structure of those proteins, allow for improved sequence coverage and PTMs detection. Unlike traditional bottom up approach, where proteins are digested into peptides prior to their MS/MS fragmentation, TD-MS does not require any sample preparation which reduces the risk of artefacts generation [20]. In addition, the main advantage of TD in comparison to BU approach, is that it is the only approach where the information provided on the primary structure level can be concomitantly correlated with the intact protein information [202-203]. For these reasons, TD have recently gained a particular interest in characterization of therapeutic

proteins intact mass and their PTMs assessment [25-26, 203-207]. However, TD-MS is still challenging for large proteins such as mAbs, which requires a mild controlled digestion to generate smaller subunits, referred to as middle-down approach (MD-MS) [28-29, 119, 208-209].

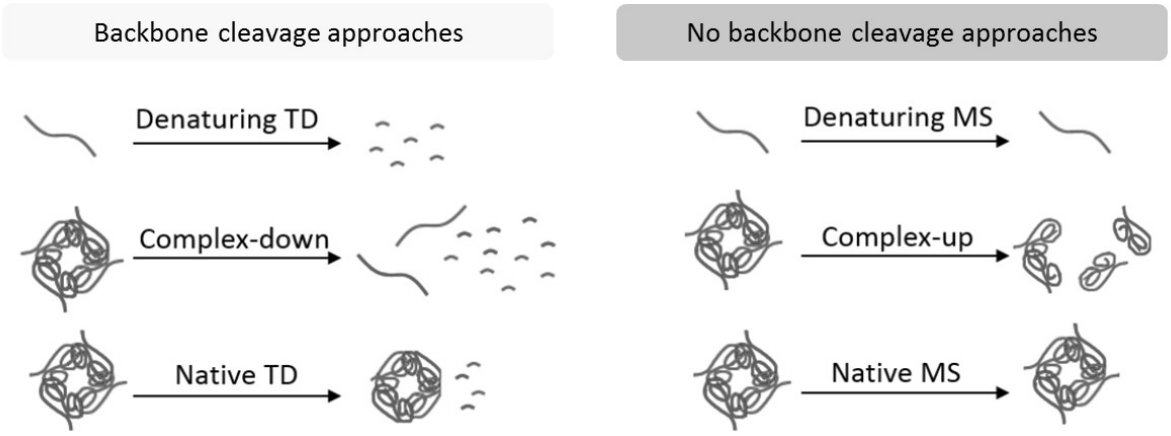


Figure 25: Approaches based on protein backbone fragmentation such as top-down, and other approaches that do not induce backbone cleavage. Adapted from Lermyte *et al.* (2019). [200]

5.1 Key milestones in TD-MS

Since the recent advances of ESI ionization revolutionized the applications of tandem MS/MS in 1989 [135, 210], MS have been essential in proteomics research. With the development of electron-capture dissociation (ECD), Zubarev's and Mclafferty's groups first studied proteins fragmentation in the 90s [211-212]. Even so, Keller's group in 1999 was the first to term to this approach as "top down" to distinguish it from classical bottom up [20]. A couple of years later, the field of TD proteomics was supported by the development of tailored TD-MS softwares starting from ProsightPC, implemented in 2001 by Keller's lab [213-216]. The year after, TD-MS of large proteins (45 kDa) using ECD in FTMS was achieved [217]. With the rising benefits of ETD [217], this fragmentation technique was implemented on a quadrupole linear-trap instrument (QLT) in 2004 for top-down experiments [34]. While the FTICR was the original workhorse in TD-MS [160, 218-219], the high resolving power and sensitivity of the orbitrap made it interesting for TD-MS experiments [30, 206-207, 219-225]. An example of the use of LTQ-orbitrap with the possibility of performing MS^n orbitrap detection , was highlighted for protein sequencing [220]. The low efficiency of classical fragmentation techniques paved the way for UVPD dissociation [226], which was proved to allow near complete characterization of a 29 kDa protein in a hybrid linear orbitrap MS (using the 193 nm laser) in 2013 [227]. Fornelli *et al.* in the same year, provided extensive characterization of larger proteins (30-80 kDa) using ETD on a ToF MS [23]. With development of recent activation techniques such as AI-ETD that is based of ion activation prior to ETD fragmentation, proteins in comparable mass range (30-70 kDa) could be characterized through a TD-MS approach using a quadrupole-Orbitrap-linear ion trap hybrid MS system in 2018 [24]. The field of TD proteomics was particularly advanced through the

fragmentation of large IgGs (~150 kDa) by ETD on a ToF and on orbitrap instruments [205-206], leading to a maximum sequence coverage of 33%. Several groups allowed increasing the sequence coverage of mAbs and mAb conjugates through MD-MS approaches [25, 28-30, 119, 208, 219, 228]. ADCs were particularly challenging through TD-MS characterization, but thanks to the complementarity of several techniques, these challenges could be tackled. Brodbelt's group in 2020 perfectly emphasized this point by combining EThcD, ETD and UVPD for an improved ADC fragmentation [29]. The plethora of fragments generated from several activation techniques, gave rise to a mixture of terminal and internal fragments, where the latter type was less studied and often miss-assigned. Recently, in 2021 the group of Loo focused on this particular point and developed a new algorithm for internal fragments assignment, in addition to terminal ones, called ClipsMS [229]. The added value of internal fragments assignment was further emphasized by the same group through ADC near complete fragmentation, in 2023 [129](**Figure 26**).

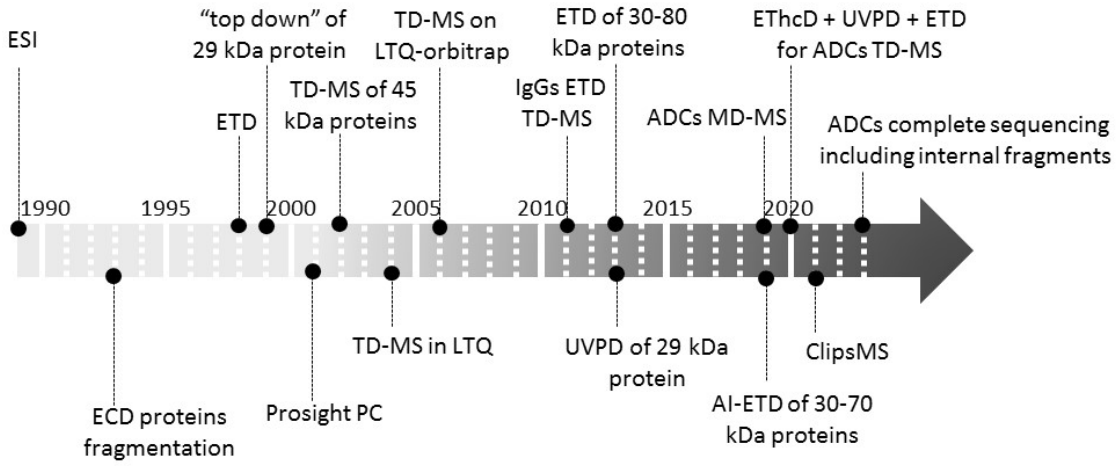


Figure 26: key milestones in TD-MS field.

5.2 Fragmentation techniques

a) Collision-based dissociations

The main collision-based dissociations are CID and HCD. In CID, precursor ions kinetic energy is increased through their acceleration, which allows their collision with neutral gas molecules in the collision cell (argon, nitrogen or helium) [230]. When the ion's kinetic energy is increased, it can be converted into internal vibrational energy that leads to peptide bond cleavage upon collision with neutral molecules. As CID typically takes many small steps of energy conversion to result in dissociation, it is referred to as a "slow heating activation method". The fragment ions generated upon peptide bond CID cleavage are mainly *b* and *y* ions [34, 231-232](**Figure 27**). CID preferably breaks weak peptide bonds, thus labile PTMs are fragmented at the MS/MS stage. In contrast, HCD circumvent this challenge as it allows ions to accumulate higher internal energy, resulting in additional fragmentation

pathways. Unlike CID, HCD is performed in a multipole collision cell, which is solely available on Orbitrap instruments [33].

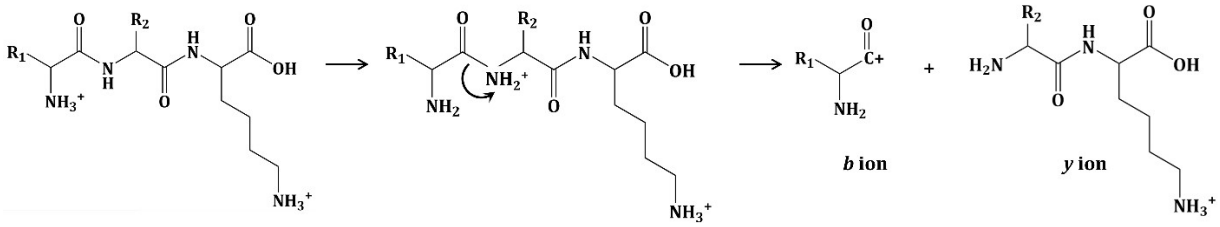


Figure 27: Mechanism of cleavage upon collision-based dissociations. Adapted from Syka *et al.* [34]

b) Electron-based dissociations

Electron-based dissociation techniques (ExD) are good alternatives to collision-based dissociations, and allow high sequence coverage for intact proteins characterization. In particular, electron capture dissociation (ECD), introduced in 1998 has provided unique MS/MS cleavage of biomolecules [233]. ECD is based on the capture of multiply charged ions to single electrons with low energy (< 0.2 V), that are directly introduced to trapped gas phase ions. The major products of electron capture of $[M + nH]^{n+}$ cations are the charge-reduced species $[M + nH]^{(n-1)+}$ that are followed by protein backbone cleavage by radical-driven reaction. Typically, the generation mechanism of *c*, z, c• and *z*• fragment types, is proposed by the Cornell and the Utah-Washington mechanisms [233-234]. After ECD, electron transfer dissociation (ETD) was described by Syka *et al.* (2004) [34], which is based on electron transfer from a radical anion (generally fluoranthene or anthracene) to multiply charged precursor ions in the ion trap. Similarly to ECD, the electron-driven pathways lead to the generation of *c*, z, c• and *z*• fragment ions (**Figure 28**) [235-236]. In fact, the fragmentation efficiency of both ECD and ETD relies highly on the charge state and thus charge density of the precursor ion [237]. Therefore, when reacting the radical anion with low charge density ions, a phenomenon of ECnoD or ETnoD [238] is usually observed due to absence of fragmentation with only charge reduced species [239]. Activated ion ETD and ECD (AI-ETD and AI-ECD, respectively) were recently implemented for top-down fragmentation of intact proteins, thanks to their higher efficiency compared to ExD techniques [24, 26, 240-242]. Indeed, the first infrared photo-activation step consists in making the protein in vibration and breaking noncovalent interactions, while the second step allows efficient electron-based fragmentation.

ETD and ECD capabilities are often combined with HCD and CID dissociations, to yield complementary sequence coverage. In fact, ions collisional dissociation are often preferred to bring an additional energy prior to ExD and thus increasing the fragmentation efficiency. Those hybrid techniques (*i.e.* EThcD and ETciD)[221, 243] have been demonstrated promising in several top-down studies [30].

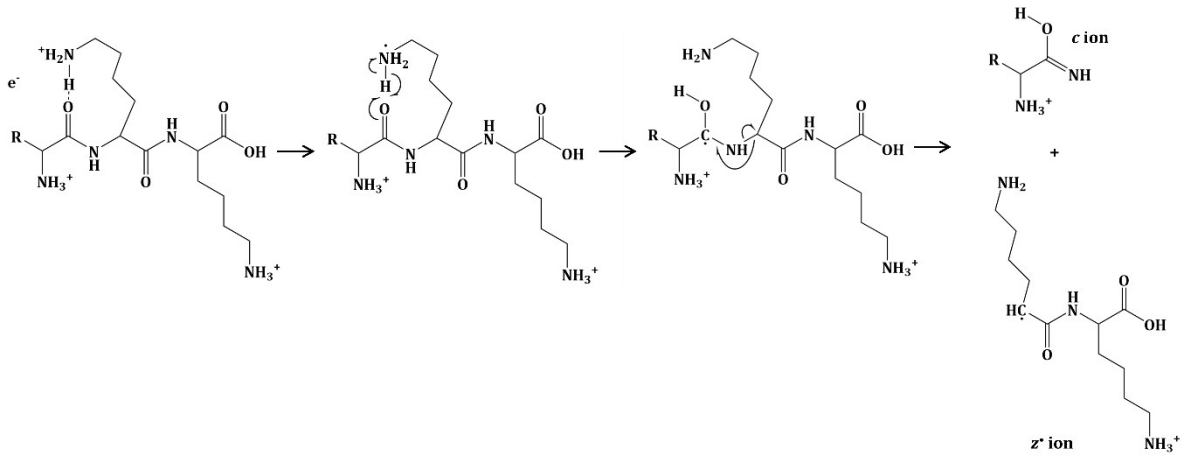


Figure 28: Mechanism of dissociation upon electron-driven fragmentation techniques. Adapted from Syka *et al.* [34]

c) Photo-dissociation techniques

Photo-dissociation techniques rely on the use of laser beams to produce photons, which are absorbed by proteins and thus inducing their fragmentation [244]. The photons could be either in the infrared domain; infrared multiple photo-dissociation (IRMPD) or ultraviolet domain (ultraviolet photo-dissociation (UVPD) [35, 226, 244].

In IRMPD, a CO_2 laser at 10.6 µm wavelength is used to generate photons with a relatively low energy (0.1 eV). Those photons are subsequently accumulated and thus lead to the protein backbone cleavage, with mainly *b* and *y* type ions [245-246].

In contrast, the photons generated by UVPD lasers (at different wavelengths including 266 nm, 213 nm, 193 nm, and 157 nm), have higher energy (6.4–7.9 eV) which offers a more extensive backbone fragmentation [247]. UVPD involves different dissociation pathways through two main mechanisms: i) direct dissociation, which relies on electrons excitation that leads to their direct fragmentation and ii) internal conversion, where the internal energy of ions is converted to vibrational energy which induces their fragmentation. These two mechanisms work together to yield high sequence coverage upon generation of all type of ions (*a*, *b*, *c*, *x*, *y*, and *z*) (**Figure 29**) [247]. UVPD thus gained a lot of interest in the top-down proteomics field [28, 35, 227, 248-250], with a particular importance of 193 nm and 213 nm lasers, which were nicely compared in a work held by Brodbelt's group [251].

Similarly to other techniques, a hybrid dissociation combining the capabilities of UVPD with ETD (*i.e.* ETuvpD)[252] was reported in top-down applications.

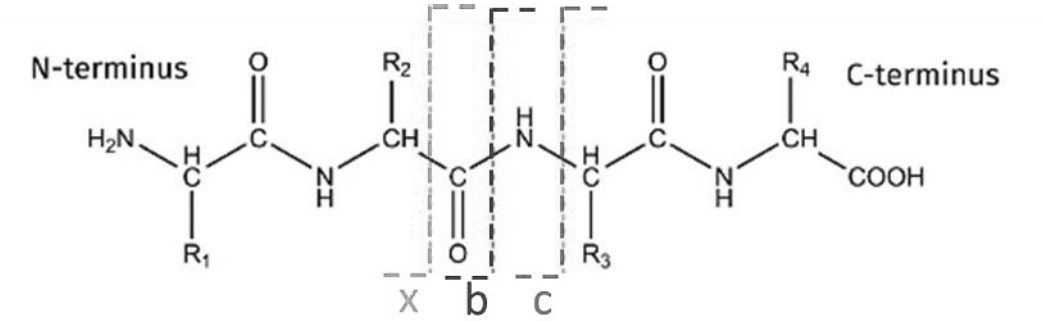


Figure 29: Annotation of peptide fragment ions obtained from UVPD. The a/b/c ions are N-terminal fragments, while x/y/z fragments are C-terminal. CID and HCD predominantly result in b and y ions; electron-mediated techniques, c and z ions; and UVPD, a complex mixture. Adapted from Hale *et al.* (2020). [253]

d) Proton-transfer charge reduction (PTCR)

Although proton-transfer charge reduction (PTCR) is not a fragmentation technique, it gained an increased interest in simplifying top down MS^2 spectra and thus identifying more ion fragments species [225, 254]. This technique is based on ion-ion proton transfer reactions, where multiply charged cation are reacted with singly charged anion (typically perfulorinated species), and thus form deprotonated cations (4).

$$4.\ [M + nH]^{n+} + A^{-} \rightarrow [M + (n-1)H]^{(n-1)+} + HA$$

Although it was first described in the 90s [255], it has only been employed recently in top-down proteomics workflows to yield high sequence coverage. Indeed, the generation of a plethora of fragments from different techniques, provides a crowded spectra with different overlapping peaks. Thus, the deprotonation of those ions through PTCR reactions allows their subsequent detection at high *m/z* range [256]. Therefore, PTCR is normally implemented on state-of-the-art instruments that can reach higher *m/z* values [225, 254].

5.3 Instrumental and data processing considerations for TD/MD-MS

For TD/MD-MS experiments that were commonly performed on FTICR instruments, Orbitrap instruments gained more popularity in recent years for their cost-effectiveness, in addition to their comparable high mass resolution and high mass accuracy than FTICR mass spectrometers. In addition, the rise of various fragmentation techniques compatible with Orbitrap analyzers, allowed selecting Orbitrap MS as the go-to instrument for top-down proteomics. In fact, the first Orbitrap instrument (The Orbitrap Eclipse Tribrid MS, Thermo Fisher) that combined four activation techniques (CID, HCD, ETD, UVPD) with PTCR, in addition to hybrid EThcD and ETciD was evaluated during this manuscript, and will be discussed in details in **Part III**.

TD-MD/MS data processing requires a few steps including: i) intact mass measurement to confirm the MW of the protein sequence, ii) deconvolution of the MS/MS spectra resulted from precursor ions fragmentation and finally iii) assigning the resulting fragments to the candidate sequence [257].

Intact mass measurement is performed as described beforehand. Depending on the used algorithm for fragments matching, the raw MS/MS data can either be converted into monoisotopic mass (M) or single charge mass (M+H+) following different algorithms [257]. This deconvolution takes into consideration several parameters such as the S/N ratio and the fit factor, depending on the quality of the spectra.

When fragment ions are deconvoluted, a monoisotopic peak list is generated and compared with a theoritical peak list generated by the algorithm. The fragments are then matched with the amino acid sequence using stringent mass tolerance (typically 10 ppm for terminal fragments and 2 ppm for internal fragments), which allows to obtain a sequence coverage. The software used in **Part III** of this manuscript (ProSight Lite) [216], uses in addition a scoring system to increase the confidence in the fragment matching with the sequence. This P-score is calculated using the following equation (5):

$$5.\ P_{f,n} = \frac{(xf)^n \times e^{-xf}}{n!}$$

Where $P_{f,n}$ is the probability of random protein matching, f is the number of monoisotopic masses in the input list, n is the number of random fragment ion hits and x is the average probability of the fragment ion mass of random protein matching. Therefore, the low the P-score, the more confidence we obtain in the sequence coverage [257].

Conversely to classical TD/MD-MS softwares, the newly developed ClipsMS software [229] (which will be evaluated during this manuscript) that aim at assigning internal fragments in addition to terminal ones, do not have any scoring system and thus its use is still limited.

6. Combining the benefits of native and top-down MS for mAbs characterization

In the aim of combining the information obtained upon nMS on the quaternary structure of complexes and the information obtained upon TD-MS on the primary structure, native top-down MS (nTDMS) have risen as a promising approach [258-260]. This approach relies on the fragmentation of the gas-phase proteins, in their native state using fragmentation techniques such as ETD and UVPD that do not disrupt the overall complex structure. Despite the limited fragmentation of native intact proteins and noncovalent protein complexes, nTDMS have emerged as a valuable approach for HOS and PTMs characterization of membrane proteins and hemoglobin [261-262]. Ongoing development of tailored softwares and sophisticated equipments are supporting the emerging use of this technique.

7. Peptide-centric approaches

In addition to protein centric approaches that focus on the analysis of protein assemblies at the intact level, complementary peptide-centric techniques are based on the identification and quantification of proteins by detecting and quantifying individual peptides [263].

7.1 Hydrogen/Deuterium exchange-MS (HDX-MS)

HDX-MS is usually used as a differential approach based on the comparison of proteins in their free and bound states upon deuterium incorporation, or of wild-type proteins and their mutants [264-265]. Mostly, solvent-exposed amide protons are the ones targeted in the protein backbone. Briefly, the proteins are first diluted in a deuterated solvent (D_2O) to probe their deuteration uptake at different exposure times. Then the deuteration reaction is quenched using acid pH (2.5) and low temperature (0°C) to avoid back-exchange. Finally, the peptides obtained upon protein digestion (usually using pepsin) are analyzed using LC-MS/MS [102] (**Figure 30**). The information obtained from HDX-MS offers information on interaction regions of proteins partners, upon comparison of different deuteration profiles. In addition, HDX-MS affords conformational dynamics of proteins by assessing solvent accessibility at different deuteration times [102]. In the context of mAb-related products developments and biosimilars assessment, HDX-MS played a pivotal role in being the go-to method for the assessment of epitope mapping of mAb-antigen interactions, ligand/receptor interactions, comparability studies and conformational changes. [130, 266-267].

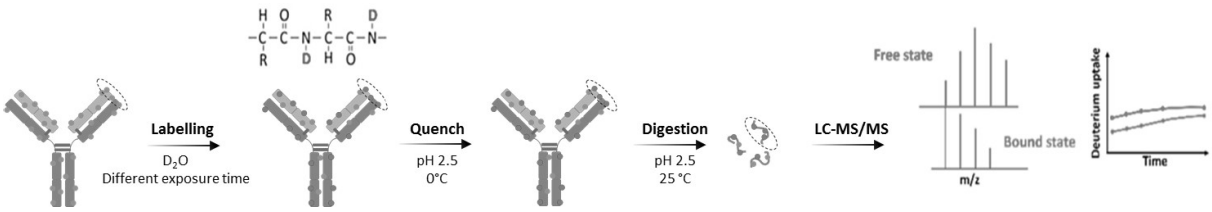


Figure 30: Schematic workflow of HDX-MS approach. Adapted from Castel *et al.* (2023). [102]

7.2 Cross-linking MS (XL-MS)

An additional tool for mapping protein-protein or protein-ligand interactions is XL-MS, which is an irreversible labelling method for higher order structures (HOS) characterization [268]. It depends on formation of new covalent bonds between spatially close residues, by chemically cross-linking proteins and proteins complexes in their native states [268]. The different cross-linking reagents are all composed of a spacer that contribute in indicating the distance between the targeted residues and a reactive group deciding the targeted amino acid. The cross-linked proteins and protein complexes are

enzymatically digested and further analyzed using LC-MS/MS The main information obtained through XL-MS experiments, is about the spatial proximity between the two targeted residues either within one protein (intraprotein cross-links) or between two different proteins (interprotein cross-links). However, even if XL-MS have been promising for the characterization of protein-protein interactions, it is still limited when it comes to mAb-related analyses [269]. A few studies described the used of XL-MS for biotherapeutics characterization, in particular for mAb-antigen binding assessment [267].

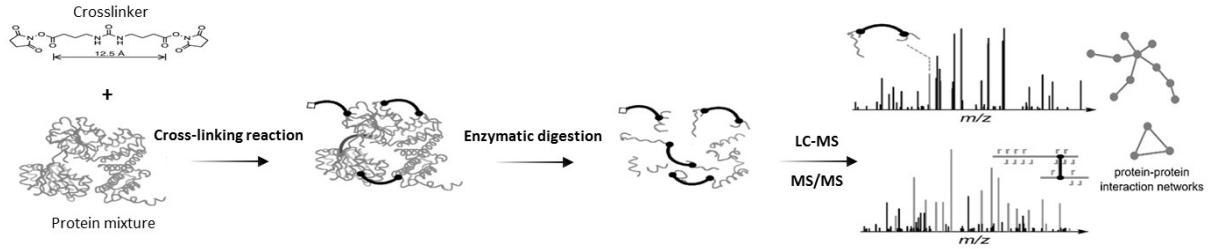


Figure 31: General workflow of XL-MS studies. Adapted from Götze *et al.* (2019). [268]

7.3 Fast photochemical oxidation of proteins MS (FPOP-MS)

FPOP footprinting is another labelling MS method, where proteins are mapped through fast and irreversible chemical reactions. FPOP is based on laser proteolysis of peroxide H_2O_2 [270]. In short, the protein sample, which is flowed with hydrogen peroxide through a fused silica capillary and irradiated by a 248 nm KrF excimer laser, is exposed to oxidation during a short time window. Further, FPOP-labelled proteins are analyzed by LC-MS/MS following their enzymatic digestion (**Figure 32**). In the context of biotherapeutics, FPOP-MS have been used to study mAb-antigen interactions, as well to perform biosimilarity assessment [102, 270-271].

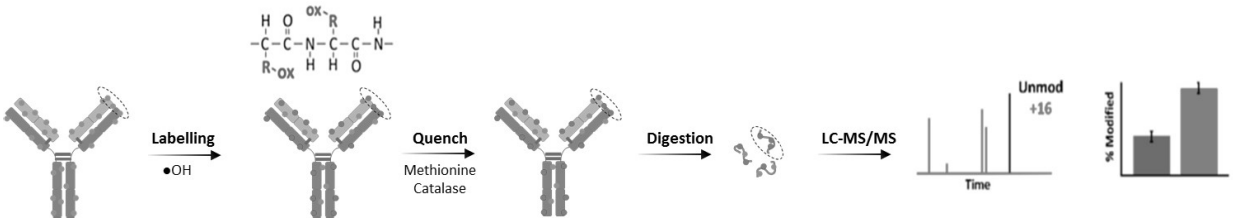


Figure 32: Schematic representation of FPOP-MS approach. Adapted from Castel *et al.* (2023). [102]

7.4 Limited proteolysis MS (LiP-MS)

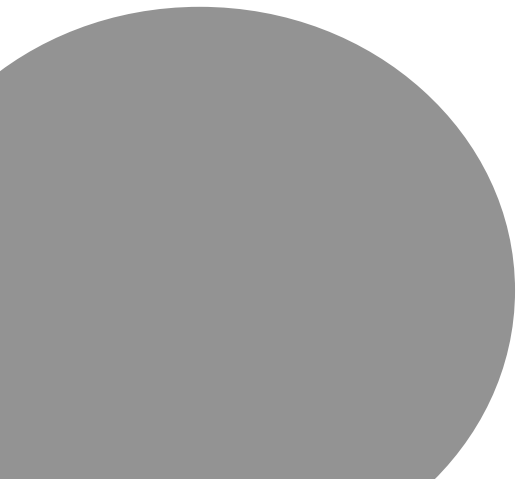
LiP-MS is a technique able to detect subtle alterations in secondary structure content, large-scale movements and more pronounced transitions such as folded and unfolded states or multimerization events [272]. In LiP-MS proteins are treated with a nonspecific protease (such as proteinase K, PK) in their native state, which induces non-tryptic cleavages. The analysis of the resulting peptides reflects chemical structure or state changes of proteins, based on comparison of their relative intensities with intact protein. LiP-MS applications in the field of biotherapeutics, have been proved upon antibody-target protein interaction [273].

8. Conclusions

The need for comprehensive characterization of mAb-related products requires more sophisticated and straightforward techniques. In this context, mass spectrometry has shown to be promising through the combination of several protein-centric and peptide-centric approaches that we have seen through this chapter. In addition to MP that has become promising in the structural characterization of mAbs and mAb complexes. Among the various techniques discussed beforehand, nMS is now the go-to method for therapeutic mAb analysis, as it affords straightforward assessment of the different CQAs required by regulatory agencies. Due to the progressive need of this technique in biopharma companies, it has been combined with separative techniques such as SEC and CEX to gain additional information on the different size and charge variants, while providing a more versatile and automated workflow. In parallel, the interest in TD-MS approach has gained increased interest due to the ability of this technique to correlate the information at the intact level, to that obtained from the primary structure for an improved characterization of protein proteoforms. For biotherapeutics, it is a promising technique to provide PTMs and drug conjugation site assessment by combining different fragmentation techniques and increase the overall sequence coverage. Therefore, during my PhD I focused more particularly on these two techniques (*i.e.* nMS and TD-MS), to provide more straightforward and optimized workflows using state-of-the-art MS platforms. My main **PhD objectives** were:

i. **Improving LC-MS coupling techniques for rapid, automated characterization of ADCs/PDCs Developing new TD/MD-MS approaches for the characterization of the primary structure of these bioconjugates,**

ii. **Developing new TD/MD-MS approaches for the characterization of the primary structure of these bioconjugates.**



Part II: Hyphenation of liquid chromatography to mass spectrometry for mAb-based products characterization

Chapter 1: Evaluation of a new LC-MS system to automate the coupling of non-denaturing chromatography to native mass spectrometry for the characterization of biotherapeutics

1. Analytical context

In the past years, the coupling of non-denaturing LC to native MS has been demonstrated as very promising and useful for the characterization of mAb-based products [10, 121, 125]. For example, size exclusion chromatography (SEC) offers a prompt and efficient on-line buffer exchange of mAb samples, in addition to its ability to characterize size variants [274-275]. On the other hand, ion exchange chromatography (IEX) especially cation exchange chromatography (CEX) has been widely used for the assessment of post-translational modifications (PTMs) based on the charge variants separation [18, 156, 167, 171, 173, 276-277]. More importantly, the hyphenation of those LC techniques to nMS is of utmost interest for the biopharmaceutical field, in order to measure precisely the masses of the different species and confirm their identity while preserving the intact mAb products.

Recent technological advances in SEC-nMS highlighted its increasing use of this coupling, especially for characterizing biotherapeutics using sub-µm columns to yield better separation [153, 278], or using shorter columns to increase the throughput of analysis [279]. In addition to the fast on-line buffer exchange SEC provides when using longer columns, another level of characterization is obtained which is the separation of size variants either, those stemming from different oligomeric states, or from middle level digestion [15]. SEC-nMS was also successfully coupled to ion mobility as additional level of separation for conformational characterization of mAb products through collision cross section (CCS) measurements in few minutes [17, 126-127, 165, 178, 280], to reach another level of separation.

Meanwhile, CEX-nMS have been used for the characterization of biotherapeutics resulting from mAb degraded studies or mAb biosimilars [156, 173, 281-282], using either Time-of-Flight (ToF) or Orbitrap analyzers for higher resolution characterization of charge variants [156]. The information obtained from SEC- and CEX-nMS represent important critical quality attributes that help the approval of an ADC/PDC by regulatory organisms such as the FDA and EMA [8]. Therefore, those CQAs need to be routinely monitored during the development process of new biotherapeutics.

However, both SEC- and CEX-nMS require a certain level of expertise and time-consuming set up of the coupling, ranging from hyphenating the LC system to the MS, to manual calibration of the MS at higher masses and manual data processing. Therefore, the lack of automation of the mentioned couplings make them difficult to implement in biopharmaceutical companies. Consequently, we wanted to implement these workflows on the BioAccord, a benchtop LC-MS platform designed for

biopharma use, in order to completely automatize these workflows, from LC-MS coupling to data treatment, and achieve straightforward characterization of different mAb-based products.

2. Objectives

This part of the thesis work consisted of evaluating a new benchtop LC-MS platform dedicated to biopharma use to automate SEC-nMS and CEX-nMS workflows. This consisted of:

- Implementation of automated SEC-nMS workflow for the rapid assessment of mAbs size variants
- Implementation of automated CEX-nMS workflow for the characterization of mAbs charge variants

3. Presentation of the LC-MS platform

3.1 Schematic of the BioAccord

The BioAccord LC-MS platform is a user-friendly system, commercialized by Waters since 2020 and dedicated for biopharma use for the high-throughput analysis of small and large molecules. It is composed of an Acquity UPLC M-Class liquid chromatography system that includes a binary solvent manager (BSM), a sample manager, a column oven and a UV detector, continuously hyphenated to a ToF mass spectrometer with an RDa™ analyzer. **Figure 33** shows a schematic representation of the instrument.

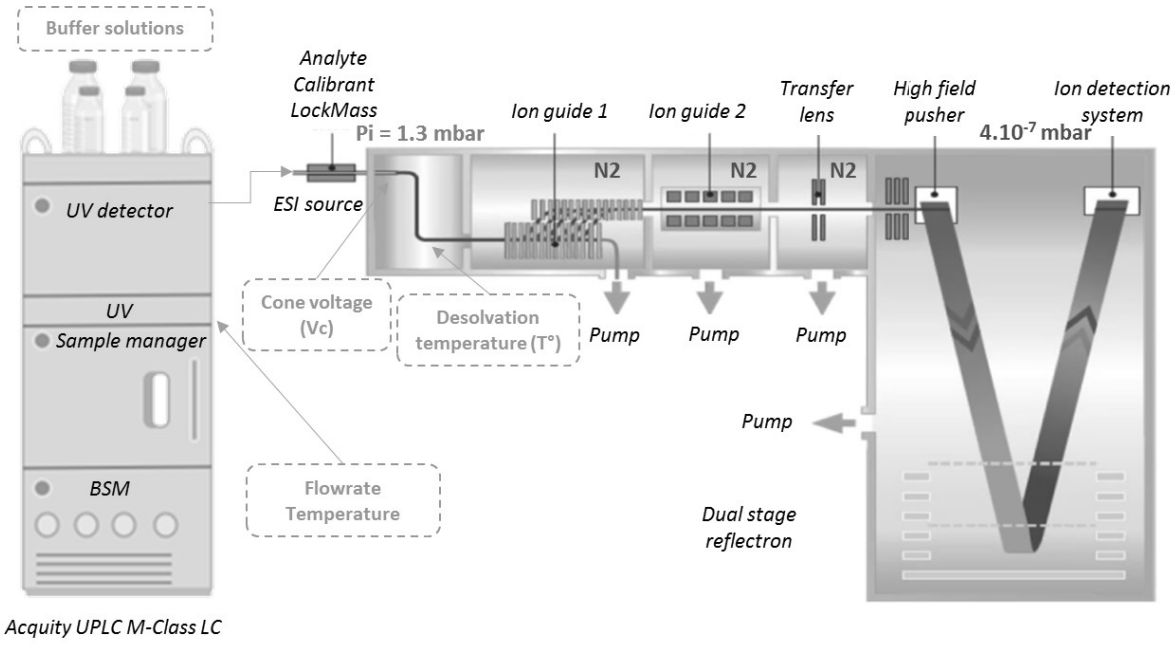


Figure 33: Representative illustration of the BioAccord LC-MS system (Waters). The fixed parameter corresponding to the source pressure (Pi) is in red while the parameters that could be varied i.e. the LC solutions, the column flowrate and temperature, the cone voltage (Vc) and the desolvation temperature (T°) are represented in green.

The MS is composed of different parts including:

- A Zspray™ ion source that allows electrospray ionization using a "Z" trajectory in order to eliminate neutral species and to improve the sensitivity of the instrument. Herein, the backing pressure is fixed at 1.3 mbar, and the cone voltage(Vc) could be modulated either to enhance the transmission and the desolvation of the ions or to activate them [283].
- Two ions guides including a StepWave, which are a series of superimposed lenses with a continuous flow of nitrogen draining the neutral molecules while directing the ions from the electrospray source towards the mass analyzer. A combination of radiofrequency (RF) and a direct current (DC) is used to enable the focalization and transmission of the ions.
- A transfer lens used to focalize the ions and direct them into the analyzer.
- A ToF analyzer with a reflectron constructed in a V form under a vacuum of 4.10^{-7} mbar with a first region where the ions are accelerated with an application of a high potential difference and a second region where the ions travel to reach the detector. The separation of ions based on their times of flight is proportional to their *m/z* ratio [284]. The kinetic energy of the ions in the ToF is determined following the equation (6) :

6. $Ec = zeU = \frac{1}{2}\, mv^2$

Where Ec is the kinetic energy; z is the ion charge, e is the electron charge, U is the difference of potiential, m is the mass of the ion and v is the ion velocity. The ion velocity could be determined by equation (7):

7. $v = \sqrt{\frac{2zeU}{m}}$

Therefore, the time of flight (t) could be related to *m/z* ratio as a function of the flight trajectory (L) following the equation (8):

8. $t = \frac{L}{v} = L\sqrt{\frac{m}{z2eU}}$ ➔ $m/z = \frac{2eUt^2}{L^2}$

The whole system is fully controlled by the UNIFI software on WatersConnect interface which allows the automation of analytical workflows.

3.2 LC key parameters

In addition to all the parameters discussed in the previous paragraph, LC parameters are also important to set prior to any SEC-nMS or CEX-nMS experiment, such as the choice of the mobile phase, the column and the gradient.

For SEC-nMS experiments, many SEC column technologies are currently available on the market. In particular, sub-3µm columns, which have a pore size less 200 Å [153]. They allow achieving improved

peaks resolution and analytes separation in a reduced analysis time. The main drawback of the latter columns is that the separation is affected by the secondary interactions of the analytes with the stationary phase. A new generation of bio-inert SEC columns has been recently developed, the MaxPeak technology, which provides a large specific surface area for size-specific interactions with the analytes, resulting in reduction of non-specific interactions and thus better peak shapes and separation.

The choice of the mobile phase is also at utmost interest, a solvent that is volatile and compatible with mass spectrometry is mandatory to enable the ionization of the proteins in the gas phase. Moreover, the pH and the ionic strength of this solvent are critical to maintain the native conformations of the proteins and proteins complexes [285]. Ammonium acetate is the most suitable buffer for this kind of experiments, which was proven by several studies [286-287]. A concentration of ~100-500 mM is usually enough to allow mAb ionization with good mass resolution. The separation is usually carried out in isocratic elution mode to focus on the size separation rather than the interactions with the stationary phase.

For CEX-nMS experiments, the choice of the column and mobile phase are both important as they affect the separation efficiency, resolution and the selectivity of the chromatographic separation [171]. State-of-the-art cation exchangers are packed with non-porous particles that enhance the peaks shapes and improve the resolution of close eluting species which is a key point in the characterization of mAbs proteoforms[18]. Volatile mobiles phases used during the chromatographic separation are either pH or salts gradients to allow the disruption of ionic interactions [173, 288]. More recently, pH-salt mediated gradients have emerged as a powerful tool for enhanced and more efficient CEX separation, which is the elution mechanism we considered for our further analyses [174].

3.3 MS crucial parameters

The BioAccord is constructed in a way to be used routinely with pre-optimized methods for various applications. This means many parameters were pre-defined by the manufacturer to reduce the calibration time and to facilitate the switch from denaturing to native conditions. The fixed parameters are presented in red and the variable ones are in green in **Figure 33**. The backing pressure (**Pi**), the cone voltage (**Vc**) and the desolvation temperature (**T°**) are crucial parameters to take into consideration to ensure the desolvation and transmission of the analytes in denaturing or native conditions.

- The cone voltage (**Vc**) is a crucial variable parameter that affects the ion's kinetic energy at the entrance of the interface region. Higher cone voltages allow sufficient desolvation of species but may cause the dissociation of noncovalent complexes. Contrariwise, lower cone voltages preserve these interactions but lead to less desolvation which prevent precise and accurate mass measurements [289].
- The desolvation temperature in the source (**T°**) is also an important parameter that can be adjusted. It is the temperature inside the ESI source where the solvent molecules are evaporated. When this value is too high, the analytes might be altered or unfolded, while at low temperature the desolvation is not enough resulting in large peaks and low sensitivity signal.
- The **Pi** is the pressure of the gas applied through the interface of the mass spectrometer (backside of the ion source). Higher pressure values lead to reducing the ions mean free path, which leads to maintaining the native conformation of the species and better transmission of high m/z species. Conversely, lower pressure values result in more energetic collisions of the ions with gas molecules that might lead to the disruption of noncovalent interactions. However, this value is fixed at 1.3 mbar on the BioAccord system, which is one of the disadvantages when working in native conditions.

Therefore, a compromise between the cone voltage (Vc) and the backing pressure (Pi) need to be found to preserve the noncovalent interactions [289].

4. Method development for automated SEC-nMS coupling for the characterization of reference mAbs

4.1 State of the art

SEC-nMS have been widely used by different groups for the characterization of newly developed biopharmaceuticals [15-17, 123-124, 159], such as the latest Trastuzumab Deruxtecan intended for the treatment of breast cancer [121], paving the way to the routine integration of SEC-nMS in biopharma environments. Latest advances illustrated also the hyphenation of SEC-nMS to ion mobility for mAb-based products characterization which emphasizes the automation of this workflow thanks to the online buffer-exchange prior to IM-MS experiments [126 17, 165, 178]. Simultaneously, a high importance have been attached to the development of new SEC columns. In 2017 Goyon et al. evaluated for the first time the new sub-3µm SEC columns which are now widely favored for SEC-nMS analyses of biotherapeutics [153]. Building on recent studies, a new way of performing high throughput analyses is to use a narrower type of SEC columns of an internal diameter of 1 mm with a flowrate of 15 µL/min

[290], which increases the protein-ionization efficiency and facilitates the detection of higher order structures. While the late advances focus on increasing the throughput of SEC-nMS experiments, maintaining the quality of the chromatographic separation should be also carefully considered to ensure accurate mass measurements of the separated species. In this context, a recent SEC separation technology was developed (Maxpeak Premier SEC columns) in late 2022 [157], to reduce nonspecific interactions between the analytes and the stationary phases, which we evaluated during this study.

The BioAccord instrument has proven to be versatile in the field of biopharma with various applications. Notably, it was used for bsAbs characterization employing several LC-MS techniques to provide in-depth characterization of their structure and PTMs [125]. In addition, it excelled in MS-based multiple-attribute monitoring (MAM) studies of mAbs, allowing a comprehensive characterization and quantification of a series of CQAs [116]. N-glycoprofiling studies is another area where the BioAccord was evaluated, aiding the analysis of glycosylation patterns that are critical for biopharmaceutical safety of mAbs [291].

4.2 Evaluation of SEC bio-inert columns for optimal LC separation

The usual columns used for SEC-nMS experiments are built following the Ethylene Bridged Hybrid (BEH) particle technology which was the first available sub-2-µm SEC chemistry commercialized by Waters. This packing material provides a stable chemistry with minimal secondary interactions. However, due to the high active surface of some biotherapeutics that leads to interactions with the column hardware, a recent technology have been developed to suppress those nonspecific interactions and improve the SEC separation. The Maxpeak protein column are packed with BEH particles in addition to a coverage with hydroxyl-terminated polyethylene oxide (PEO) that acts as a barrier surface to eliminate nonspecific interactions, and thus to achieve inertness of the SEC column. **Figure 34** depicts the two different columns technologies.

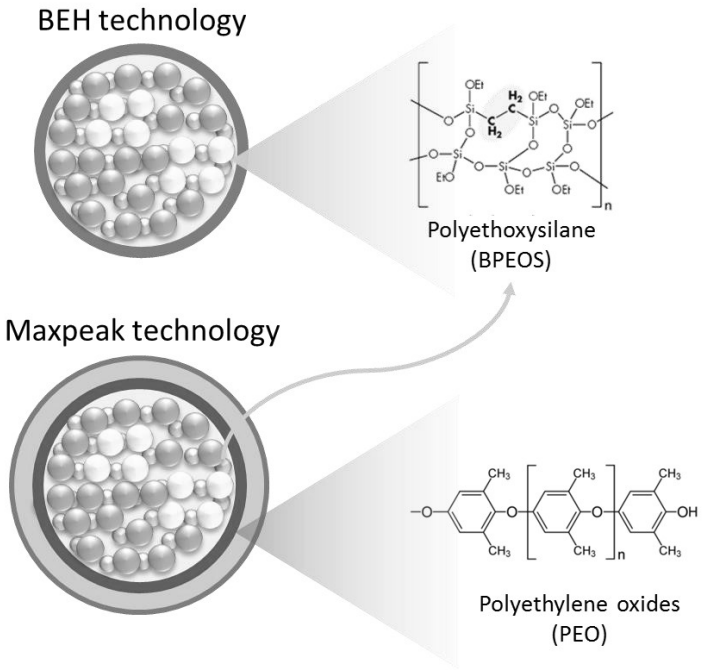


Figure 34: BEH technology compared to Maxpeak technology columns.

This Maxpeak technology has been applied to different columns (RPLC, SEC), however we were mainly interested in the SEC columns. **Table 4** resumes the characteristics of the Maxpeak SEC column we used compared to the classical BEH SEC column.

Table 4: Comparison between the characteristics of BEH and Maxpeak SEC columns (Waters) showing similar features for both columns, except for the particle technology that is modified for Maxpeak columns to increase the bio-inertness and reduce non-specific interactions, and the pore size that is slightly higher for the latter column.

Column	BEH	Maxpeak
Particle technology	BEH particles	BEH-PEO particles
Pore size	200 Å	250 Å
Particle size	1.7 µm	1.7 µm
Inner diameter	4.6 mm	4.6 mm
Length	150 mm	150 mm
pH range	2.5-8	2.5-8

We thus explored the use of those last generation bio-inert SEC columns, in comparison to classical BEH SEC columns. By doing so, we analyzed a reference glycosylated mAb (trastuzumab) at the intact and middle levels, using our go-to column (BEH SEC 200 Å 1.7 µm, 4.6 x 150 mm) in a first place, then by using the bio-inert column (Maxpeak BEH 250 Å 1.7 µm, 4.6 x 150 mm). Both columns were operated at a flowrate of 0.25 mL/min at room temperature, using a mobile phase containing 150 mM $AcONH_4$ pH 6.9.

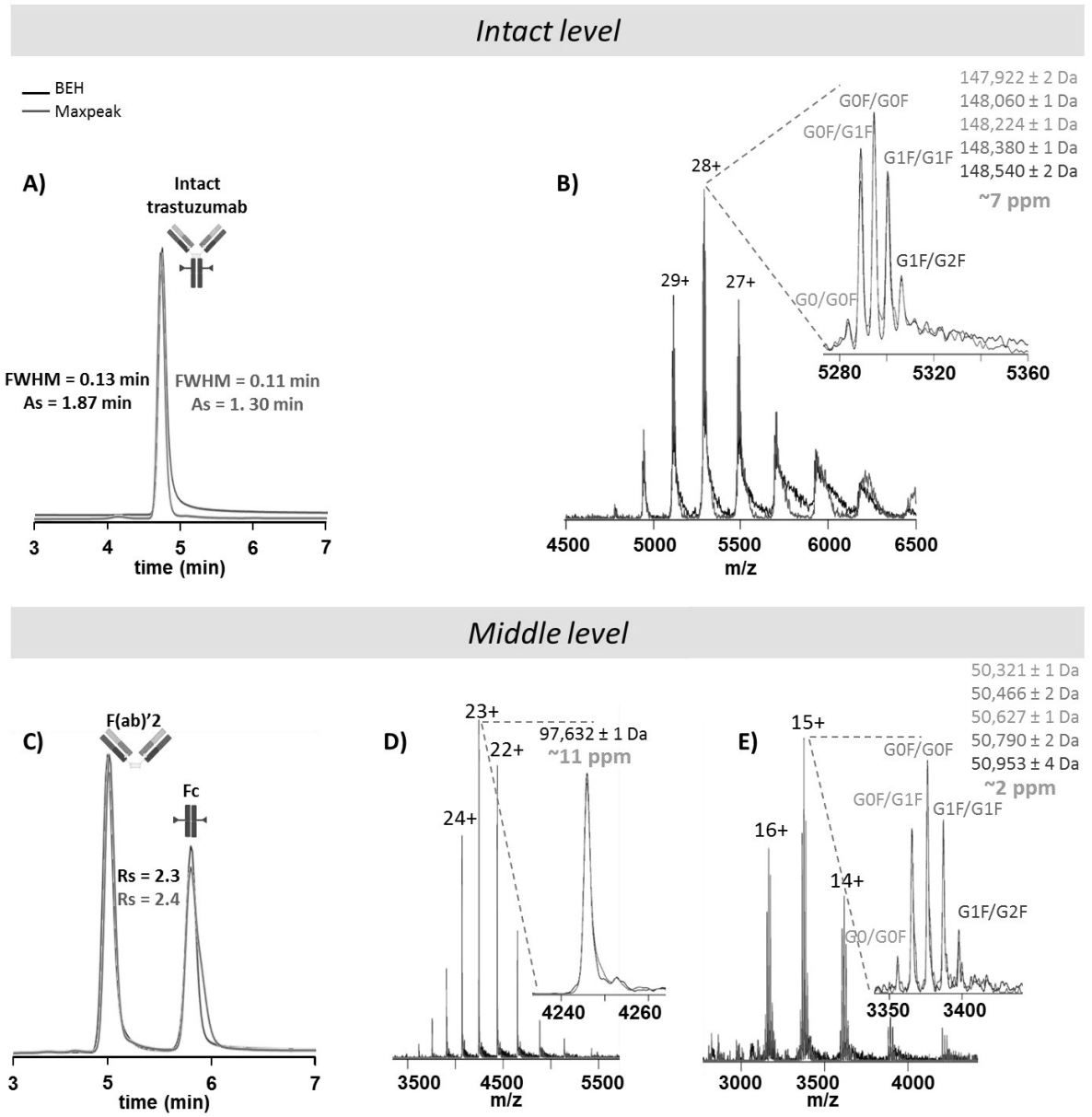


Figure 35: SEC-nMS analysis of trastuzumab on the BioAccord LC-MS system using BEH SEC column (black) versus Maxpeak SEC column (blue). SEC-UV chromatograms of **(A)** intact trastuzumab and **(C)** IdeS digested trastuzumab showing one homogeneous peak for intact mAb and two peaks for digested mAb corresponding to the separated F(ab)'2 and Fc subunits. Native MS spectra reveals the identity of each species namely **(B)** intact trastuzumab, **(D)** F(ab)'2 fragment and **(E)** Fc subunit with zoom on the most intense charge states highlighting the presence of glycoforms for intact mAb and at the Fc subunit level. The remarkable mass accuracy of less than 30 ppm for intact mAb and less than 15 ppm for subunits is highlighted in this study.

SEC-UV chromatograms reveal one peak at the intact level corresponding to intact trastuzumab eluting at ~4.7 min and two peaks at the middle level corresponding to F(ab)'2 fragment at ~5.0 min and Fc subunit at ~5.8 min (**Figure 35A, 2C**). Interestingly, when using the Maxpeak column the peaks were slightly sharper (FWHM = 0.11 min versus FWHM = 0.13 min with BEH column, for intact trastuzumab peak). The symmetry of the peaks was also enhanced when using the Maxpeak column as for the peak of intact trastuzumab the peak tailing was reduced (As = 1.30 min versus As = 1.87 with BEH column). Both columns exhibited overall good baseline separation in the case of middle level subunits analysis,

with slightly better separation when using the Maxpeak column (Rs = 2.4 versus Rs = 2.3 with BEH column). No significant difference was observed in terms of mass measurement accuracy, as for the intact trastuzumab (148,224 ± 1 Da for G0F/G1F), the F(ab)'2 (97,632 ± 1 Da) and the Fc subunit (50 627 ± 1 Da for G0F/F1F), the mass accuracy was less than or equal to 7 ppm when for intact trastuzumab, ~11 ppm for F(ab')2 subunit and ~2ppm for Fc species using both columns. However, a slight improvement in the MS peak resolution was noted when using the Maxpeak column (FWHM = 1.90 *m/z* versus FWHM = 2.01 *m/z* for G0F/G1F peak). Altogether, these results clearly demonstrate that the Maxpeak Protein SEC column provide better performances for intact and middle level analyses. This is due to the high porosity of the resin, which provides a large specific surface area for size-specific interaction with proteins, thus decreasing the unspecific interactions. Overall, this comparison allowed us to select the Maxpeak SEC column for further mAb analyses.

4.3 Optimization of MS parameters to achieve optimal native conditions

1) Tuning of MS parameters to maintain mAbs native-like structure

After optimization of the LC conditions, we next evaluated the MS parameters in order to assess their influence on better transmission of the studied mAbs under native conditions. As discussed earlier, only the cone voltage (Vc) and the desolvation temperature (T°) could be tuned to increase the transmission of high *m/z* ions while preserving the intact structure of the mAb. When the Vc value is set at 60V, we obtain insufficient desolvation leading to large MS peaks. On the other hand, using higher cone voltage values could lead to the dissociation of the intact structure of the mAb of interest, thus this value was fixed at 120V for our reference trastuzumab analysis (**Figure 36**).

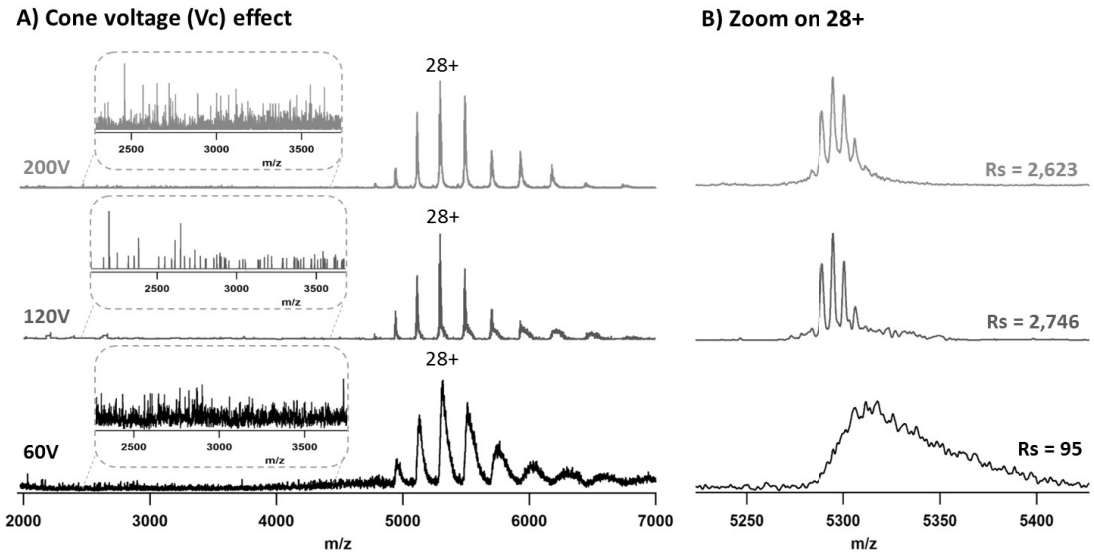


Figure 36: Effect of cone voltage for intact trastuzumab upon SEC-nMS analysis. **(A)** Full nMS corresponding to intact trastuzumab with different voltages. When using 60V the desolvation the desolvation is insufficient and at 200V we start to observe some dissociated fragments. The optimal Vc fixed at 120V show better desolvation and no significant dissociation. **(B)** Zoom on the most intense charge state 28+ showing that at the optimal conditions the different glycoforms are observed with better FWHM of the most intense species, in contrast to other conditions.

The desolvation temperature needs to be monitored as well. As explained previously, lower temperatures prevent desolvation of high *m/z* ions into the analyzer, thus the presence of adducts results in very heterogeneous and wide peaks. This is highlighted in the example of intact trastuzumab analysis at the optimized cone voltage of 120V and a desolvation temperature of 100 °C where almost no signal is observed. Conversely, at higher desolvation temperatures such as 500°C, which is the maximum temperature on the BioAccord, no significant difference is observed and comparable resolution (at 5,300 *m/z*) was observed. This parameter was thus fixed at 300°C in our case, as higher temperature do not provide an additional benefit (**Figure 37**).

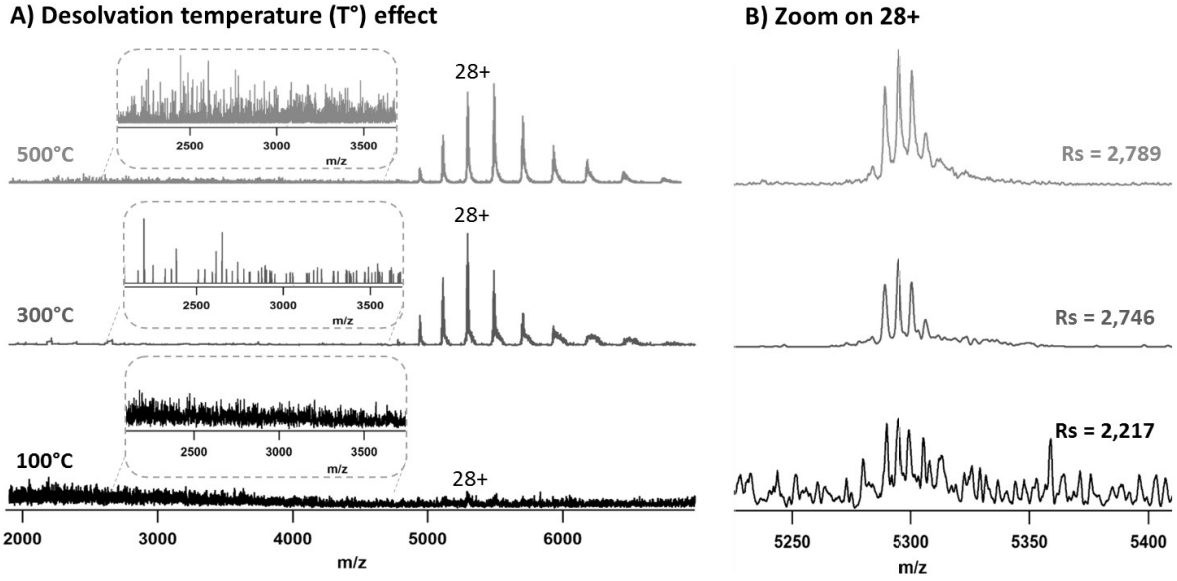


Figure 37: Effect of desolvation temperature on intact trastuzumab SEC-nMS analysis. **(A)** Full nMS spectrum at 100°C show no desolvation, at 500°C a beginning of dissociation is observed, while at 300°C a sufficient desolvation is observed with no sign of dissociation. **(B)** Zoom on the most intense charge state 28+ showing that at the optimal conditions the different glycoforms are observed with better FWHM of the most intense species, in contrast to other conditions.

Overall, it is important to find a compromise of the key parameters discussed beforehand to keep the integrity of the trastuzumab on the BioAccord LC-MS system. Therefore, for further experiments, the MS parameters for intact mAb analysis will be fixed at 120V and 300°C for cone voltage and desolvation temperature, respectively. Those optimal conditions allow obtaining enough desolvation of the mAb while keeping its native-like state in the gas phase.

2) Robustness, sensitivity and reproducibility of the BioAccord

To emphasize the robustness of the BioAccord instrument, we compared it with our go-to instrument for SEC-nMS analysis of intact mAbs within our laboratory. **Figure 38** shows a comparison of an intact trastuzumab SEC-nMS analysis using the optimized parameters on the two platforms; i) an Acquity UPLC H-Class hyphenated to a quadrupole time-of flight MS (Synapt G2) versus ii) the BioAccord in the optimized conditions.

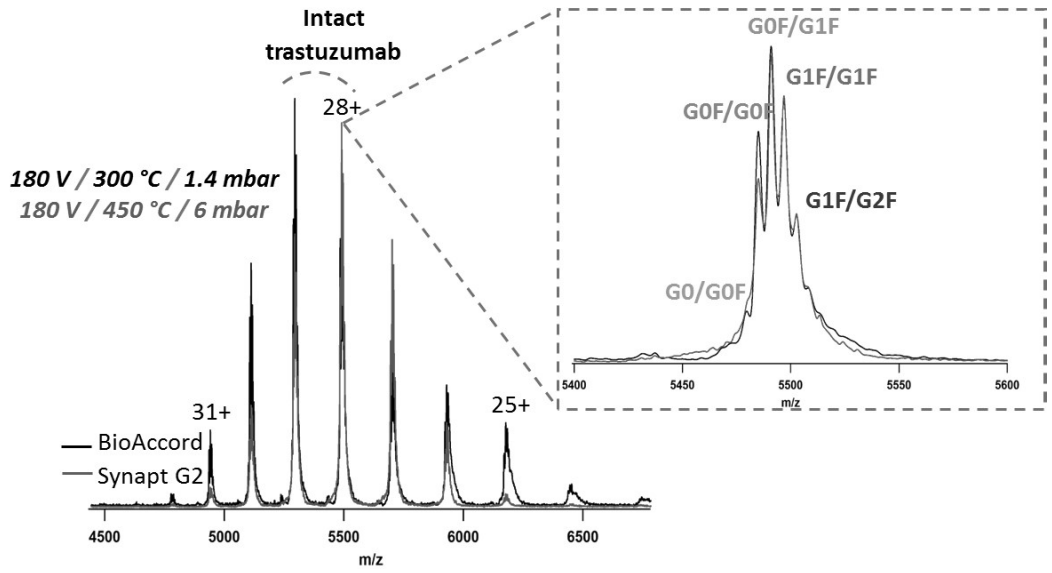


Figure 38: Comparison of SEC-nMS results of an intact trastuzumab on the BioAccord LC-MS system (black) versus Synapt G2 MS coupled to an Acquity I-Class liquid chromatography (blue) in their optimal conditions, showing similar MS profiles in both cases with a homogeneous distribution centered on the 28+ for the Synapt G2 and centered on the 29+ for the BioAccord highlighting the difference of ionization due to the difference in source pressure (Pi). On the grey dots square, zoom on the 28+ charge state showing the different observed glycoforms in both cases.

The obtained MS spectra from both systems exhibit remarkable mass accuracies (~40 ppm for G0F/G1F) with unambiguous glycoforms determination mostly due to the powerful combination of SEC separation to native MS, which results in significantly thin MS peaks. This encourages us to use the BioAccord system for our further SEC-nMS analyses without hesitation, even if its resolution (10,000 at 200 *m/z*) is slightly lower than of the Synapt G2 (until 40,000 at 200 *m/z*).

The sensitivity of the instrument was also evaluated as we managed to decrease the injected quantity of intact trastuzumab to 2 µg, allowing us to characterize the different glycoforms (**Figure 39**). Indeed the high mass accuracy and sensitivity of the BioAccord, allowed us to achieve a straightforward characterization of the reference studied mAb.

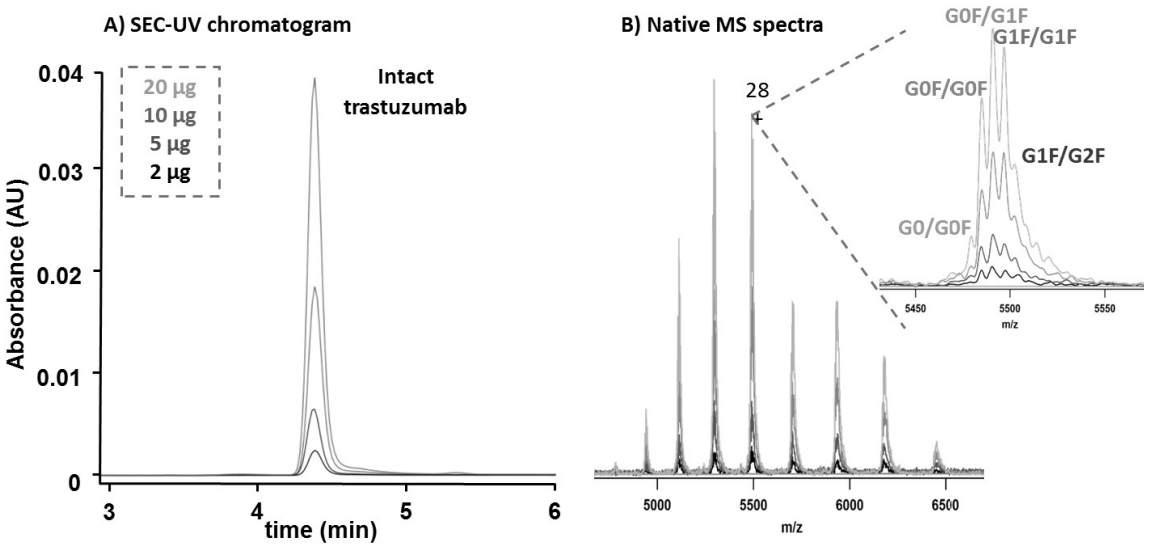


Figure 39: SEC-nMS analysis of intact trastuzumab on the BioAccord LC-MS system in the optimized conditions (Vc = 120V and T° = 300°C), with different quantities injected ranging from 2 to 20 µg**. (A)** SEC-UV chromatogram of the different injections showing that the intensity of the chromatographic peak decreases proportionally with the injected quantity, as expected. **(B)** The corresponding nMS spectra of each peak revealing the identity of the different glycoforms, focused on the charge state 28+, and showing that even at 2 µg the glycoforms could be characterized.

Finally, we benchmarked SEC-nMS analysis of 10 reference mAb samples to highlight the reproducibility of our results. **Table 5** summarizes the mass accuracies and resolutions obtained upon analyses of these mAb products on the BioAccord LC-MS system.

Table 5: Benchmarking of SEC-nMS analysis on the BioAccord LC-MS system of 10 reference mAbs with different PIs and different theoretical masses**.** The experiments were performed in the prior optimized conditions of the instrument. The experimental masses were observed upon precise mass measurement with mass accuracies ranging from 73.0 to even 5.5 ppm. Those analyses highlight the high resolution of the MS with resolution of the 28+ charge state above 1200 overall.

mAb	PI	Experimental mass (Da)	Theoretical mass (Da)	Mass accuracy (ppm)	Resolution based on 28+ charge state
Trastuzumab	9.0	148,060 ± 1	148,056	27.0	2,849.9
Ipilimumab	9.2	147,991 ± 1	147,992	6.5	3,410.6
Bevacizumab	8.3	149,202 ± 0	149,191	73.1	2,211.4
Pertuzumab	9.0	148,100 ± 0	148,102	13.0	2,531.2
Avelumumab	8.5	146,598 ± 1	146,594	24.8	2,103.0
Pembrolizumab	7.6	148,895 ± 1	148,887	54.7	2,859.4
Durvalumab	8.9	148,972 ± 0	148,973	5.5	1,822.4
Nivolumab	8.0	146,235 ± 1	146,240	31.9	3,146.6
Panitumumab	6.8	147,104 ± 1	147,094	68.0	1,208.1
Atezolizumab	8.8	144,367 ± 1	144,356	73.7	2,104.8

4.4 Conclusions

Altogether, these first results highlight the high resolution, sensitivity and mass accuracy of the BioAccord LC-MS system. Furthermore, the versality and automation of the SEC-nMS analysis method, which takes ~1 hour for 10 mAbs analysis, makes it an interesting workflow to be implemented for quality control studies in biopharma laboratories.

5. Development of CEX-nMS method for reference mAb charge variants analysis

5.1 State of the art of CEX-nMS

The post-translational modifications (PTMs) occurring during the development of new ADCs are considered as critical quality attributes (CQAs) [113], they must fall within specific ranges before product release. Charge variants are among the CQAs that can affect the drug efficacy and safety and therefore need to be monitored. Historically, cation exchange chromatography (CEX) coupled to ultraviolet (UV) detection was the reference technique used for the qualitative and quantitative characterization of charge heterogeneity during the quality control of biotherapeutics [18]. The buffers used for the separation are limited to their incompatibility with mass spectrometry, thus the fractions need to be buffer exchanged before their MS analysis, which is time consuming. The use of volatile compatible solvents made it possible for on-line MS coupling, where CEX-MS was used for mAbs charge variants characterization [281]. Thereafter, CEX to native mass spectrometry (CEX-nMS) have been described as the go-to analytical method for the determination of multiple acidic and basic species [19]. Several studies demonstrated that the use of MS compatible salts (such as ammonium acetate) using pH gradients resulted in sufficient separation of the desired species [173, 282], albeit the addition of salt gradients called salt mediated gradients proposed recently by Yan et al. (2018) [155] allowed to increase the sensitivity and the identification of minor variants. Beyond the coupling of CEX to classical resolutive mass spectrometers such as quadrupole time-of-flight (Q-ToF) MS, Fussl et al. (2020) reported the hyphenation of CEX to high-resolution nMS implemented on Orbitraps, the remarkable mass accuracy allowed obtaining higher isoform coverage. In a recent study conducted by Van Schaick et al. (2023), CEX-MS was coupled to ion mobility in order to map the conformational changes induced by PTMs on different IgGs through CEX-CIU-MS experiments. However, the mentioned workflows above require considerable expertise in order to perform the instrumental coupling and to adapt the LC and MS parameters to the mAb type. In this context, we aimed at implementing CEX-nMS on the BioAccord LC-MS platform to automate the workflow and increase the throughput of our analyses.

5.2 Optimization of LC parameters and MS parameters for intact mAb CEX-nMS analysis

The starting point for implementing a robust and versatile CEX-nMS workflow is the CEX separation. To achieve this, we used in-house salt-mediated pH gradient; mobile phase A consisted of 50 mM $AcONH_4$ at pH 5.0 and mobile phase B was 160 mM $AcONH_4$ at pH 8.6. For the intact and IdeS digested trastuzumab analyses, a linear gradient from 50 to 70% B in 10 minutes was applied. **Figure 40** summarizes the obtained results at the intact level.

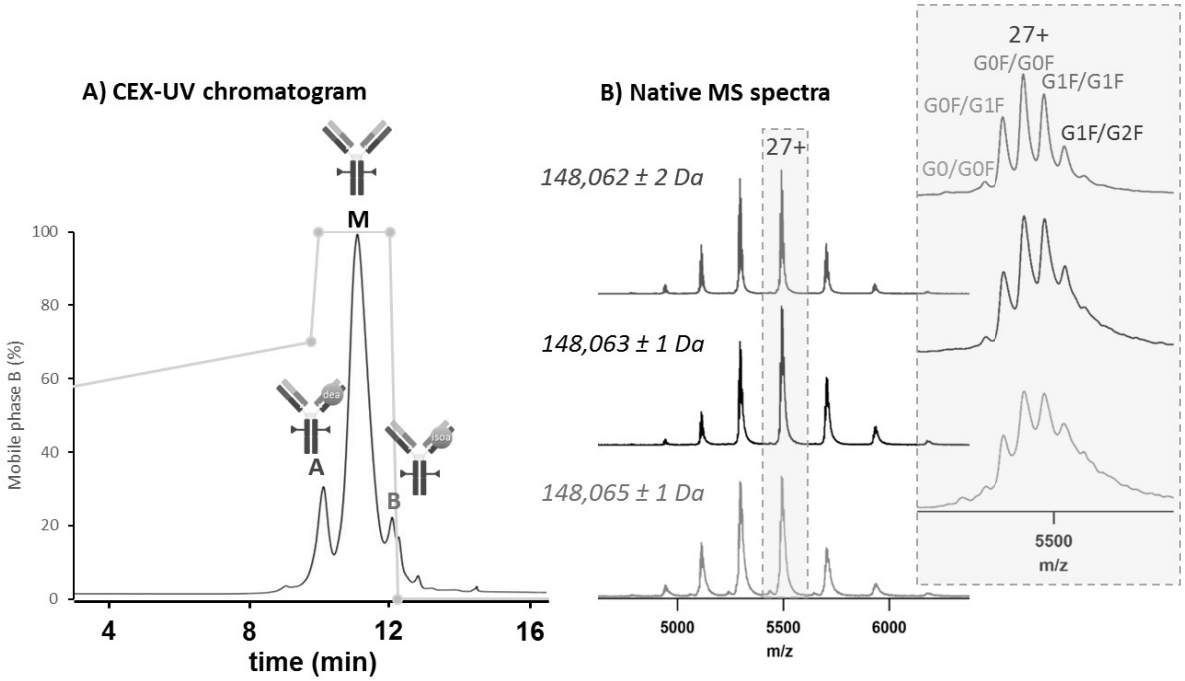


Figure 40: CEX-nMS analysis of intact trastuzumab. **(A)** CEX-UV chromatogram showing the different separated species with main peak, acidic and basic variants. **(B)** The corresponding nMS spectra of the identified species, black spectrum is the distribution of the main species, blue is the acidic variant and green is the basic variant, the given mass corresponds to G0F/G0F species. A zoom on the charge state 27+ shows the different observed glycoforms.

According to the UV chromatogram obtained upon CEX separation, many species could be identified at the intact level that could correspond to the different charge variant species. The main peak eluting at 11.06 min with FWHM of 0.65 min was clearly separated from other species with an acceptable resolution (Rs = 2.01 with acidic species eluting at 10.09 min). The nMS data revealed that this peak corresponded to the main species of trastuzumab (peak M**, Figure 40A**) with a mass of 148,063 ± 1 Da for G0F/G0F within a 40 ppm mass accuracy. The acidic species could correlate to deamidated variants (+1 Da) while the basic species may correspond to isomerization of Asp residues as reported in literature [19, 155]. Overall, the global information obtained from the intact level analysis gives us an idea about the coexistence of several species. However, the limited resolution of the BioAccord hampers the confirmation of variants with low mass differences, as the observed masses are very similar.

5.3 Implementation of CEX-nMS analysis for middle-level mAb characterization

To confirm further the PTMs prior detected at the intact level, we analyzed the same trastuzumab sample upon IdeS digestion to obtain information at the subunit level. The use of the same gradient and chromatographic conditions allowed a clear separation of the Fc and F(ab')2 subunits (Rs = 16.48 between the main Fc peak and the first eluting variant of the F(ab')2). The Fc fragment was observed as a single species eluting at ~5.69 min while the F(ab')2 exhibited 3 different peaks corresponding to acidic variants, main species and basic variants eluting at ~10.10 min, ~11.47 min and ~12.30 min, respectively. The nMS spectra show the deconvoluted masses obtained within a mass accuracy of 20 ppm for the Fc subunit (**Figure 41B**) with the different expected glycoforms. The F(ab')2 fragment revealed masses that could correspond to one deamidation (+1 Da, 40 ppm) and one isomerization (50 ppm) suggesting that the PTMs occur on this subunit. Although the detected masses are very similar, the analysis on the subunit level provided us with specific information on the PTMs type and localization.

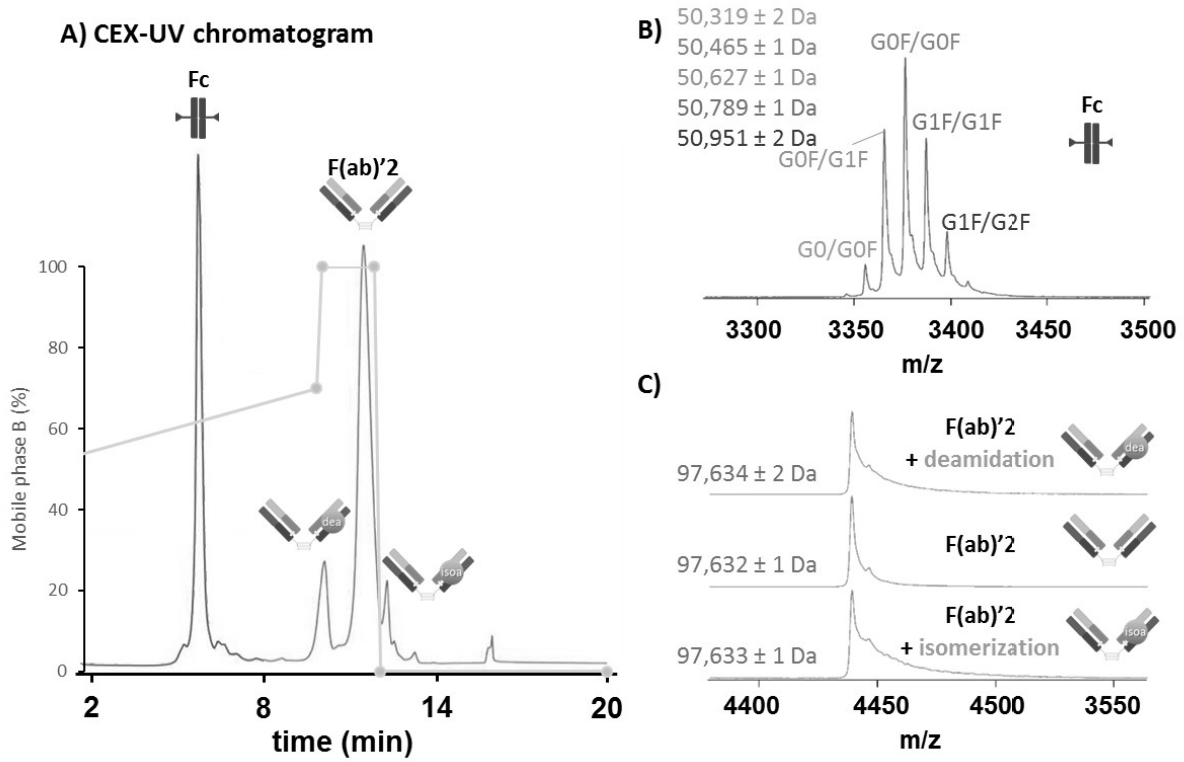


Figure 41: CEX-nMS analysis of IdeS digested trastuzumab. **(A)** CEX-UV chromatogram of the separated subunits showing one Fc species and three F(ab)'2 species confirming the acidic and basic variants are F(ab)'2 species. **(B)** The corresponding nMS spectra of the Fc subunit zoomed on the most intense charge state showing the different glycoforms. **(C)** A zoom on the most intense charge state of the F(ab)'2 fragment from nMS spectra showing the different PTMs occurring on this subunit, confirming PTMs occur on the F(ab)'2 level.

Overall, our implemented CEX-nMS method on the BioAccord LC-MS for the analysis of mAbs, at the intact and subunit levels, pinpoint to the suitability of this instrument for this approach. However, the trickiness of characterizing low mass difference PTMs require more optimization to enhance the

separation of the different species. In some cases, the addition of acetonitrile to IonHance mobile phases improves slightly the chromatographic resolution and peak shape; however, an uncontrolled amount of this organic solvent could potentially disrupt the native structure of antibodies. We therefore kept the optimized method for future CEX-nMS experiments [19].

5.4 Conclusions

The implementation of CEX-nMS workflows allowed straightforward characterization of a reference trastuzumab at the intact and middle level, allowing fast assessment of acidic and basic variants. Our results confirm the applicability of this workflow on the BioAccord system; however, the reproducibility of the CEX-nMS analysis was tricky, probably due to the high possibility of generating more adducts in CEX-nMS due to less desalting efficiency compared to SEC-nMS. Therefore, additional studies will be performed during the next months to tackle this limitation.

6. Establishment of high throughput analysis and data processing methods for biopharma use

The automation of SEC-nMS and CEX-nMS workflows consist on implementing handy analysis and processing methods. The latter is possible with the UNIFI software through the WatersConnect interface that allows us to create automatic processing methods. For mAb products characterization upon SEC-nMS or CEX-nMS analysis, our automated processing methods consist on i) peak integration for quantification of the different size or charge variants, ii) intact mass deconvolution of the species among the different chromatographic peaks and iii) determination of D0 quantity, avDAR and DLD in case of the analysis of conjugated mAbs aka antibody-drug conjugates that will be discussed in the next chapter (ADCs).

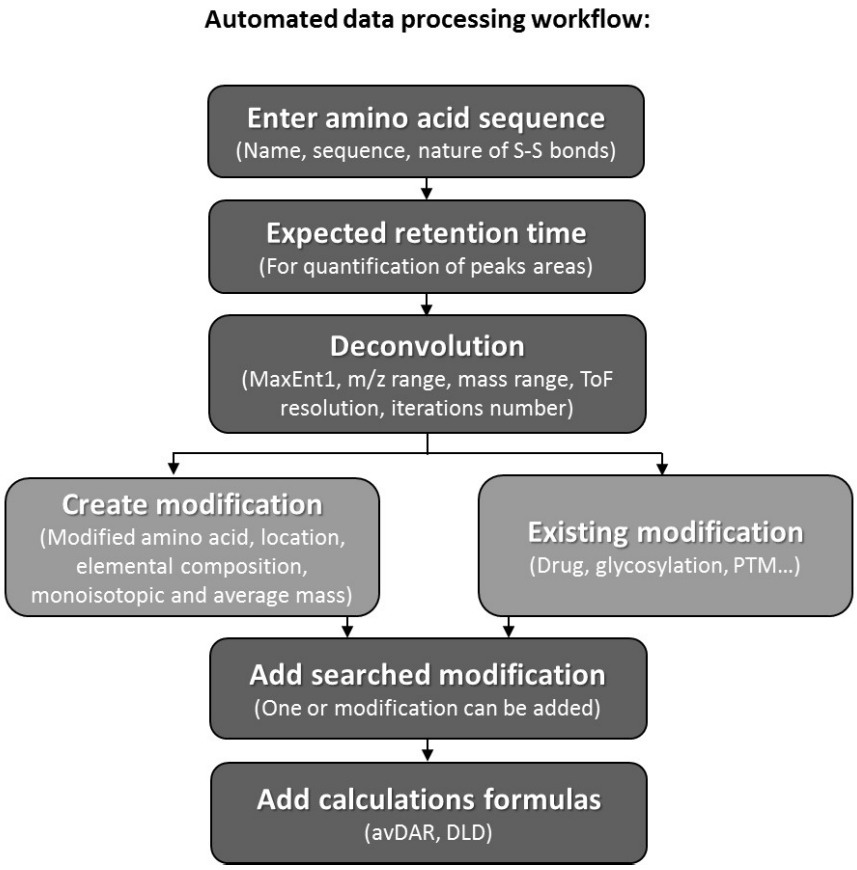


Figure 42: Schematic of the automated data processing workflow developed for SEC-nMS and CEX-nMS analysis on the Unifi software. The first step consists in giving information about the sequence. The second step corresponds to entering the expected retention time of the eluted species that will serve for quantification matter. Next, to allow the deconvolution of the quantified species, m/z and mass ranges should be noted with deconvolution type such as MaxEnt1. In case some modifications are expected, such as PTMs, glycosylation or payload in case of ADCs, the modification should be added to the modification library. Finally, formulas that will help calculating avDAR and DLD should be addressed.

Figure 42 shows an example of the workflow we implemented for SEC-nMS and CEX-nMS data processing. The first step is to enter the amino acid sequence of the analyzed component, specifying whether the disulfide bonds are oxidized or reduced which automatically calculates the expected mass. This gives us a direct information on the mass accuracy when comparing with the species theoretical mass. Next, the expected retention time range of the species of interest and the potential subunits should be determined, in order to integrate the peak areas and quantify relatively each species. Each integrated species is further deconvoluted based on the resulting raw MS spectra. For deconvolution, we decided to use the maximum entropy method (MaxEnt1) which produces simple molecular mass spectra that account for acquired m/z data within a specified mass range. There are several parameters to take into account, such as the ToF resolution that we recommend to be fixed at 10,000 and the iterations numbers at 20 iterations if the quality of the spectra is good, if not, increasing the iterations numbers could increase the measurement confidence. The most important step for ADCs characterization is the determination of CQAs namely the quantity of D0, avDAR and DLD. To do so, the potential modifications should be prior created in the scientific library by determining the modified

amino acid (specific or all amino acids) and its location (side chain, N-terminus or C-terminus), the elemental composition of the drug/payload and the monoisotopic or average mass of this latter. One modification or more could be used during the processing search. Finally, the validation step requires adding formulas for the calculation of the average DAR and the DAR quantities based on the MS deconvolution intensities for DLD determination. Our developed processing method was easily combined with our prior optimized SEC-nMS and CEX-nMS analysis methods, which allowed the characterization of 10 ADCs in less than 1 hour for SEC-nMS, and less than 3 hours for CEX-nMS.

Overall, this workflow is very useful for the batch analysis of a mAb stemming from different conjugation conditions using the same payload, in order to screen the optimal conjugation strategy. It is also applicable for quality control (QC) analysis to highlight the effect of stress on mAb products. However, for ADCs resulting from different conjugation strategies applied on different mAb samples, the automatic processing is challenging as the peaks are eluting in different retention times, with different payload and unconjugated mAb masses, thus manual processing is required.

7. Conclusions

In this chapter, the new benchtop BioAccord platform have been evaluated through SEC-nMS and CEX-nMS experiments. The handy tuning of few LC parameters such as the choice of the SEC and CEX column, buffers and gradients with the optimization of two main MS parameters namely the cone voltage and the temperature of desolvation made it our go-to instrument for routine analyses. The robustness, sensitivity and reproducibility of SEC-nMS experiments was highlighted through the characterization of reference mAbs. For CEX-nMS, the adaptability of this workflow on the BioAccord have been demonstrated, albeit more studies on other mAbs need to be conducted to prove the reproducibility of these experiments. A workflow of the automated data processing was also suggested in the end of this chapter for mAbs characterization in the aim of the versatile integration of UV peaks with deconvolution of the integrated species. Additionally, the automated analysis and processing methods developed could be useful for the routine quality control of ADCs that will be discussed in the next chapter.

Chapter 2: Application of the automated SEC-nMS method for the characterization of site-specific ADCs/PDCs derived from Ugi conjugation

1. Analytical context

Biopharmaceutical companies focus on the development of new class of therapeutics namely protein-drug conjugates (PDCs) [86, 105] and antibody-drug conjugates (ADCs) [1, 5-6, 70-71]. These products are distinguished by their high specificity and drug efficiency compared to their unconjugated counterparts. The oncological field is directly benefiting from these entities as many were intended for cancer therapy [110, 292]. Taking into consideration their high molecular weight, their complexity and their heterogeneity resulting from their bioconjugation, powerful analytical tools have to be developed for their straightforward characterization [17, 121-122, 124]. In the frame of our ITN Targeted Cancer Therapies (TACT) program, we collaborated with Chaubet's group who aims at developing a new generation of ADCs and PDCs resulting from Ugi reaction [32]. The need to screen several reaction conditions (pH, temperature, reaction time, concentration, etc.) to improve the conjugation strategy makes it necessary to use versatile, rapid and automated LC-MS workflows to be able to analyze ~20 samples/day.

2. Objectives

Herein, we applied our previously optimized SEC-nMS method on the BioAccord LC-MS system, to monitor and optimize the Ugi reaction on different systems:

- mAb-based conjugates (~150 kDa) resulting from Ugi reaction using different conditions, to determine their precise mass along with informative CQAs
- Newly developed PDCs: Anticalin-drug conjugates (AcDCs, ~20 kDa) generated upon Ugi reaction to investigate the applicability of this strategy on smaller proteins

3. Optimizing Ugi reaction for ADCs development using SEC-nMS

3.1 The Ugi reaction

The Ugi multicomponent reaction used by Chaubet's group is a two-steps reaction: 1) simultaneously conjugating covalently two amino acid residues; the side-chain amine and the carboxylate groups of two neighboring lysine and aspartate/glutamate with and azide-containing aldehyde and an isocyanide, 2) followed by a linker-drug derivatization through strain-promoted azide-alkyne cycloaddition (SPAAC) [32, 293-294]. In our case, the mAbs were conjugated to BCN-iminobiotin giving access to incorporation of a payload with an average mass of 769 Da (**Figure 43**).

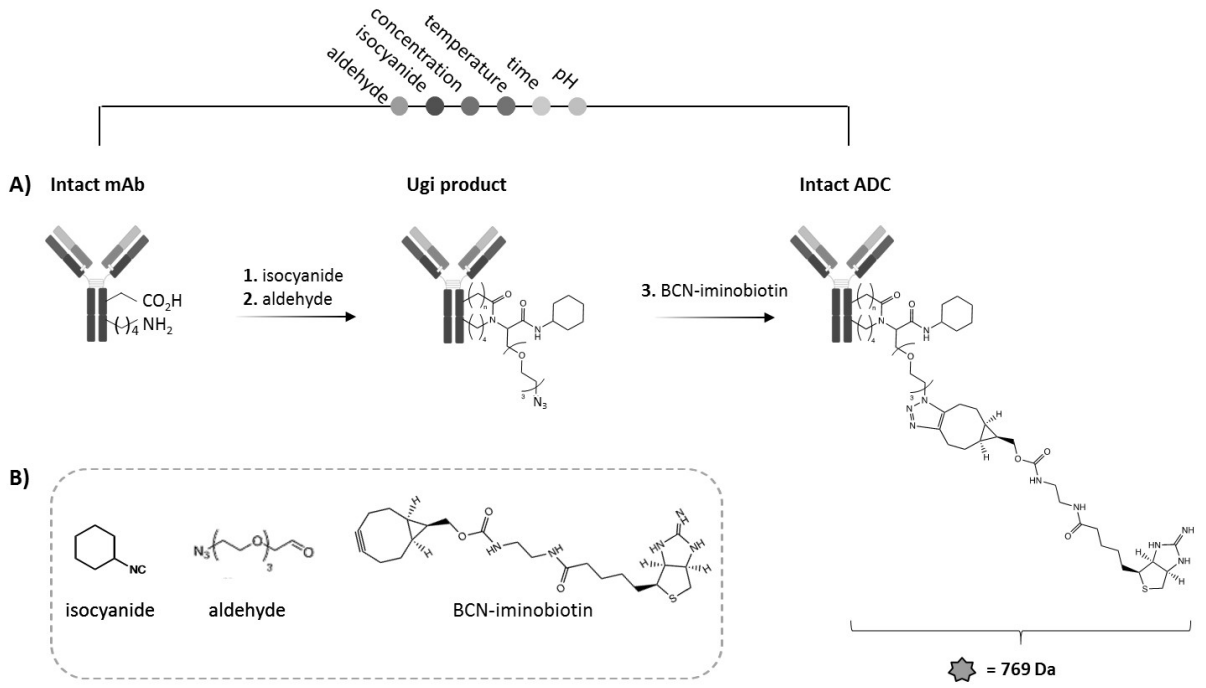


Figure 43: Schematic representation of conjugation of intact mAbs using Ugi multicomponent reaction. **A**) Ugi reaction resumed in two steps: first step consists in conjugating the mAb with azide-containing aldehyde and isocyanide through the side-chain amine and the carboxylate groups of two neighboring lysine and aspartate/glutamate. The second step consists in BCN-iminobiotin functionalization through SPAAC reaction. The parameters varied to optimize this conjugation reaction are depicted in colorful circles on top (pH, time… etc.). **B**) The reagents used during different steps of Ugi reaction.

The residues involved in the Ugi reaction can also partook in either intra-residue Ugi reaction on Passerini three components side reaction, where only the aspartate/glutamate reacts, preventing the chemoselective and regioselective multicomponent reaction and leading to different payload increment masses. Those conclusions were drawn from previous work performed within Chaubet's group and our team after characterization of Ugi conjugated trastuzumab by SEC-nMS and peptide mapping studies [32]. Therefore, the need to improve the specificity of mAb conjugation through Ugi reaction requires the optimization of a large number of parameters, in particular aldehyde and isocyanide choice, as well as several reactions conditions (temperature, time, pH… etc.), which generates a significant amount of samples to be monitored through SEC-nMS experiments.

3.2 SEC-nMS for the characterization of ADCs

First, we wanted to evaluate our previously developed SEC-nMS method on the BioAccord LC-MS system by analyzing a model conjugated trastuzumab. The same sample was injected through an Acquity I-Class LC coupled to a Synapt G2 HDMS instrument for comparison purpose. Both instruments were operated in their optimized conditions:

- **BioAccord**: 20 µg injected through a Maxpeak BEH 250 Å 1.7 µm, 4.6 x 150 mm column (Waters) using 150 mM ammonium acetate at pH 6.9 as a buffer. Cone voltage and desolvation temperature were both fixed at 180V and 300 °C, respectively. Source pressure was 1.4 mbar.
- **Synapt G2**: 20 µg injected through a Maxpeak BEH 250 Å 1.7 µm, 4.6 x 150 mm column (Waters) using 150 mM ammonium acetate at pH 6.9 as a buffer. Cone voltage and desolvation temperature were fixed at 180V and 450 °C, respectively. Source pressure was 6 mbar.

The SEC-UV signal in both cases shows a homogeneous main peak corresponding to monomeric ADC (95% and 97% of the signal for the BioAccord and the Synapt G2, respectively), with hallmarks of HMWS (5% and 3% of the signal for the BioAccord and the Synapt G2, respectively). The identity of the latter peak, which corresponded to the dimer, was confirmed on the Synapt G2 but not on the BioAccord due to limited m/z range. The native MS spectra of the main peak show similar MS signals in both cases, with identical DAR values, DLD and overall accurate mass measurement (~10 ppm) (**Figure 44**). Altogether, these results are in-line with what we obtained when comparing two mAb analyses on the two instruments, leading to the conclusion that our SEC-nMS method is adapted to high throughput ADCs characterization.

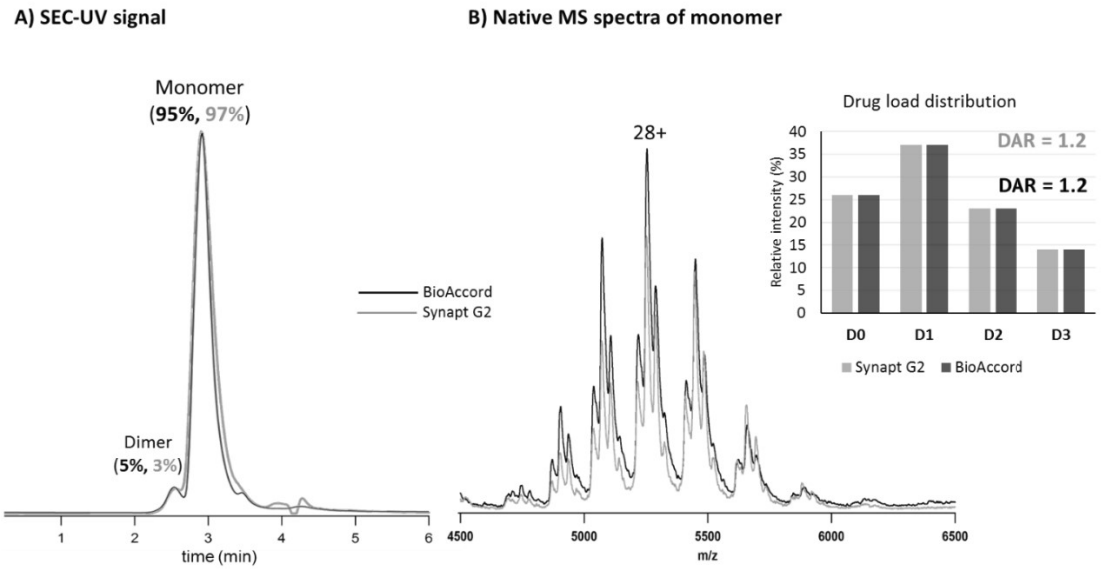


Figure 44: SEC-nMS analysis of a conjugated trastuzumab on the BioAccord LC-MS system compared to its analysis on the Synapt G2 using the optimized conditions (Chapter 1). **(A)** SEC-UV signal showing the main species corresponding to the ADC's monomer with hallmarks of dimer. **(B)** Native MS spectra of the ADC's monomer on both instruments centered on the charge state 28+ with the corresponding DLD and avDAR value that are comparable for both instruments. The signal of the BioAccord is shown in black while the signal of the Synapt G2 is blue.

3.3 High throughput SEC-nMS analyses of ADCs generated upon Ugi reaction

We next used SEC-nMS to screen multiple experimental conditions to achieve optimal Ugi reaction. Different parameters could influence the progress and selectivity of the reaction. Herein, we will show only two examples of our extensive series of experiments, in particular the investigation of pH variation and the use of different aldehydes.

a) pH optimization:

One of the main parameters that can control the accessibility of amino acid side chains due to the difference in their pKa values is the pH of their containing buffer. Thus, our collaborators in an effort to direct the selectivity of their Ugi reaction, they studied its influence on the outcome of the conjugation.

In case of changing pH upon Ugi reaction on trastuzumab, our high throughput SEC-nMS experiments allowed to perform precise mass measurement of the different DAR species (from D0 to D3 species), confirming the conjugation of one to three payloads and the presence of unconjugated trastuzumab. These results also allowed observing the effect of pH variation on the conjugation reaction. Shown in **Figure 45**, slight change was observed in the avDAR values and the DLD when varying the pH. At pH range of 6.5-7.0, a maximum avDAR value of 1.30-1.34 was reached with lower D0 quantities. Conversely, at pH lower than 6.5 and higher than 7.0, the avDAR was notably lower and the quantity of unconjugated mAb (D0) was higher. This study allowed our collaborators to select 6.5-7.0 as their optimal pH range for their further reaction optimization. Of note, the SEC-nMS analyses were performed in monoplicates in order to gain time and provide a fast feedback to our collaborators.

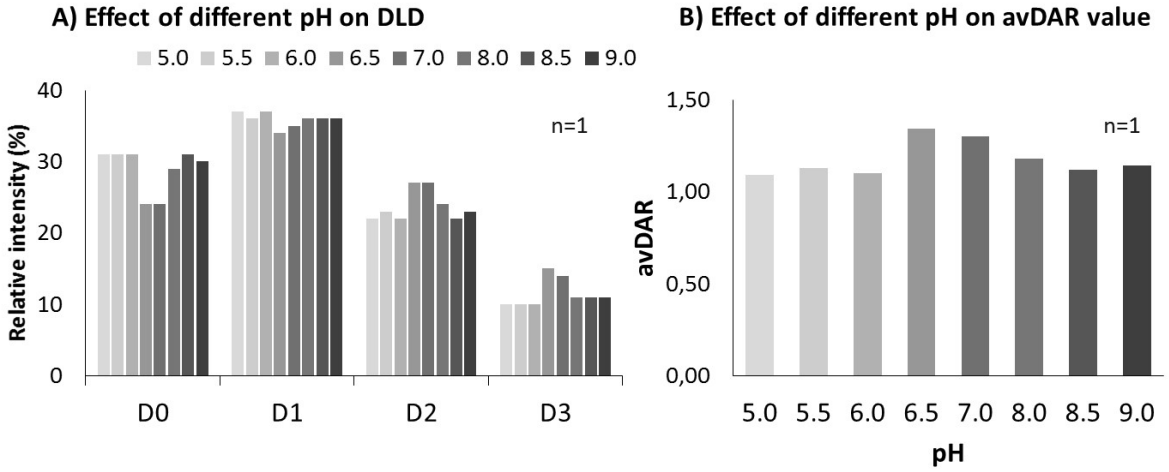


Figure 45: DLD and average DAR values calculated upon the analyses of 8 trastuzumab conjugates upon Ugi reaction using different pH. (**A**) DLD showing similar behavior where slight difference in D0-D3 relative intensities is observed at pH 6.5 and 7.0, as expected. (**B**) The avDAR values showing no significant change when the pH is between 5.0 and 6.0 and between 8.0 and 9.0, with higher values corresponding to pH 6.5 and 7.0.

b) Aldehyde optimization:

Next, our collaborators evaluated five different aldehydes that are important compounds for the success of the Ugi reaction. SEC-nMS analyses revealed significant variations in the avDAR value and D0 relative intensity as a function of the tested aldehyde (**Figure 46**). Indeed, using aldehydes 1, 2 and 5 led to the observation of additional species corresponding to D4 conjugates, with lower quantities of unconjugated trastuzumab (D0). In some cases, the unconjugated trastuzumab quantity was

significant such as for aldehyde 4, which highlights the lower efficiency of the reaction when using this more important aldehyde. On another note, the observation of highly conjugated species (such as D4) could be indicating the potential loss of selectivity. Thus, in this case the avDAR and the DLD were highly informative on the conjugation reaction efficiency and selectivity, results that could be supported by peptide mapping studies.

Overall, our high throughput automated SEC-nMS experiments allowed to test a large number of conditions (~20 samples/day) to improve the conjugation reactions and in this case, to select the optimal aldehyde that led to higher conjugation selectivity and efficiency.

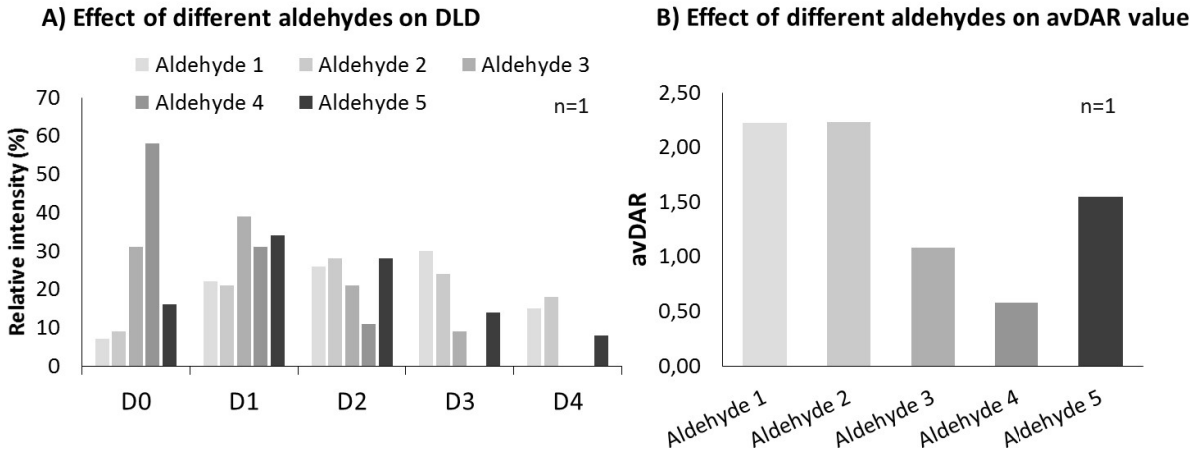


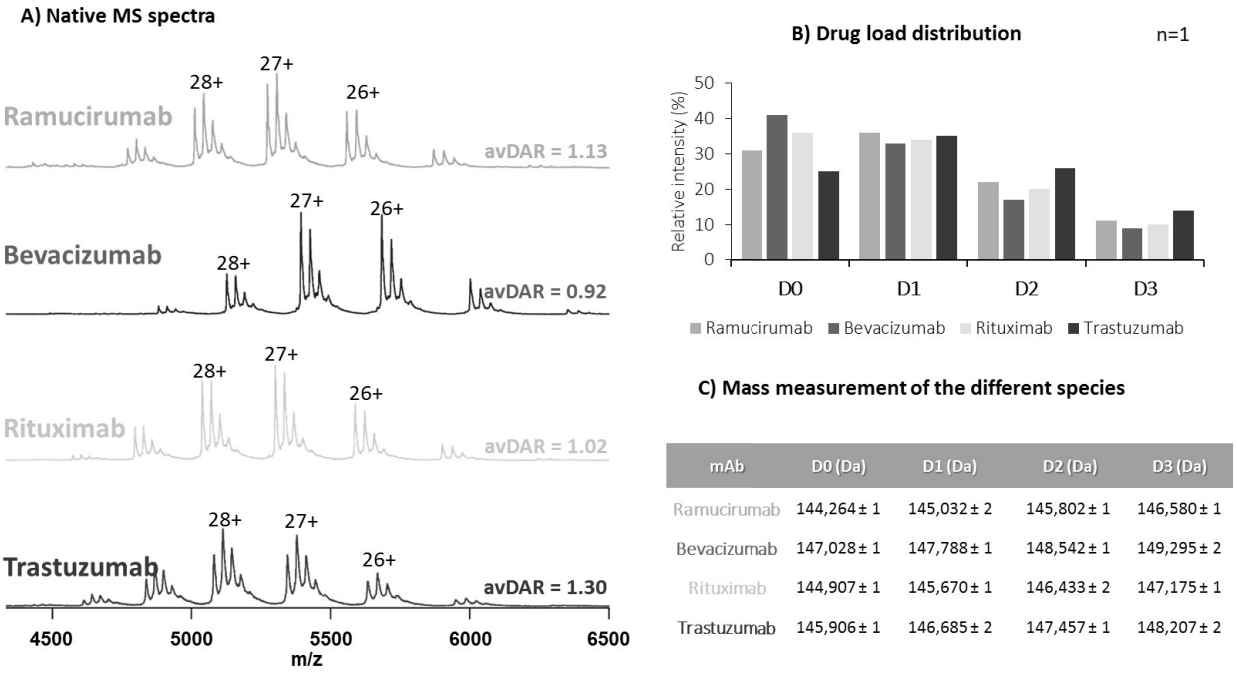
Figure 46: DLD and average DAR values calculated upon the analyses of 6 trastuzumab conjugates upon Ugi reaction using different aldehydes. (**A**) Different DLD observed while varying aldehydes with aldehyde 1, 2 and 6 showing until D4 species. (**B**) The avDAR values showing major differences upon aldehyde variation with aldehyde 1 and 2 giving higher avDAR values.

c) Versality of SEC-nMS for a variety of mAbs:

In the following study, we aimed to evaluate the optimal conditions upon other mAbs bio-conjugation namely bevacizumab (pI = 8.3), ramucirumab (pI = 9.1) and rituximab (pI = 9.4) (**Figure 47**). Our SEC-nMS results showed similar MS profiles centered on the charge state 27+ for all mAb-conjugates, with bevacizumab slightly shifted to higher *m/z* due to its higher molecular weight, as expected. Thanks to the mass accuracy of our instrument (~27 ppm for intact mAbs), we could also characterize the different DAR species varying from D0 to D3 conjugates. The mass increment observed between two species each time, corresponds to the added payload (+ 769 Da). Interestingly, trastuzumab (pI = 9.1) showed higher avDAR value of 1.30 with only 25% of unconjugated mAb (D0), compared to other mAb conjugates showing nearly similar avDAR values around ~1.00 and similar DLD with significant quantity of D0 species (31%, 40% and 37% for ramucirumab, bevacizumab and rituximab, respectively).

As shown in our previous batch analysis comparing different pH values, the choice of the pH was not crucial to achieve a better specificity and control of the Ugi reaction. However, antibodies with different pIs could match the selected pH (6.5-7 in this case) and could facilitate the incorporation of

the BCN-iminobiotin payload. For example, mAbs with higher pIs (trastuzumab, ramucirumab and rituximab) have more available amine groups that are reactive at the selected pH range. This could explain the observed higher avDAR values and different DLD in comparison to bevacizumab. However, the variations in avDAR and DLD could also arise from several factors, such as the stability, affinity and target binding of the chosen antibodies.



C) Mass measurement of the different species

mAb	D0 (Da)	D1 (Da)	D2 (Da)	D3 (Da)
Ramucirumab	144,264 ± 1	145,032 ± 2	145,802 ± 1	146,580 ± 1
Bevacizumab	147,028 ± 1	147,788 ± 1	148,542 ± 1	149,295 ± 2
Rituximab	144,907 ± 1	145,670 ± 1	146,433 ± 2	147,175 ± 1
Trastuzumab	145,906 ± 1	146,685 ± 2	147,457 ± 1	148,207 ± 2

Figure 47: SEC-nMS analysis of different ADCs resulting from mAbs conjugation upon Ugi reaction. **(A)** Native MS spectra of the different conjugates with their corresponding avDAR values (from top to bottom: ramucirumab, bevacizumab, rituximab and trastuzumab conjugates) showing similar distribution overall, with bevacizumab shifted to higher m/z. **(B)** Drug load distribution of each ADC showing different conjugation profiles for each mAb conjugate. **(C)** Masses of the identified species with standard deviation of masses obtained from different charge states.

Overall, upon obtaining an information on the homogeneity and stoichiometry of the analyzed conjugates, our implemented high throughput SEC-nMS method allowed monitoring the impact of Ugi reaction on different mAb products with CQAs assessment.

d) Automated data processing:

The large number of samples analyzed requires the use of automated data processing methods. For this purpose, we used our previously established method on UNIFI software (Waters). **Figure 48** illustrates the steps followed to perform automated data processing for trastuzumab conjugates with varied pH as an example.

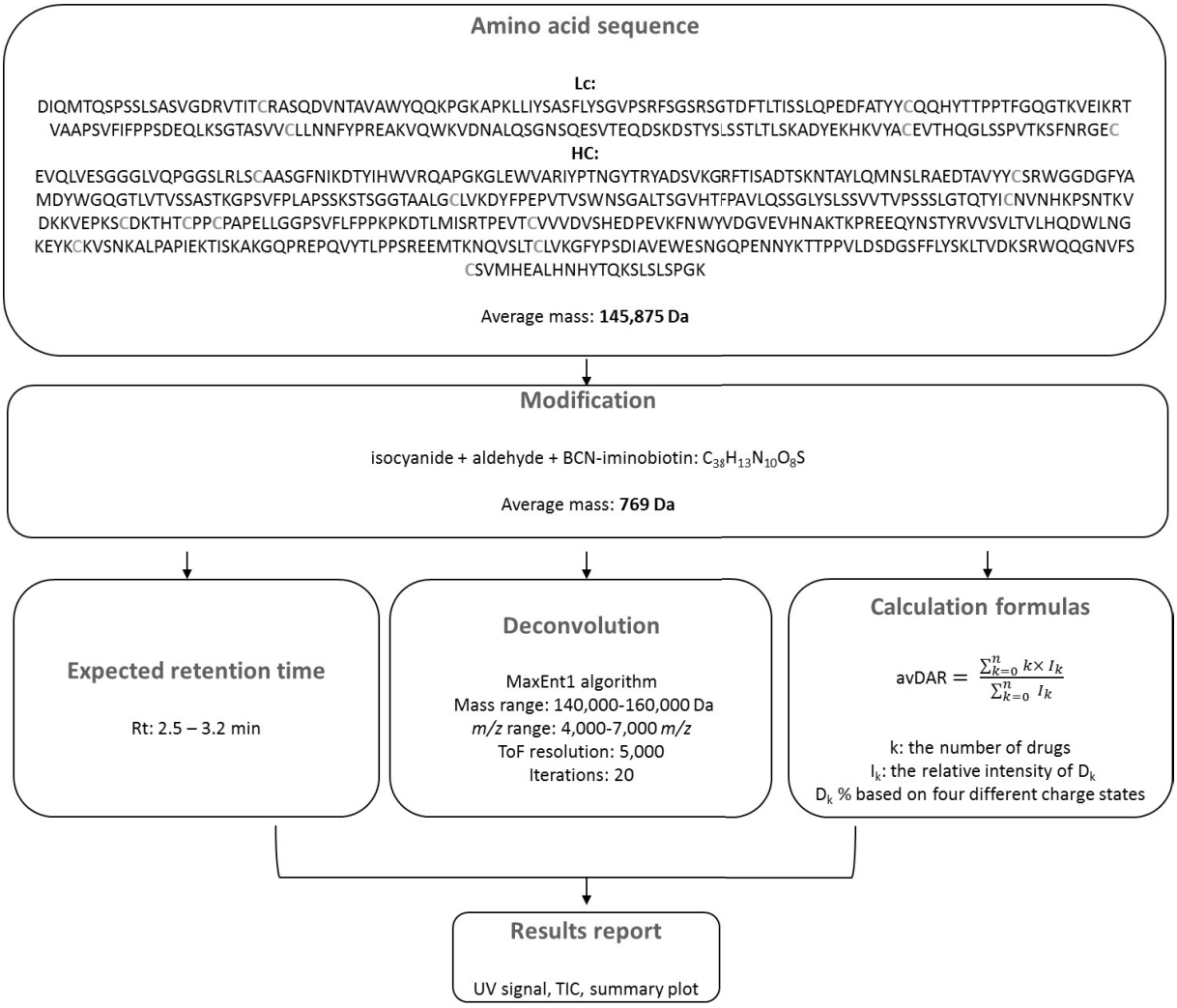


Figure 48: A schematic of the workflow followed for data processing of a large number of samples, including trastuzumab conjugates with varied pH. The first steps consist in providing information on the amino acid sequence and the modification. Next, the expected retention time, the parameters of deconvolution and the formula to calculate the avDAR are designated. Finally the information needed for the report are selected.

First, we entered the amino acid sequence corresponding to the intact mAb (ex: trastuzumab) with maintained disulfide bridge, which allowed to calculate a theoretical mass of 145,875 Da. An important step was to create and integrate the modification in our research, which includes the conjugation of a BCN-iminobiotin molecule after isocyanide and aldehyde additions (769 Da). Next, we chose to integrate the peaks between 2.5 and 3.2 min corresponding to the monomeric species of the conjugated antibody. This enabled us to quantify the monomeric species, which account for 95-98% for all samples with hallmarks of dimer (290 kDa, 2-5%). Then, this peak of interest is integrated to see the masses of the eluted species. Using the MaxEnt1 algorithm, we selected a mass range between 140,000 Da and 160,000 Da, targeting the *m/z* 4,000-7,000 region where the ADCs are observed. When the peak width model is set to ToF, we can modify the ToF resolution value; a value of 5,000 is sufficient to obtain an accurate mass measurement. To lower the data processing time, the number of iterations was set to 20. This corresponds to the number of times the algorithm repeats the mass calculation to

match the theoretical mass of the species. Finally, after adding the formulas that are essential for calculating avDAR and DLD, the method was applied to process up to 20 samples/day. **Figure 49** represents the type of report we obtained while screening eight different samples with different pH, showing the UV signal and the total ion chromatogram (TIC) as well as the summary plot with different avDAR values of each injection. This automated method and the obtained report allowed us to save an enormous amount of time (6 min run + less than 1 min for data processing for each sample). Therefore, we were able to provide our collaborators with a rapid response about the pH variations studies, which led to select the pH giving the highest avDAR value. However, when it comes to heterogeneous samples with spectra with higher background noise and less desolvated peaks, manual validation is necessary to ensure a correct data processing.

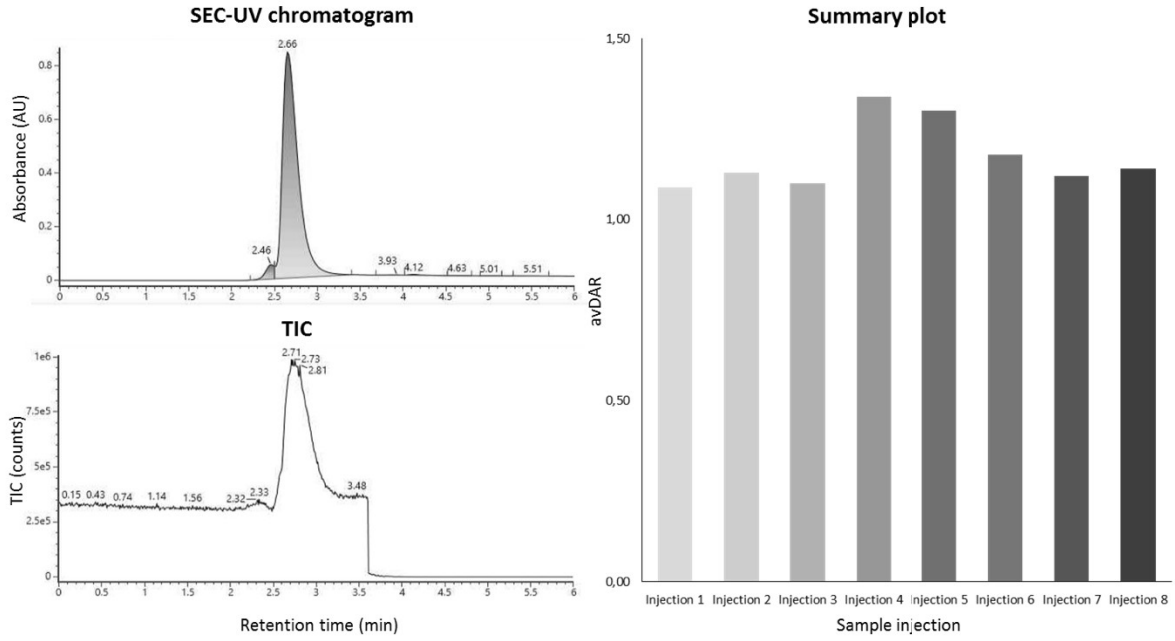


Figure 49: Overview of results obtained after analysis of 8 samples batch upon screening different pH values for trastuzumab Ugi conjugation. The first panel (top right) correspond to the SEC-UV chromatogram with integrated monomeric and dimeric peaks. The second panel (bottom right) is the total ion chromatogram. The last panel (left) is a summary plot of all the injections with the calculated avDAR values.

e) Peptide mapping supporting our SEC-nMS results:

To support our SEC-nMS results that reveal different avDAR and DLD values when varying the different parameters, peptide mapping was an important step to assess the specificity of the conjugation. Peptide mapping is conventionally used to obtain sequence information and localize PTMs along with drug conjugation sites. When performed on one of the optimal Ugi trastuzumab conjugates (Aldehyde 4 pH 7.4), after digestion with trypsin, we first determined a sequence coverage of 100% for the light chain (Lc) and 86% for the heavy chain (HC). The MS/MS data revealed a modification of +769 Da, corresponding to Ugi conjugation, on the Lc at position D1 and on the heavy chain at position E1 (**Figure**

50). These results are in agreement with trastuzumab bearing one Ugi payload on each chain (leading to D4 species as observed in SEC-nMS), which are in-line with our SEC-nMS data. Those data supported by peptide mapping studies, allowed us to screen a large number of conditions to achieve site-selective trastuzumab Ugi bio-conjugation.

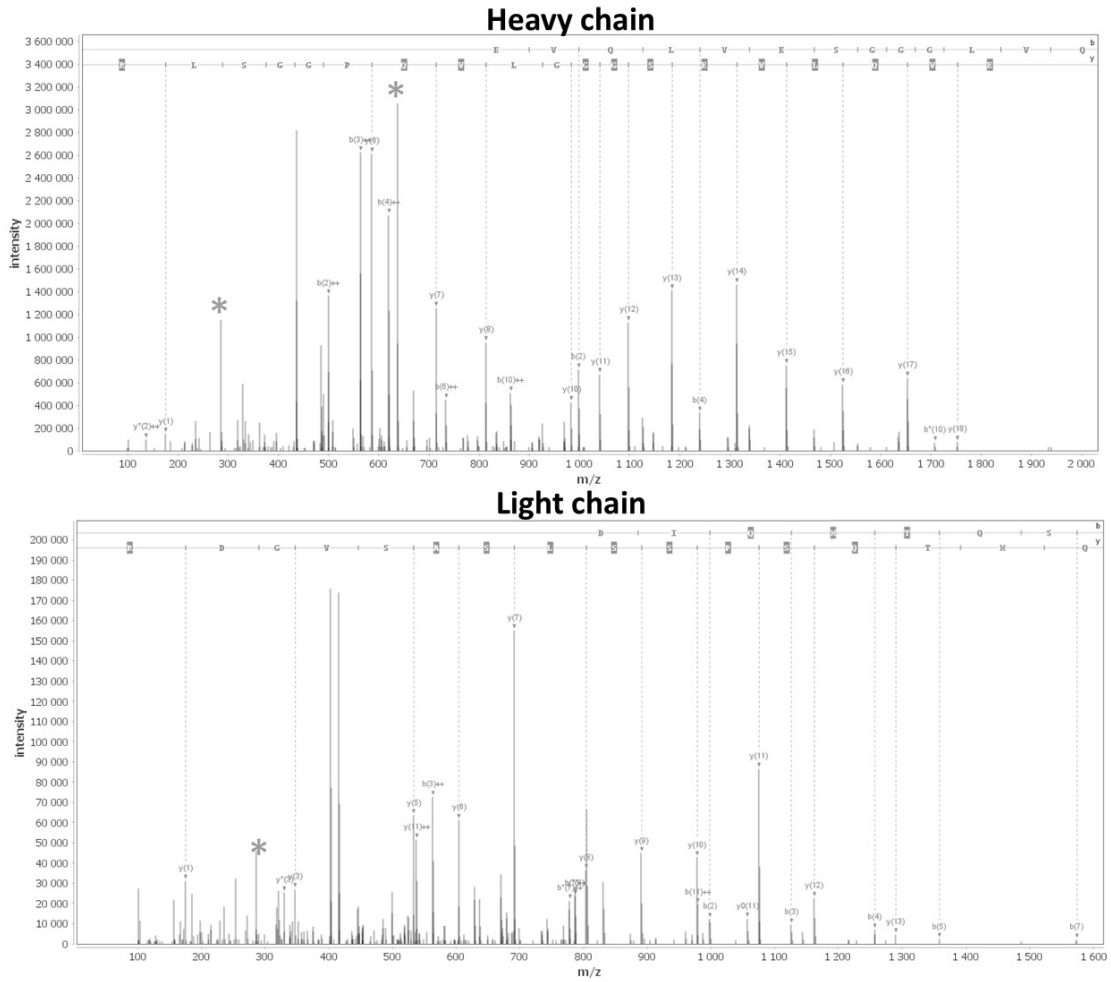


Figure 50: MS/MS spectra of trastuzumab conjugate upon peptide mapping analysis. Fragments ions corresponding to the conjugated sites are depicted in orange star confirming one conjugation on the light chain and one on the heavy chain.

4. Characterization of AcDCs upon Ugi reaction using SEC-nMS

Beyond screening the optimal conditions for mAb's Ugi bio-conjugation, along with Chaubet's group we aimed at investigating Ugi reaction on smaller protein systems, in particular on Anticalins (Collaboration with Chaubet's group) [295-296].

4.1 Anticalin proteins for immuno-oncology

Anticalin proteins are small entities derived from natural lipocalins that are abundant in human plasma proteins. Thanks to their small size, their compact fold, their targeted specificity and their low production cost they emerged as promising alternatives to antibodies in the field of oncology and diagnostics [295-296]. They thus represent an interesting system for specific Ugi conjugation in order to

enhance their specificity and their therapeutic efficiency. In this context, along with our collaborators, we aimed at evaluating Ugi reaction for the D11vs surrogate Anticalin (a variant with higher affinity toward the CD98hc from mouse, targeting medical application in humans)[297-298].

4.2 Anticalin and Anticalin-drug conjugates (AcDC) analysis using SEC-nMS

First, we performed SEC-nMS experiments on the unconjugated form of the D11vs Anticalin variant. The sample was buffer exchanged into 150 mM $AcONH_4$ pH 6.9 through an Acquity BEH SEC column (200 Å, 1.7 µm, 2.1 mm X 150 mm), cone voltage and desolvation temperature were set at 60V and 100 °C, respectively. **Figure 51A** shows the SEC-UV signal of the Anticalin that exhibited one single peak eluting at ~3.45 min (As= 1.2, FWHM = 0.12 min). The corresponding native MS spectra (**Figure 51B**) revealed the presence of three different forms: denatured Anticalin, dimer of Anticalin (42,601.8 ± 0.5 Da) and the main species, which is the monomeric Anticalin with a mass of 21,300.7 ± 0.1 Da, in line with Anticalin with one disulfide bridge. Additionally, a zoom on the monomeric species revealed the presence of 2 additional proteoforms of Anticalin that may correspond to N-ter PyroQ glutamination (-17 Da) and protein oxidation (+16 Da). However, the UV signal does not show any peak that could correspond to dimeric Anticalin, which means that the observed dimer in the MS spectra could correspond to a non-specific dimer formed during the electrospray process.

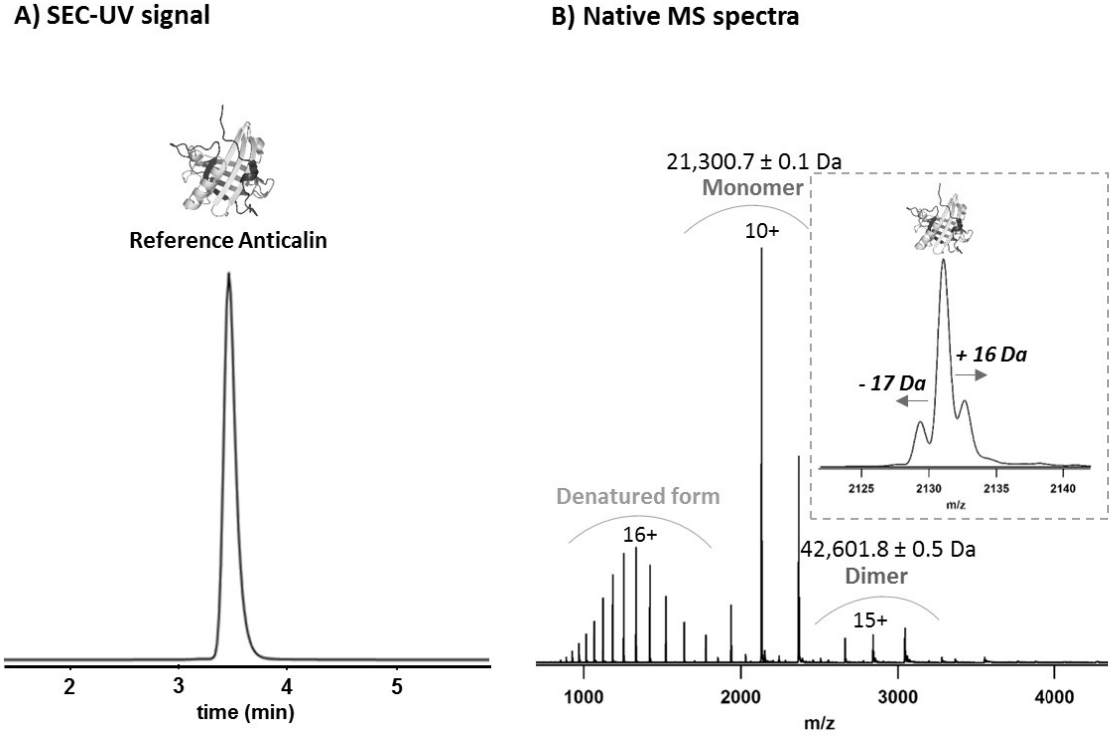


Figure 51: SEC-nMS analysis of reference Anticalin D11vs variant. **(A)** SEC-UV signal showing one homogeneous species at 3.45 min **(B**) Native MS spectra of the identified peak, showing three different regions namely denatured form of Anticalin, monomer and dimer forms. On the grey square, a zoom on the charge state 10+ of the monomer region showing the triplet profile of the peaks due to presence of modifications on the protein.

Overall, SEC-nMS analysis of Anticalin allowed determining the oligomeric state of this protein, along with the assessment of minor modifications (Oxidation and PyroQ glutamination) through accurate mass measurement.

Next, our versatile SEC-nMS method was used to investigate Ugi site-specific conjugation on the Anticalin D11vs, along with other Anticalin D11vs mutants, namely mutated (K46R)-D11vs, truncated (Q1-D6)-D11vs, and the double mutant (Q1-D6)-(k46R)-D11vs protein. **Table 6** shows the first 43 or 48 amino acids of the proteins' sequences, which is the region where the mutation and truncation occurred.

Table 6: The first 43-48 amino acids involved in the area where the truncation/mutation of Anticalin variants occurred. D11vs variant is the reference protein. (K46R)-D11vs is the mutant with modification of K46 with an arginine. (Q1-D6)-D11vs is the truncated form lacking the first six amino acid at the N-terminal. (Q1-D6)-(K46R)-D11vs is the double mutant with K46 modification and Q1-D6 truncation. The truncated or mutated amino acids are highlighted in orange. The theoritical masses were calculated based on the amino acid seqeuces that are presented in the experimental section of this manuscript.

Name	First 43-48 amino acids of the sequence	Theoritical mass (Da)
D11vs	------	21,302.12
(K46R)-D11vs	NLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDRDP… 43	21,330.14
(Q1-D6)-D11vs	------	20,782.66
(Q1-D6)-(K46R)-D11vs	NLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDKDP… 43	20,810.67
	QDSTSDLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDKDP … 48	
	QDSTSDLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDRDP … 48	

The aim of this study is first to see if Anticalins partook efficiently in the Ugi multicomponent reaction, therefore a first step of this project was to investigate the addition of an increment mass of +308 Da corresponding to an isocyanide with an aldehyde (**Figure 52**). The second objective is to investigate if the mutation/truncation modify the Ugi conjugation.

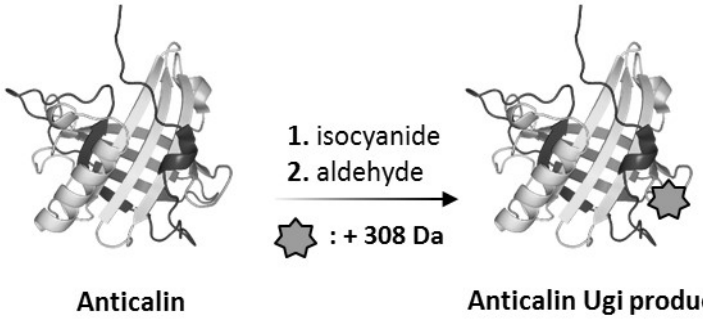


Figure 52: Schematic representation of Ugi reaction on Anticalin proteins, which includes the conjugation of reference Anticalins to one isocyanide and aldehyde (similarly used for trastuzumab) inducing the addition of an increment mass of 308 Da. A pink star depicts the addition of one payload (isocyanide + aldehyde).

Figure 53 shows the results of SEC-nMS analysis of the conjugated reference Anticalin and its mutants upon Ugi reaction. The SEC-UV chromatogram reveals one major main peak for all proteins eluting around ~3.42-3.50 min. Of note, a slight peak tailing is observed due to the increased hydrophobicity of the conjugated proteins after payload incorporation, which increases non-specific interactions between the analyte and the stationary phase of the column.

The identity of the main peaks revealed thanks to the coupling with nMS, allowed to characterize each species:

- For reference Anticalin D11vs, D1 species was highly abundant with presence of D0 and D2 species with an avDAR of 1.08,
- For the truncated (Q1-D6)-D11vs, the main species was also D1 with presence of D0 and D2 in lower quantity with an avDAR of 1.06,
- For the (K46R)-D11vs mutant, the D0 was the highest species followed by D1 and D2 species which exhibited a lower avDAR of 0.49,
- In addition, for the double mutant (Q1-D6)-(k46R)-D11vs, only two species were observed, mainly D0, which was the most abundant, and D1 with lower abundance and with the lowest avDAR value of 0.16.

The (Q1-D6)-D11vs Anticalin was truncated at the first six amino acids where Passerini reaction tends to occur (in aspartic acid at position 2 (D2) according to peptide mapping results, **Table 7**). However, the absence of this reactive site did not boost Anticalin Ugi conjugation and similar DLD was observed compared to the reference Anticalin. This could mean that the Passerini reaction is negligible or occurs at a very slow rate. On the other hand, the (K46R)-D11vs mutant and the double mutant (Q1-D6)-(K46R)-D11vs, both lack the key lysine residue at position 46 (K46) which is involved in Ugi reaction (results obtained by peptide mapping studies, **Table 7**). The absence of this site affected directly the conjugation efficiency, where lower avDAR values (0.49 and 0.16, respectively) were observed.

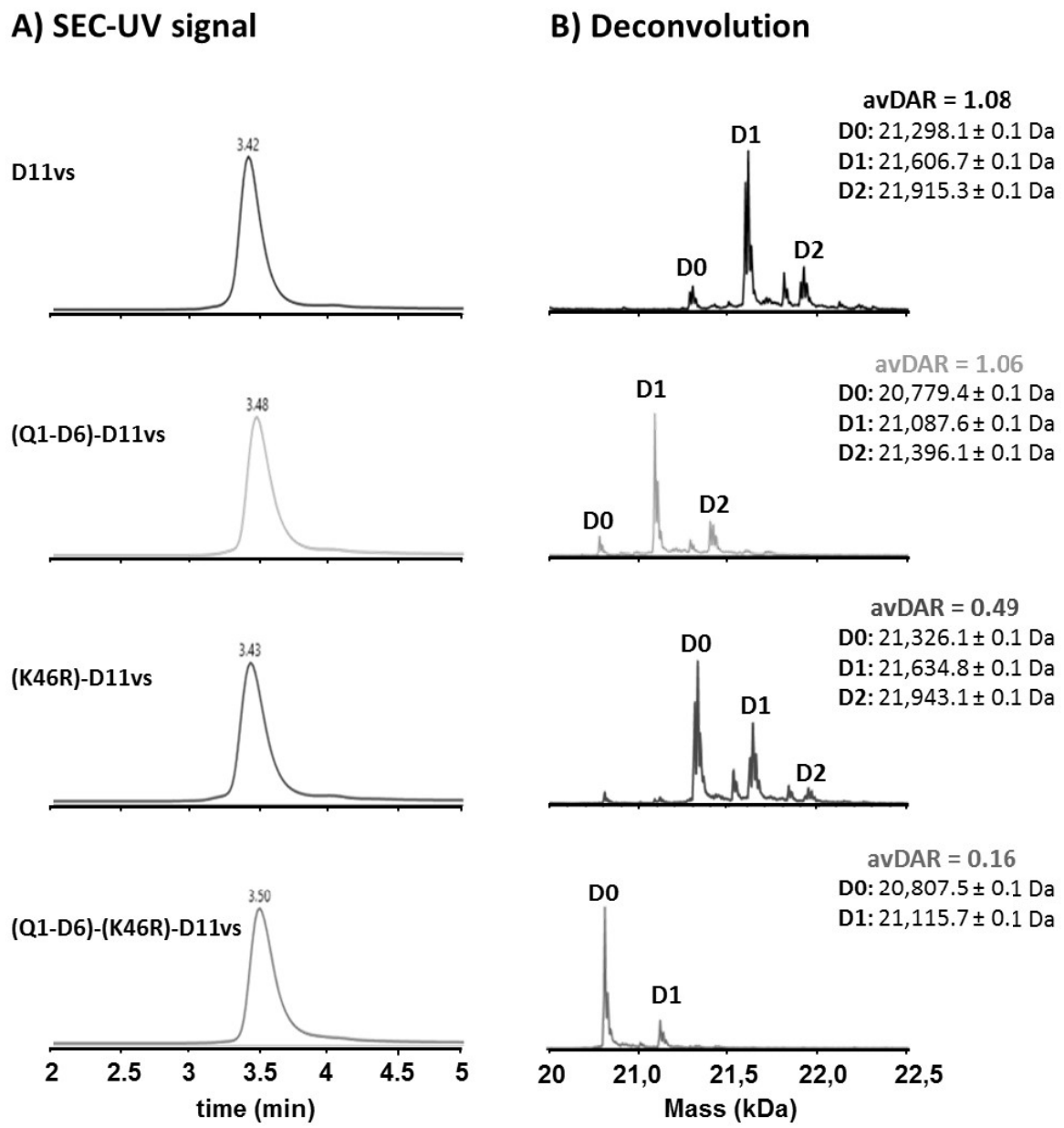


Figure 53: SEC-nMS analysis of Anticalin-drug conjugates upon Ugi reaction performed on reference D11vs Anticalin and its mutants. **(A)** SEC-UV signal of unconjugated Anticalins with their experimental masses (black) and their conjugated counterparts (blue shades). **(B)** Deconvolution of the identified species from native MS spectra of unconjugated Anticalins (black) and the conjugated ones (blue shades) with the different DAR species annotated and the calculated avDAR value.

In conclusion, our SEC-nMS results supported with peptide mapping data summarized in **Table 7**, enabled us first to prove the occurring of the Ugi multicomponent reaction on smaller proteins in comparison to mAbs; in particular Anticalin proteins. Moreover, our experiments performed on different Anticalin variants, could highlight the impact of mutations/truncation on the specificity and efficiency of the Ugi reaction through avDAR and DLD assessment. These analyses, together with others carried out on different batches to screen various Ugi reaction parameters, have enabled us to select the optimal conditions for a successful Ugi reaction. The results are summarized in a paper currently in preparation.

Table 7: Peptide mapping results showing the sequence coverage obtained after MS/MS fragmentation of trypsin digestion Anticalins, with the modified amino acids upon Ugi, Passerini or both reactions. The conjugation that was not observed is annotated as N.O.

Protein	Sequence coverage	AA	Ugi	Passerini
D11vs	100%	D2, K46, E44	Yes	Yes
(Q1-D6)-D11vs	97%	K46, E44	Yes	N.O.
(K46R)-D11vs	91%	D2	Yes	N.O.
(Q1-D6)-(K46R)-D11vs	100%	-	N.O.	Yes

5. Conclusions

This chapter highlights the important role of SEC-nMS method to characterize newly developed ADCs and PDCs. During the screening process of several conjugation parameters for Ugi reaction on antibodies, and the applicability of the optimal parameters on different Anticalin mutants, SEC-nMS was a powerful tool to give precise mass measurement and CQAs assessment promptly. These results suggest that this workflow should be implemented in bio-conjugation and quality control laboratories in academia and biopharmaceutical companies, to monitor newly developed conjugates.

6. Communications/publications

Benazza R., Hernandez-Alba O. and Cianférani S. **New MS-based approaches for the characterization of PDCs**. *Oral presentation*. *Journées Pratiques Françaises des Sciences Analytiques (JPFSA).* April **2022**, Marrakech, Morocco.

Koutsopetras I., Vaur V., Benazza R., Diemer H., Sornay C., Ersoy Y., Rochet L., Longo C., Hernandez-Alba O., Erb S., Detappe A., Skerra A., Wagner A., Cianférani S. and Chaubet G. **Site-Selective Protein Conjugation by a Multicomponent Ugi Reaction.** *ChemEurJ*, **2023**.

Chapter 3: Benefits of SEC-nMS for the characterization of newly developed bispecific antibodies (bsAbs)

1. Analytical context

Beyond the interest in ADCs as a new class for cancer therapy, bispecific antibodies (bsAbs) are emerging as promising antibodies formats to fight against cancer [4, 88]. Owing to their capacity to bind both to tumor and to T-cell antigen sites of antibodies, they represent an ideal candidate for the rich clinical antibody pipeline. In fact, there are more than 160 bsAbs and multispecific mAbs currently in clinical trials, with nearly 100 combination of antigen-binding moieties [87, 299]. This plethora of complex products requires robust and powerful analytical tools to monitor their development and LC-MS was widely used it this context [125]. Albada's group (Wageningen University, Netherlands) which is part of the ITN TACT program, aimed at generating synthetic bispecifics upon a new bio-orthogonal click chemistry.

2. Objectives

The aim of this chapter is to evaluate the capacity of the BioAccord LC-MS platform for the characterization of more complex and larger bsAbs. Additionally, we aim to provide our collaborators with a direct snapshot on the synthetized bispecifics through SEC-nMS analysis in order to better optimize the click chemistry reaction.

3. Results

3.1 Bio-conjugation strategy

The studied bispecifics were obtained via a bio-orthogonal tyrosine-based click chemistry reaction (**Figure 54**) aiming at conjugating two Fab fragments (trastuzumab anti-HER2: 48,332 Da or muromonab anti-CD3 referred to as OKT3: 48,141 Da) to a deglycosylated anti-HER2 trastuzumab. The final expected product are 2:2 HER2:HER2 (*will be referred to Trastu:Trastu bsAb in this chapter*) or 2:2 HER2:CD3 (*will be referred to Trastu-OKT3 bsAb in this chapter*) bsAb constructs, depending on the used Fab fragment. Briefly, the trastuzumab is deglycosylated using peptide-N-glycosidase F (PNGase F, 37°C, 16h, pH 7.4) to remove N-glycans. The deglycosylated mAb is then functionalized on each Fc domain with one trans-cyclooctene (TCO) by subsequent treatment with mushroom tyrosinase (mTyr), inducing the addition of a theoretical increment mass of 2x 535 Da (4°C, 16h, pH 5.5). This handle will allow the attachment of Fab fragments to the intact antibody through bio-orthogonal click chemistry.

Of note, the Fab fragments were prior re-bridged using a pyridazinedione (PD) construct [82] to provide tetrazine handles to enable linkage with TCO handles.

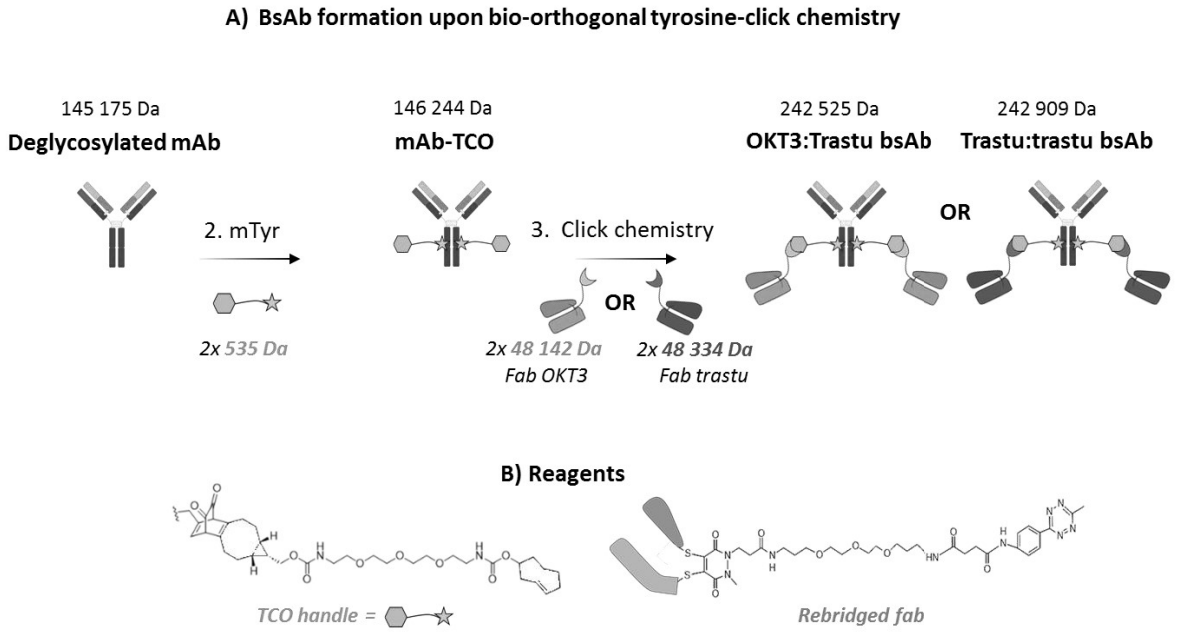


Figure 54: Schematic representation of the development of bispecific antibodies (bsAb) through bio-orthogonal tyrosine-based click chemistry. (**A**) BsAb formation reaction resumed in three steps: **1.** deglycosylation of intact mAb using peptide-N-glycosidase F (PNGase F), **2**. Functionalization of the Fc part with trans-cyclooctene (TCO) handles using mushroom tyrosine (mTyr) and **3.** The attachment of the TCO-functionalized mAb to Fab fragments, prior rebriged using pyridazinediones (PD). Fab fragments used are either from trastuzumab or OKT3 antibodies. The theoretical masses of the used TCO and Fab fragments, along with the theoretical masses of the final products are represented on the scheme. (**B**) Structures of the reagents used for bsAb formation namely TCO and rebriged Fab fragments.

3.2 SEC-nMS characterization of bispecific antibodies (bsAbs)

Initially, we tested our previously developed SEC-nMS method on the BioAccord platform upon characterization of the Trastu-OKT3 bsAb as an example, in comparison to the reference deglycosylated trastuzumab. The samples were analyzed using the Maxpeak SEC column (BEH 250 Å 2.5 µm, 4.6 x 150 mm column (Waters) using 150 mM ammonium acetate at pH 6.9 as a buffer. Cone voltage and desolvation temperature were both fixed at 180V and 300 °C, respectively. Source pressure was 1.4 mbar. As shown in **Figure 55**, the UV signal of the deglycosylated mAb exhibited one single symmetric peak at ~4.33 min with (Main peak depicted as M, 100%, FWHM= 0.110 and As= 1.13). In contrast, the bsAb sample exhibited several peaks observed in the UV profile, with the main peak (peak M) eluting slightly earlier at ~3.95 min and considerably larger in comparison to the unconjugated sample (M, 78% FWHM = 0.271 min, As =2.56). This major peak presented a tailing at ~4.12, thus we could observe two regions (peak 2 and peak 3) probably due to the co-elution of

different bsAb constructs (i.e. 2:2 or 2:1 Trastu-OKT3 bsAb, expectedly). In addition, several LMWS (peak 4, ~5.43 min) and HMWS (peak 1, ~3.62) were observed in the bsAb UV profile.

Through direct nMS coupling, we were able to unambiguously reveal the mass of the deglycosylated trastuzumab (145,175 ± 1 Da, 1 ppm). Whereas for the bsAb sample, the identification was more challenging. First, LWMS indicated presence of unbound OKT3 Fab fragment (48,143 ± 1 Da, 12 ppm). Then, the major peak (peak 2) and the peak tailing (peak 3) corresponded to 2:2 (242,530 ± 8 Da, 21 ppm) and 1:2 (194,386 ± 4 Da, 18 ppm) bsAb constructs, respectively. However, the bsAbs isotopic distribution was not fully observed due to the limited *m/z* range of the BioAccord (until 7,000 *m/z*). This hampered the precise mass measurement of those species due to consideration of only few charge states. Moreover, the HMWS observed in the UV signal were not identified in the MS because they probably ionize at higher *m/z* range (>7,000 *m/z*), or because they do not ionize well compared to the other species. Therefore, the results shown through bsAb analysis on the BioAccord platform reveal the limitations of this instrument to characterize larger proteins (>150 kDa) with the SEC-nMS approach. Those limitations could be tackled upon expansion of the *m/z* range to broaden the BioaccoAd's scope of applications in the future.

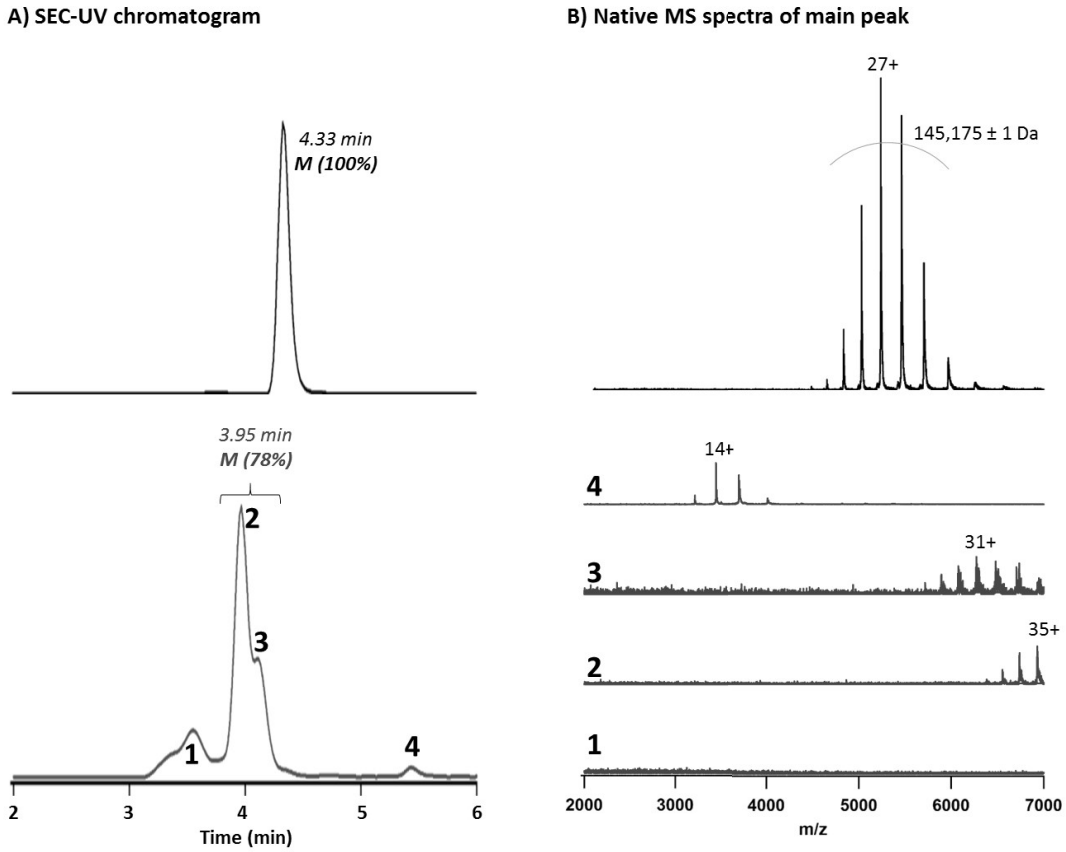


Figure 55: SEC-nMS analysis of deglycosylated trastuzumab (**black**) and Trastu-OKT3 bsAb (**blue**) on the BioAccord LC-MS system. (**A**) SEC-UV chromatogram showing one unique species for deglycosylated trastuzumab while revealing different species for the treated bsAb. (**B**) Native MS spectra of the main peak (peak M) for the deglycosylated trastuzumab (black) and the different species named 1, 2, 3 and 4 observed in the bsAb sample.

To get more informative mass measurement on the produced bsAbs, we next performed the same analysis on the Synapt G2 where the *m/z* range is extended to 10,000 m/z. The deglycosylated trastuzumab shows a homogeneous sample with a single peak at ~4.33 min. The TCO functionalized trastuzumab exhibits two more peaks at ~3.52 min (7%) and ~3.77 min (22%) in addition to a major peak (71%) at ~4.35 min. Interestingly, the UV signal of the formed bispecifics exhibited at least 4 peaks each. For the Trastu-OKT3 bsAb, the major peak observed at ~3.95 min (54%) overlapped with a peak tailing at ~4.12 min (32%). Similarly, the Trastu-trastu bsAb showed one major peak at ~3.97 min (54%) that presented a peak tailing at ~4.12 min (24%). Although these two overlapped species are not baselined resolved from the main species, a relative quantification was performed upon a Gaussian fitting. Additionally, both samples exhibited peaks corresponding to LMWS at ~5.43 min (6% and 3% for Trastu-OKT3 and Trastu-Trastu bsAbs, respectively). Finally, HMWS were observed around ~3.50 min (8% and 20% for Trastu-OKT3 and Trastu-Trastu bsAbs, respectively). The relative quantification was based on the areas of the chromatographic peaks (**Figure 56**). Native MS revealed the identity of each chromatographic peak (**Figure 56**). As expected, the main peak of the deglycosylated trastuzumab sample corresponded to the monomer (145,175 ± 1 Da, 1 ppm). For the TCO functionalized trastuzumab, a mass of 146,246 ± 1 Da (12 ppm) was observed for the main peak, in-line with the attachment of 2 TCO handles (+ 2,148 Da). Trastuzumab-TCO (Trastu-TCO) sample showed additional species with masses that corresponded to $(Trastu\text{-}TCO)_2$ dimer (3.77 min, 292,495 ± 8 Da, 23 ppm). However, the species eluting at 3.52 min could not be identified in the MS probably due to insufficient ionization. The major peaks in the conjugated samples (Trastu-OKT3 and Trastu-trastu bsAb) highlighted the formation of 2:2 bsAbs mainly corroborating the conjugation of two Fab fragments to one trastuzumab (242,530 ± 8 Da, 21 ppm and 242,889 ± 4 Da, 20 ppm for trastuzumab-OKT3 and Trastu-Trastu bsAbs, respectively). Additionally, masses corresponding to 1:2 bsAb formats were observed (**Table 8**). Similarly to Trastu-TCO, the HMWS in this case were not observed in the nMS spectra; the latter were not observed on the BioAccord either due to the limited *m/z* range. **Table 8** summarizes the masses of all the identified species.

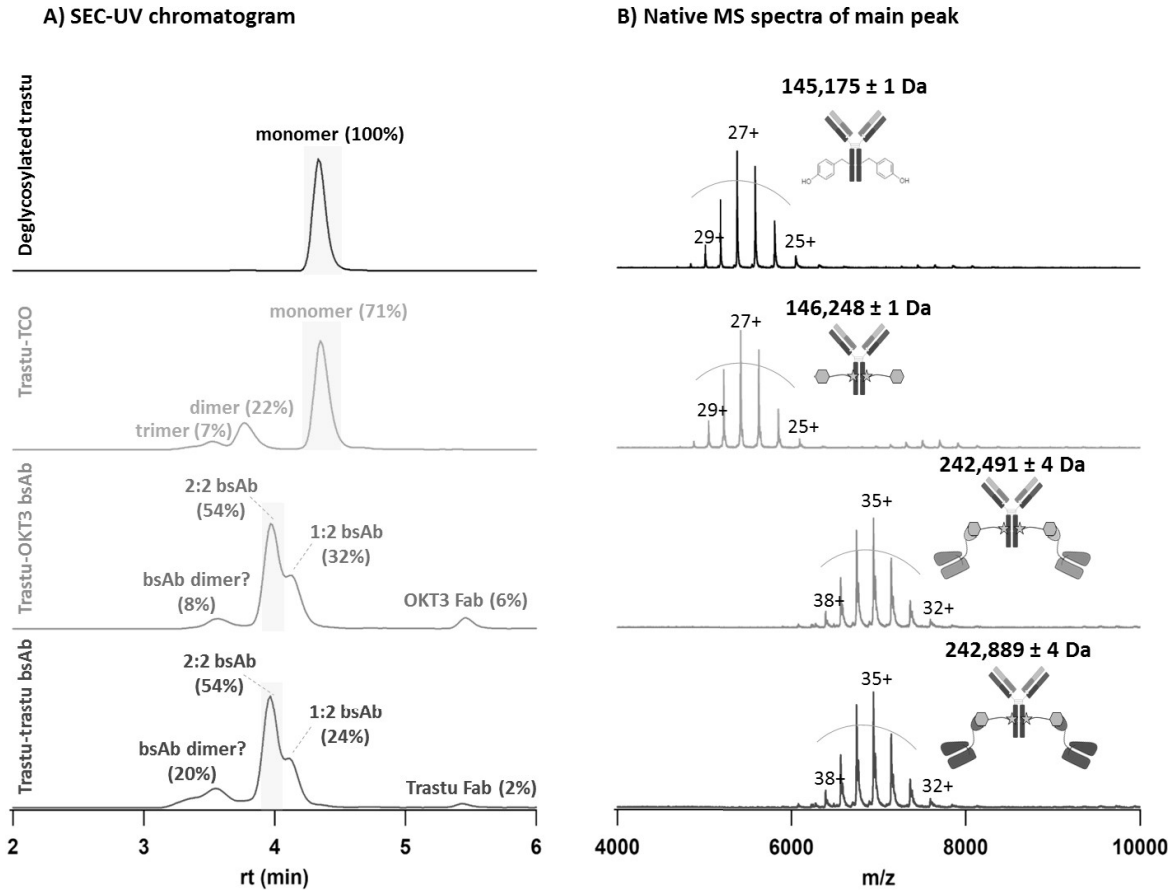


Figure 56: SEC-nMS analysis of trastuzumab samples obtained upon bispecific formation via bio-orthogonal tyrosine-based click chemistry. (**A**) SEC-UV chromatogram of deglycosylated trastuzumab (**black**), Trastu-TCO (**light blue**), Trastu-OKT3 bsAb (**green**) and Trastu-Trastu bsAb (**dark blue**). Relative quantification of each species is performed upon the integration of chromatographic peak areas. (**B**) Native MS spectra of the major peak from each sample namely, monomer of deglycosylated trastuzumab (**black**), monomer of Trastu-TCO (**light blue**), bsAb species of Trastu-OKT3 (**green**) and bsAb species of Trastu-Trastu (**dark blue**). The experimental masses are provided with a standard deviation obtained from at least four different charge states. The masses of other minor species are summarized in **Table 8**.

Beyond the accurate mass measurement of the different species, SEC-nMS results confirm the formation of bispecifics upon the bio-orthogonal tyrosine-based click chemistry reaction developed by Albada's group. The analysis of size variants highlights the co-existence of different bispecific formats ranging from 2:2 to 1:2 forms, paving the way to further optimization of the reaction conditions to obtain more homogeneous bsAb products.

Table 8: Masses of the identified species upon SEC-nMS analysis of deglycosylated Trastu, Trastu-TCO, Trastu-OKT3 and Trastu-trastu bsAbs, showing the peak area and retention time of each species, along with the experimental masses compared to theoretical ones.

Sample	Species	Area (%)	Retention time (min)	Experimental mass (Da)	Theoretical mass (Da)	Mass accuracy (ppm)
Deglycosylated Trastu	Monomer	100	4.33	145,175 ± 1	145,175	1
Trastu-TCO	N.A	7	3.52	N.A	/	/
	Dimer	22	3.77	292,495 ± 8	292,488	23
	Monomer	71	4.35	146,246 ± 1	146,244	12
Trastu-OKT3 bsAb	N.A	8	3.62	N.A	/	/
	1:2 bsAb	54	3.95	242,530 ± 8	242,525	21
	1:1 bsAb	32	4.12	194,386 ± 4	194,383	18
	OKT3 Fab	6	5.43	48,143 ± 1	48,142	12
Trastu-trastu bsAb	N.A	20	3.53	N.A	/	/
	1:2 bsAb	54	3.97	242,889 ± 4	242,909	20
	1:1 bsAb	24	4.12	194,586 ± 2	194,575	18
	Trastu Fab	2	5.43	48,334 ± 1	48,334	8

3.3 Mass photometry as additional tool

To determine the stoichiometry of bispecific binding, i.e. 1:2 or 2:2 bsAb constructs, mass photometry (MP) [133-134] could be used in complementary with SEC-nMS experiments to give a snapshot on the reaction products through direct mass measurement, in native conditions. MP is based on interferometric scattering microscopy (iSCAT); it measures the interference between the light scattered by single particles and the light reflected by the measurement surface [196]. The variation in reflectivity of proteins corresponds proportionally to their molecular mass, thanks to their optical properties and density uniformity. Consequently, the signal measured is directly correlated with the protein's molecular mass. The benefits of this technique were highlighted in literature upon characterization of several proteins, in particular antibodies, antibody-antigen and antibody-receptor interactions [132, 134, 196-197]. The increased interest in this technique is mainly due to using low sample quantities (100 pM – 100 nM) and analyzing samples in their storage/purification solution without

prior buffer exchange. Moreover, a key point which is interesting in our case, is the wide mass range that MP offers (up to 5 MDa), which could enable us to determine the identity of HMWS previously unrevealed by nMS [300]. The MP analyses provided different profiles for each sample (**Figure 57)**, similarly to results reported earlier from SEC-nMS experiments. The deglycosylated trastuzumab revealed one homogeneous distribution (100% of the sample counts) with a mass of 146 ± 14 kDa, as expected. Upon addition of TCO handles, three distinct distributions are observed, namely, 152 ± 21 kDa corresponding to TCO functionalized trastuzumab, 306 ± 29 kDa which is in-line with the Trastu-TCO dimer mass and an additional species that could correspond to Trastu-TCO trimer (464 ± 31 kDa).

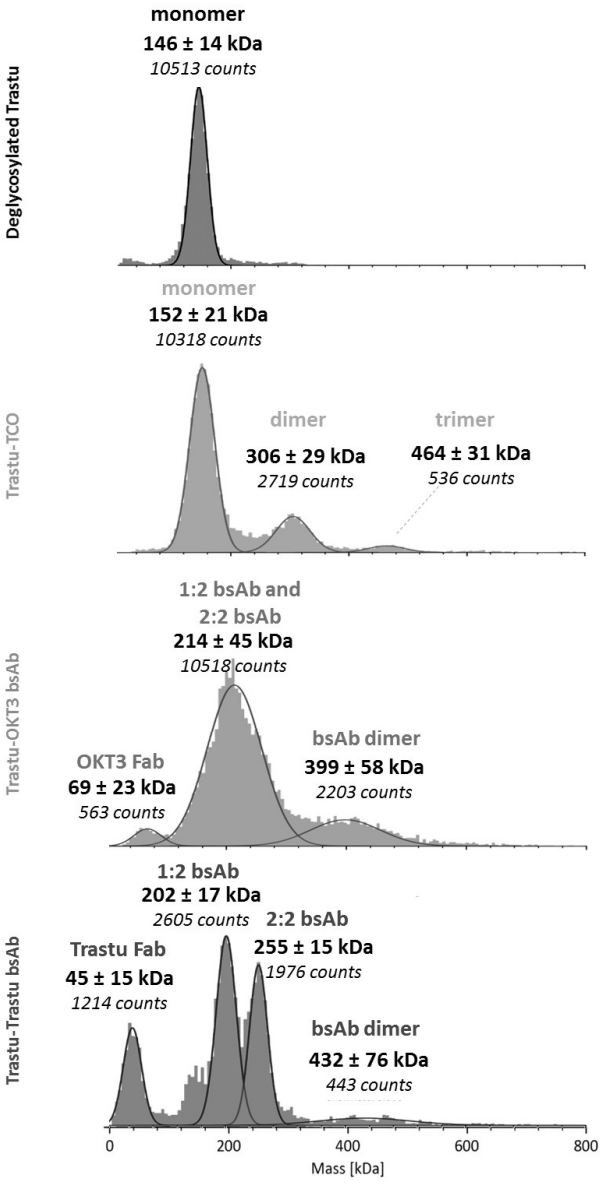


Figure 57: Mass photometry (MP) results of products obtained during development of bispecific antibodies through biogenic tyrosine-based click chemistry. **(A)** Profile of deglycosylated trastuzumab using PNGase F leading to one homogeneous species with mass corresponding to monomeric trastuzumab. **(B)** Profile of TCO functionalized trastuzumab with three species corresponding to monomer, dimer and trimer of (Trastu-TCO). **(C)** Profile of trastuzumab-OKT3 bsAb with major signal corresponding to overlapping species of different bsAb forms (1:1 and 1:2) and minor signal corresponding to OKT3 Fab fragments and bsAb dimer. **(D)** Profile of trastuzumab-trastuzumab bsAb with four species corresponding to trastuzumab Fab fragment, 1:2 and 1:1 bsAb forms and finally bsAb dimer.

The latter species was previously unassigned by nMS due to its high molecular weight, a hypothesis that is further confirmed thanks to our MP results. The single particle count analysis allowed quantifying 76%, 20% and 4% of the Trastu-TCO monomer, dimer and trimer, respectively, values that are comparable with SEC-UV quantification (**Table 9).** We also evaluated our MP approach for the bsAb constructs analysis, which resulted in the observation of three distinct distributions for the Trastu-OKT3 bsAb sample; OKT3 unbound Fab fragment (69 ± 23 kDa, 4%), bsAb 1:2, 2:2 or both constructs (214 ± 45 kDa, 79%) and bsAb dimer (399 ± 58 kDa, 17%). The Trastu-trastu bsAb revealed similar distributions with more distinguishable bsAb constructs namely; Trastu unbound Fab (45 ± 15 kDa, 19%), 1:2 Trastu-trastu bsAb (202 ± 17 kDa, 42%), 2:2 Trastu-Trastu bsAb (255 ± 15 kDa, 32%) and finally bsAb dimer (432 ± 76 kDa, 7%). Based on MP results, each species was relatively quantified and compared with nMS data as summarized in **Table 9.**

Table 9: Comparative quantification of species obtained during development of bispecific antibodies through biogenic tyrosine-based click chemistry upon SEC-nMS and MP experiments. Relative quantification of SEC-nMS experiments is obtained through the integration of peaks from SEC-UV chromatogram. MP relative quantification is based on the number of counts. Both quantifications are overall comparable, with major signal corresponding to bsAb monomer forms at the end of each reaction.

Sample	Species	Relative quantification based on SEC-nMS (%)	Relative quantification based on MP (%)	Δ (%)
Deglycosylated Trastu	Monomer	100	100	0
Trastu-TCO	Trimer	7	4	3
	Dimer	22	20	2
	Monomer	71	76	5
Trastu-OKT3 bsAb	BsAb dimer?	8	17	9
	2:2 bsAb	54	79	7
	1:2 bsAb	32		
	OKT3 Fab	6	4	2
Trastu-Trastu bsAb	BsAb dimer?	20	19	1
	2:2 bsAb	54	32	22
	1:2 bsAb	24	42	20
	Trastu Fab	2	7	3

Overall, MP offered mass measurement that could match with those obtained upon SEC-nMS experiments. Moreover, the identity of some species, namely bsAb dimers and Trastu-TCO trimer was

revealed uniquely upon MP analysis thanks to the wide mass range and the ability of analyzing low concentration samples, contrary to nMS, which requires higher µM concentrations to enable the characterization of proteins. Nevertheless, MP still encounters some limitations such as the lower resolution compared to nMS. This is highlighted in the case of Trastu-OKT3 bsAb analysis, where although the average mass is in agreement with nMS results, it could correspond to different bsAb constructs possibilities (2:2 or 1:2 bsAbs). Moreover, both species could be co-existing within the observed range (169-259 kDa) of the broad peak, however results afforded with SEC-nMS confirmed the identity of these peaks, highlighting the synergy between both techniques.

In terms of relative quantification, our MP data were comparable in the case of deglycosylated trastuzumab and trastuzumab-TCO with a neglected difference of relative quantity ranging from 0-5%. However, in the case of the trastuzumab bsAbs, some differences were noticed. For Trastu-OKT3 bsAb, only one peak was observed in MP that could correspond to different overlapping species and therefore this species was estimated to 79%. Conversely, in SEC-UV, both species were separated (54% and 24% for 2:2 and 1:2 bsAbs, respectively), therefore the sum of both species represents 86% of the UV signal, which is comparable with MP data (+7% in the case of SEC-UV). Additionally, the bsAb dimer was noticeably higher upon MP analysis (+13%) compared to SEC-UV. For Trastu-trastu bsAb, the bsAb dimer (20% and 19% using MP and nMS, respectively) and the Fab (7% and 2% using MP and SEC-UV, respectively) relative quantities were comparable for both techniques (19% and 20% using MP and SEC-UV, respectively). However, using MP the 1:2 bsAb format was the most abundant (42% versus 24% for nMS) while the 2:2 bsAb was less abundant in MP (32%) than in SEC-UV (54%). These different quantifications could be due to errors in Gaussian fitting during MP data processing. This limitation could be overcome by screening different MP conditions (replicate analysis, different buffers or different concentrations). However, during this project the majority of the samples were used for SEC-nMS analysis as a matter of priority, thus MP was only evaluated in this context.

Overall, the high sensitivity (nM) of MP allowed to characterize the bsAb constructs in a straightforward manner, along with the identification of HMWS thanks to the broad mass range. Albeit the limited resolving power of MP prevented distinguishing between different bsAb conformations, nMS fills this gap by confirming the identity of those species through accurate mass measurement which overall highlight the complementarity of these two techniques.

4. Conclusions

In this chapter, we aimed at using the BioAccord LC-MS platform for the characterization of bispecific antibodies. However, the limited mass range (until 7000 m/z) do not allow the observation of 1:2 and 2:2 forms of bispecifics that appear at higher m/z ranges, which makes the BioAccord an unsuitable

system for the analysis of bispecifics in native conditions, for the moment. The experiments conducted on a another ToF instrument (Synapt G2) coupled to SEC separation, allowed accurate mass measurements for the identified size variants namely, deglycosylated trastuzumab, functionalized trastuzumab and the different forms of trastuzumab bispecifics. Furthermore, SEC allowed the relative quantification of the different species, which confirmed that the 2:2 bispecific construct was the main species Moreover, we suggested here MP as a complementary technique to have a direct snapshot on the bsAb constructs in their storage buffer, using sub-nM concentrations. MP and nMS both showed several benefits and boundaries upon bsAb characterization, which makes their combination a key point to assess the development of novel bsAb constructs.

5. Scientific communications

Shajan I., Rochet L., Tracey S. R., Jackowska B., Benazza R., Hernandez-Alba O., Cianférani S., Scott C. J., Chudasama V., van Delft F. L. and Albada B. **Rapid Access to Potent Bispecific T-cell Engagers with Biogenic Tyrosine Click Chemistry**. *Bioconjugate Chem*, **2023**.

Chapter 4: Development of a SEC-MS method in denaturing conditions (dSEC-MS) for adapted and specific in-depth analysis of rebriged mAb-based formats

1. Analytical context

In the chapters discussed beforehand, we highlighted some of the interesting applications of native mass spectrometry (nMS) coupled to non-denaturing chromatographic techniques, mainly SEC, for the characterization of biotherapeutics. Indeed this approach evolved as a powerful tool for the straightforward characterization of ADC/PDCs. However, for some specific conjugation strategies, it is worth noting that routine and common LC-MS methods are not always suitable depending on the required information. In this context, we focused our analytical development to investigate one of the emerging site-specific conjugation strategies, disulfide rebridging reaction [80-83, 301]. This reaction consists in reducing the solvent-accessible disulfide bonds of intact antibodies or Fab fragments in mild conditions, followed by incorporation of one payload per disulfide bond upon covalent attachment through a rebridging agent **(Figure 58)**.

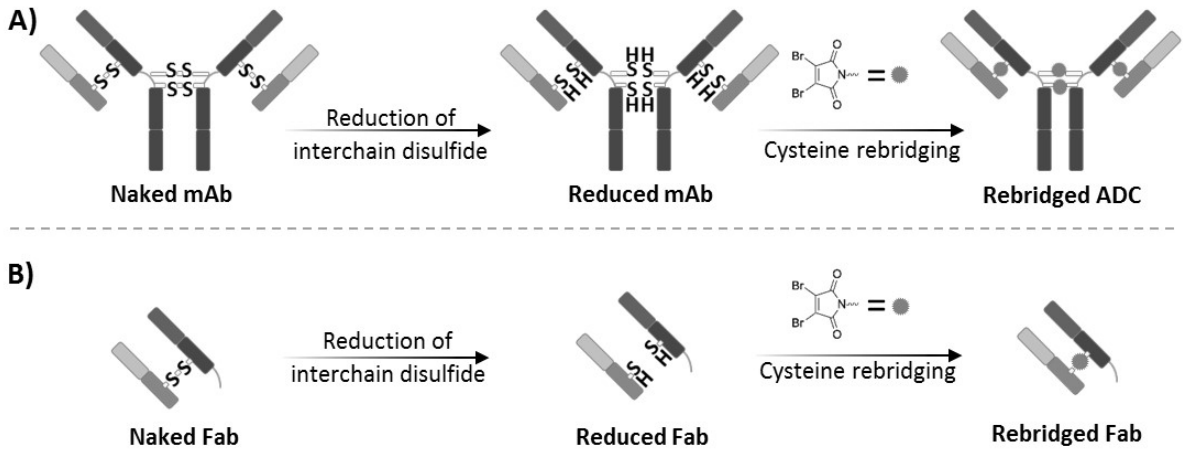


Figure 58: Conjugation of mAb subunits upon cysteine rebridging reaction following disulfide bridges reduction. **(A)** Schematic representation of rebridged intact mAb formation leading to the incorporation of four payloads between heavy and light chains of the antibody. **(B)** Schematic representation of Fab fragment rebridging leading to the incorporation of one payload between the Fc and Fd subunits.

In fact, incomplete conjugations could lead to a mixture of covalently rebriged and unbound species that are depicted in **Figure 59**, which requires the simultaneous detection of these species. To do so, sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is the reference method to reveal the approximative mass of the reaction products [83]. However, it suffers from low resolution, low mass accuracy and low sensitivity. Therefore, new approaches should be developed for a more comprehensive characterization of these products.

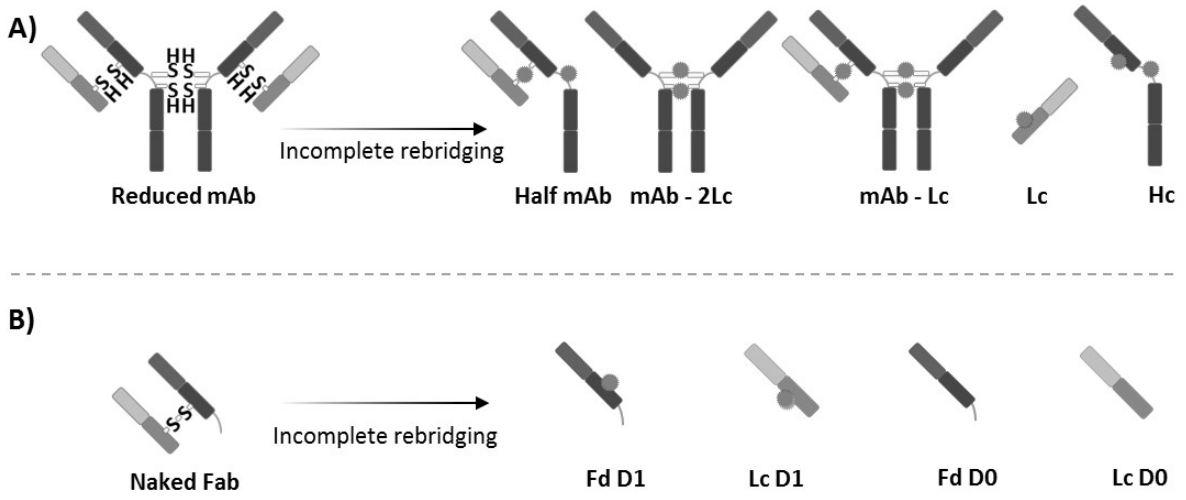


Figure 59: Incomplete conjugation of mAb subunits upon cysteine rebridging reaction following disulfide bridges reduction. **(A)** Incomplete intact mAb conjugation leads to a mixture of by-products such as half mAb, mAb without 2 Lcs, mAb lacking one Lc, Lc and Hc subunits. **(B)** Incomplete Fab fragments conjugation could lead to a mixture of by-products such as Fd with one payload, Lc with one payload and unconjugated Fd and Lc subunits.

Schematic representation of incomplete rebridging of a reduced mAb leading to the development of heterogeneous products.

2. Objectives

In this chapter, we aim at developing a straightforward approach to monitor the optimization of a one-pot reduction-rebridging reaction in collaboration with Chaubet's group (University of Strasbourg, France). This conjugation reaction aims at incorporating covalently one (in the case of Fab rebridging) or four (in the case of intact mAb rebridging) hypervalent iodine (EBZ) [302] molecules as payloads, leading to a mass increase of +552 Da per rebridged drug incorporation. The EBZ reagent synthetized by Waser's group (EPFL, Switzerland), was used thanks to its ability to react through its two cysteine reactive sites that have significant difference in their reactivity (**Figure 60**).

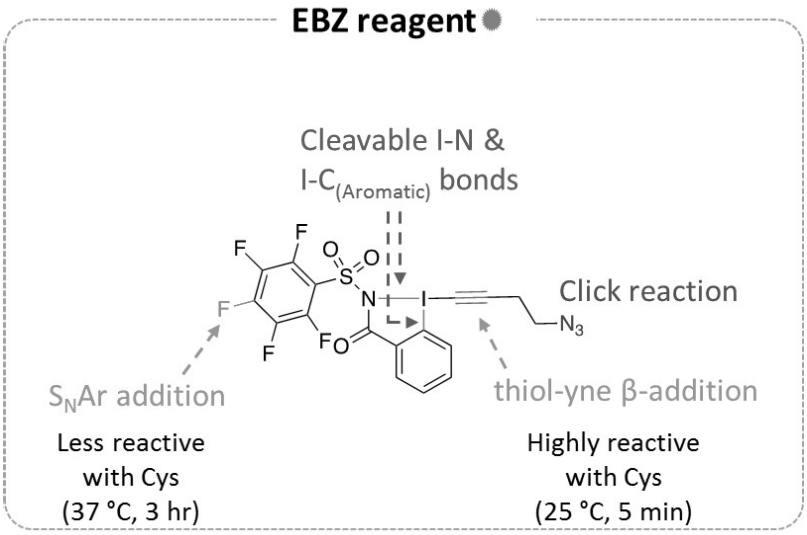


Figure 60: Structure of the EBZ reagent used for the rebridging reaction of intact mAb and Fab fragments. The two cysteines reactive sites react using SnAr addition (pink) and thiol-yne β-addition (green), respectively. In case of cleavage of the EBZ reagent at the cleavable I-N and I-C bonds level, by-products could be obtained.

We first illustrate the benefits and the limitations of classical LC-MS methods, particularly SEC-nMS and RPLC-MS, to deeply characterize the resulted species. Afterwards, we propose an adapted dSEC-MS method, combining the ability of SEC to separate size variants and the denaturing MS capabilities to distinguish between covalent and non-covalent rebriged assemblies. Our all-in-one developed dSEC-MS method aims at screening the optimal conditions for obtaining rebriged moieties along with the assessment of the avDAR, DLD of the latter and the rebridging efficiency.

3. Results

3.1 SEC-nMS analysis as our go-to method for the characterization of rebridged formats

1) SEC-nMS of Fab fragments

Considering that SEC-nMS is our go-to method for the characterization of mAb-based biotherapeutics, we used it for the analysis of the rebriged Fab and mAb samples. Of note, this study was performed on two rebridged Fab fragments upon changing the storage solution of the rebridging reagent (Fab#A and Fab#B) and one rebridged trastuzumab. The conditions are summarized in **Table 10**.

Table 10: Summary of the reaction conditions used for the one pot reduction-rebridging of Fab-A, Fab-B and rebridged trastuzumab. The changed parameters are highlighted in different colors or in bold letters showing changes in mainly EBZ reconstitution buffer. The quantities of EDTA and TCEP are adjusted for rebriged intact trastuzumab.

Sample name	Preparation conditions
Fab#A	80 µg Fab-Trastuzumab in BBS***** was reacted with **5** equiv. EBZ (**in 10 mM DMSO solution**) and **5** equiv., TCEP (in 15 mM H_2O solution) and was incubated at 37 °C for 5 h
Fab#B	80 µg Fab-Trastuzumab in BBS***** was reacted with **5** equiv. EBZ (**in 10 mM ACN solution**) and **5** equiv. TCEP (in 15 mM H_2O solution) and was incubated at 37 °C for 5 h.
Rebriged trastuzumab	100 µg Trastuzumab in BBS***** was reacted with **10** equiv. EBZ (**in 10 mM ACN solution**) and **10** equiv., TCEP (in 15 mM H_2O solution) and was incubated at 37 °C for 5 h.

*****BBS buffer contains: 25 mM H_3BO_3, 25 mM NaCl, 2 mM EDTA, pH 8.0

SEC-UV signal of unreacted Fab and rebridged Fab samples show similar profiles. Peak I was identified as a dimeric aggregate (~90 kDa) and the most intense peak (peak II) was attributed to the monomeric forms of Fab with D1 and D0 species being detected for both Fab#A and Fab#B. Although the D1 species is the most intense population (48192 ± 1 Da and 48188 ± 1 Da for Fab#A and Fab#B, respectively) in both cases, D1/D0 relative intensities are significantly different, leading to different avDAR values for Fab#A (0.8 ± 0.0) and Fab#B, (0.5 ± 0.0), thus corroborating that the different reaction conditions do not lead to the same degree of conjugation. Of note, an additional minor species is observed, corresponding to a truncated form of the Fab domain (-KTH residues in the C-terminal side of the Fd

domain) due to trastuzumab papain over-digestion (depicted as a grey circle in **Figure 61**, 47 271 ± 1 Da, 47 274 ± 1 Da and 47 275 ± 1 Da for untreated Fab, Fab#A and Fab#B, respectively). Considering the relative intensities of D1/D0, a rebridging yield of 79 ± 1% and 55 ± 5% was calculated for Fab#A and Fab#B, respectively.

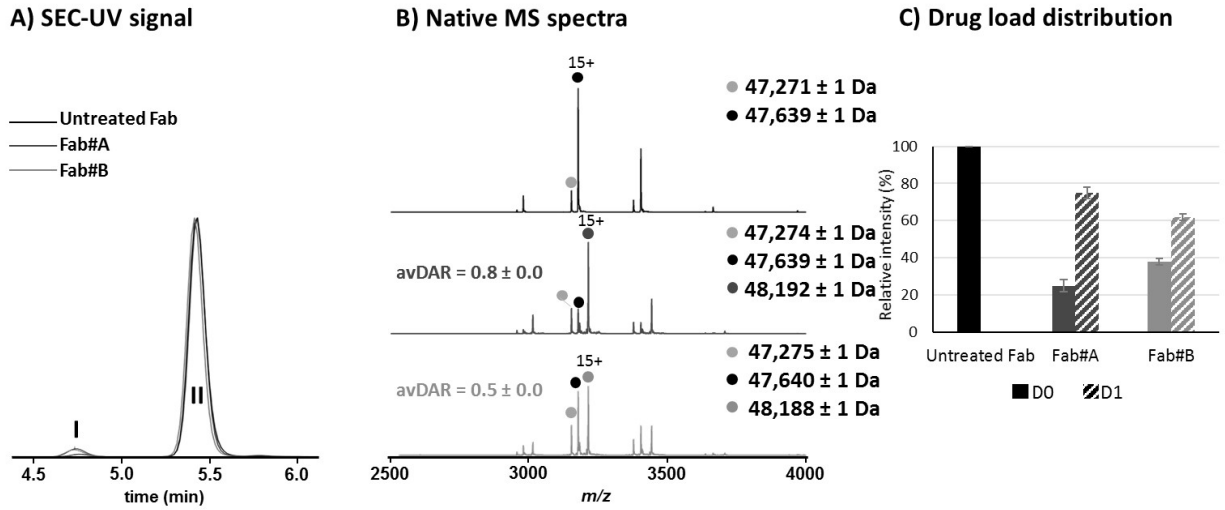


Figure 61: SEC-nMS analysis of unreacted Fab (black), rebridged Fab#A (blue) and rebriged Fab#B (green). **(A)** SEC-UV signal showing one major peak at 5.3 min with a neglected peak at 4.7 min, same profile is observed for the three samples **(B)** The corresponding native MS spectra of the main identified species based on the UV signal that corresponds to monomer of Fab fragments. Unconjugated Fab annotated as D0 is observed for the three samples, with Fab lacking KTH amino acids (-366 Da, theoretically) due to over-digestion. Fab with one payload (+550 Da, theoretically) annotated as D1 is observed in case of Fab#A and Fab#B, as expected. The avDAR of the conjugated Fab fragments is shown. **(C)** Drug load distribution based from the native MS spectra of each sample. The standard deviation of the avDAR and DLD values is calculated from triplicate analysis.

2) SEC-nMS of mAb samples

SEC chromatogram of untreated and rebriged trastuzumab revealed two peaks **(Figure 62A)**. Peak I was attributed to the dimer (~290 kDa), while the most intense one (peak II) was identified as the monomer. The rebriged trastuzumab peak is slightly shifted to a higher retention time and show a higher peak width peak (FWHM = 0.1 min for unreacted mAb and FWHM = 0.2 min for rebridged mAb), in line with what was prior described in the literature when comparing the SEC between an ADC and its unconjugated mAb counterpart. This is explained by the fact that the conjugated molecules results in increased hydrophobicity of the targeted ADC, thus increasing non-specific interaction between the analyte and the stationary phase of the column.

The corresponding nMS spectra of the main species reveal a single species for naked trastuzumab (145 864 ± 2 Da) and at least four species for conjugated trastuzumab, with masses correlating with D1, D2, D3 and D4 species. The relative intensities lead to an avDAR value of 2.5 ± 0.1. Note that the reaction aimed at conjugating four payloads (i.e. avDAR of 4.0 expected). Therefore, the results pinpoint to an incomplete rebridging reaction **(Figure 62B, 27C)**.

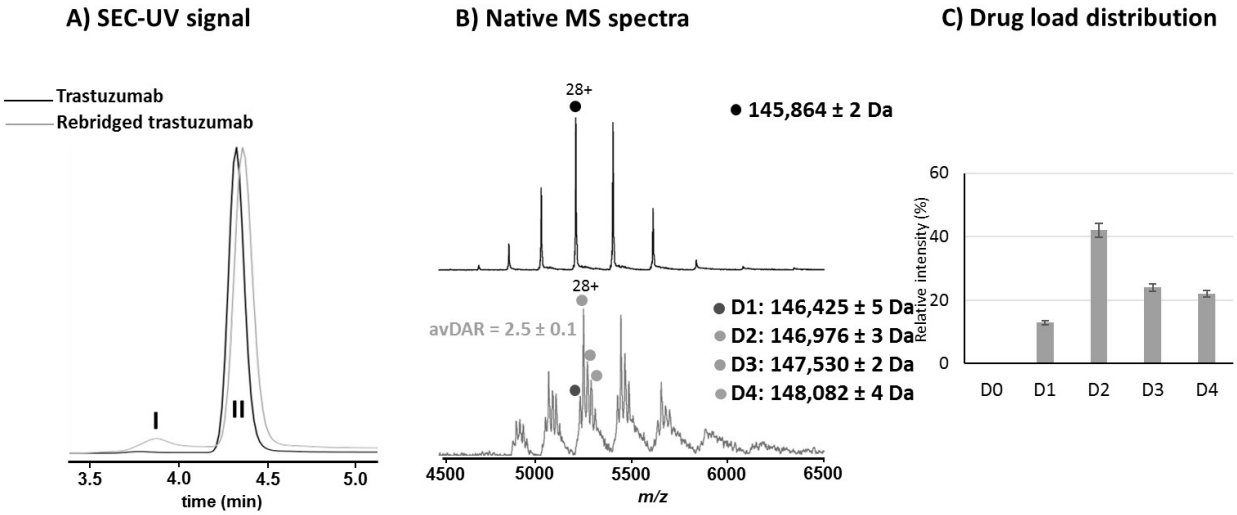


Figure 62: SEC-nMS analysis of unreacted trastuzumab (black) and rebridged trastuzumab (blue). **(A)** SEC-UV signal shows one main peak in both cases around 4.3 min and a minor peak at 3.7 min. **(B)** The nMS spectra of the main identified species that corresponds to monomeric trastuzumab. For the untreated trastuzumab one species is observed corresponding to D0 as expected. The nMS spectra of rebriged trastuzumab reveals four masses corresponding to D1, D2, D3 and D4 species highlighting the partial incorporation of four drugs. The avDAR value of 2.5 confirms the heterogeneous conjugation of four drugs instead of exclusively producing a D4 species. **(C)** Drug load distribution of the conjugate trastuzumab based on nMS spectra showing the relative intensities of D1-D4 species with D2 species being the most intense, in-line with the avDAR value. The standard deviation of the avDAR and DLD values is calculated from triplicate analysis.

3) Conclusions

The results obtained upon SEC-nMS experiments allowed a straightforward characterization of rebridged Fab/mAb and their conjugated counterparts, providing precise mass measurement of the identified species along with avDAR and DLD assessment. Taking into account that the rebridging reaction aims at covalently attaching subunits upon payloads incorporation, in case of insufficient reaction, free subunits are expected. However, in native conditions, those moieties are not observed probably to the fact that they are maintained by non-covalent interactions that are preserved in native conditions.

3.2 RPLC-MS analysis of rebridged mAb-based formats

1) RPLC-MS of Fab samples

In order to investigate the co-existence of covalent and non-covalent species upon the rebridging reaction of Fab and mAb samples, we performed RPLC-MS analysis as it is the most routine LC-MS method in classical denaturing conditions. The combination of organic solvents with acidic pH, we are able to induce the denaturation of the proteins studied, leaving only the strong covalent interactions maintained.

Interestingly, this time the chromatographic profile of the conjugated Fab fragments showed three different regions, with one major peak around ~12.87 min for both reacted Fab samples. The UV signal

of the untreated Fab showed only two peaks at 12.71 min and 13.20 min. Mass deconvolution of the identified peaks demonstrated the presence of only naked species for untreated Fab (D0 and D0-KTH), while a mixture of intact Fab fragments and free Lc and Fd species was highlighted in the case of Fab#A and Fab#B samples, as expected **(Figure 63, Table 11)**. In fact, the unrebridged species in this case correspond to Lc and Fd subunits conjugated to only fragments of the payload **(Figure 64)**.

Furthermore, the relative intensities of D0/D1 species based on the MS spectra of conjugated species, allowed the calculation of comparable avDAR values than those obtained through SEC-nMS experiments (0.8 ± 0.0 and 0.6 0.8 ± 0.0 for Fab#A and Fab#B, respectively).

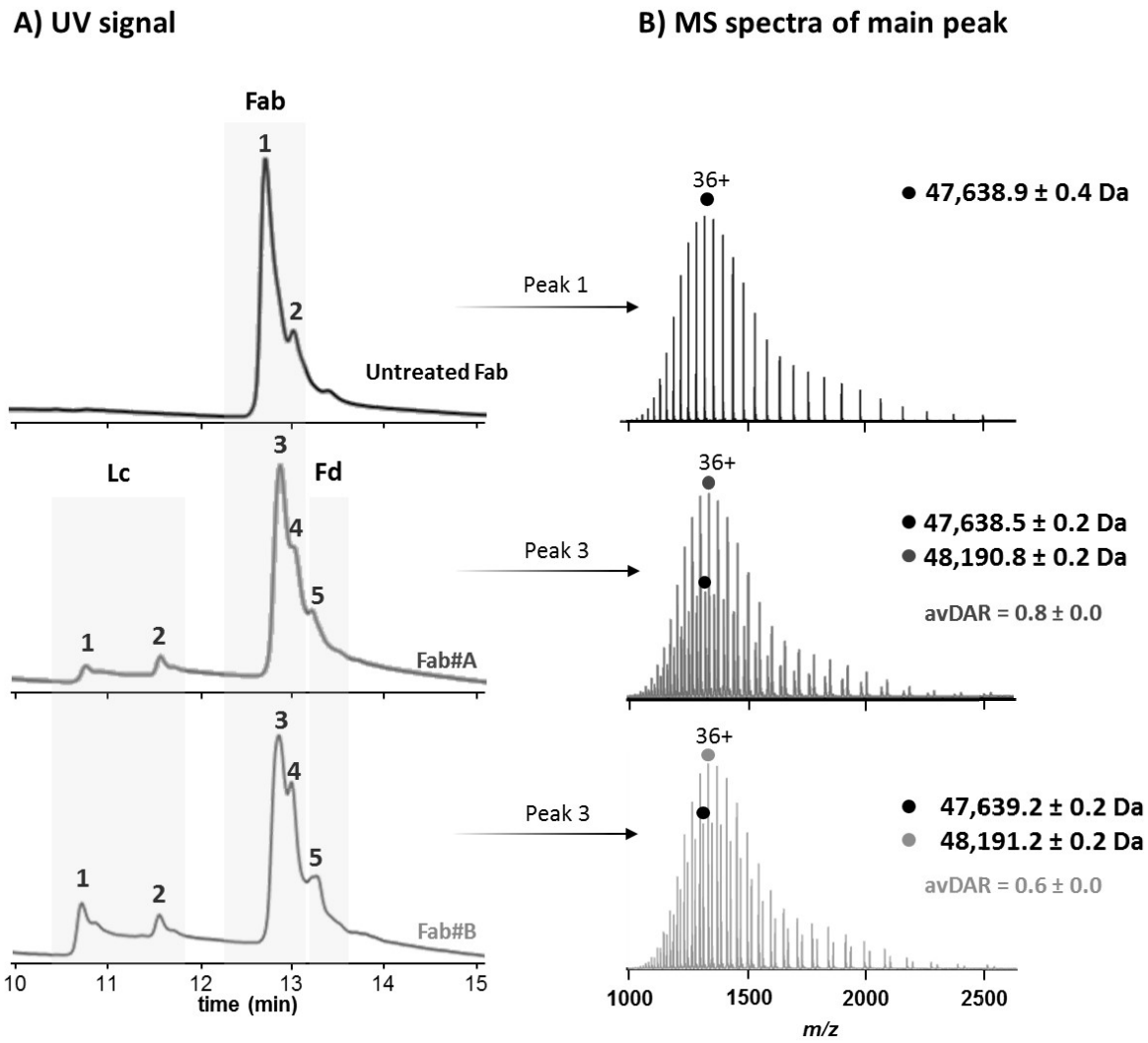


Figure 63: RPLC-MS experiments of untreated Fab (black), Fab#A (blue) and Fab#B (green). **(A)** UV signal showing one peak for untreated Fab with a minor peak shoulder due to presence of over-digested Fab species. The conjugated Fab#A and Fab#B exhibit more peaks resumed in three regions with major peak corresponding to intact Fab species. **(B)** The corresponding MS spectra of the most intense peak i.e. the peak corresponding to intact Fab species centered on the charge state 36+. The spectra reveal masses of unconjugated Fab in the three cases, with an additional mass corresponding to D1 species for Fab#A and Fab#B. Average DAR values are calculated based on the intensities of four charge states and the standard deviation of the avDAR was calculated from triplicate analysis.

Table 11: Peak identification for untreated Fab, Fab#A and Fab#B based on RPLC-MS analysis. The retention time of the peaks observed in **Figure 28** is indicated here. The identities of all species are confirmed by the observed masses that is comparable to the expected mass. Mass accuracy is calculated based on the difference of error between experimental and theoritical mass. The standard deviation of the observed mass is obtained upon mass detection from different charge states.

	Peak	Observed rt(min)	Species	Observed mass (Da)	Expected mass (Da)	Mass accuracy (ppm)
Untreated Fab	1	12.71	Fab D0	47,638.9 ± 0.4	47,637.1	37
	2	13.20	Fab - KTH	47,272.5 ± 0.6	47,270.9	33
Fab#A	1	10.77	Lc	23,439.3 ± 0.2	23,439.0	12
	2	11.57	Lc+457 Da	23,896.5 ± 0.2	23,896.0	19
	3	12.87	Fab D0	47,638.5 ± 0.2	47,637.1	28
			Fab D1	48,190.8 ± 0.2	48,189.1	34
	4	13.02	Fab-KTH	47,272.0 ± 0.2	47,270.9	22
	5	13.20	Fd+93 Da	24,293.5 ± 0.5	24,293.1	15
			Fd+457 Da	24,658.3 ± 0.8	24,657.1	49
Fab#B	1	10.72	Lc	23,439.8 ± 0.1	23,439.0	32
	2	11.53	Lc+457 Da	23,896.8 ± 0.2	23,896.0	33
	3	12.85	Fab D0	47,639.2 ± 0.2	47,637.1	43
			Fab D1	48,191.2 ± 0.2	48,189.1	42
	4	13.00	Fab-KTH	47,272.5 ± 0.5	47,270.9	34
	5	13.20	Fd+93 Da	24,293.6 ± 1.2	24,293.1	21
			Fd+457 Da	24,658.0 ± 0.2	24,657.1	35

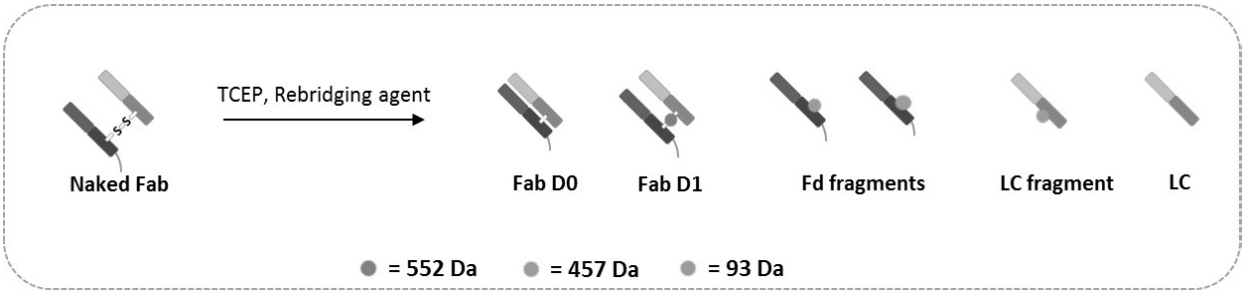


Figure 64: Scheme of incomplete rebridging of a Fab fragment with the actual different rebriged and by-products obtained upon incorporation of one EBZ reagent, showing the rebriged D1 obtained with presence of D0 and the different subunits conjugated with fragments of the payload. The exact masses are shown in **Table 6**.

2) RPLC-MS of mAb samples

Similarly to mAb samples, the UV signal of rebridged trastuzumab was more informative than in native conditions as it reveals additional signals, while one unique chromatographic peak at 13.30 min is observed for intact trastuzumab **(Figure 65)**. The identity of each species was further revealed by MS, which is summarized in **Table 12**. Albeit here we could prove the co-existence of rebridged trastuzumab species with unbound subunits, the relative quantification of these species remain a challenge as many of them overlapped due to their similar hydrophobicity.

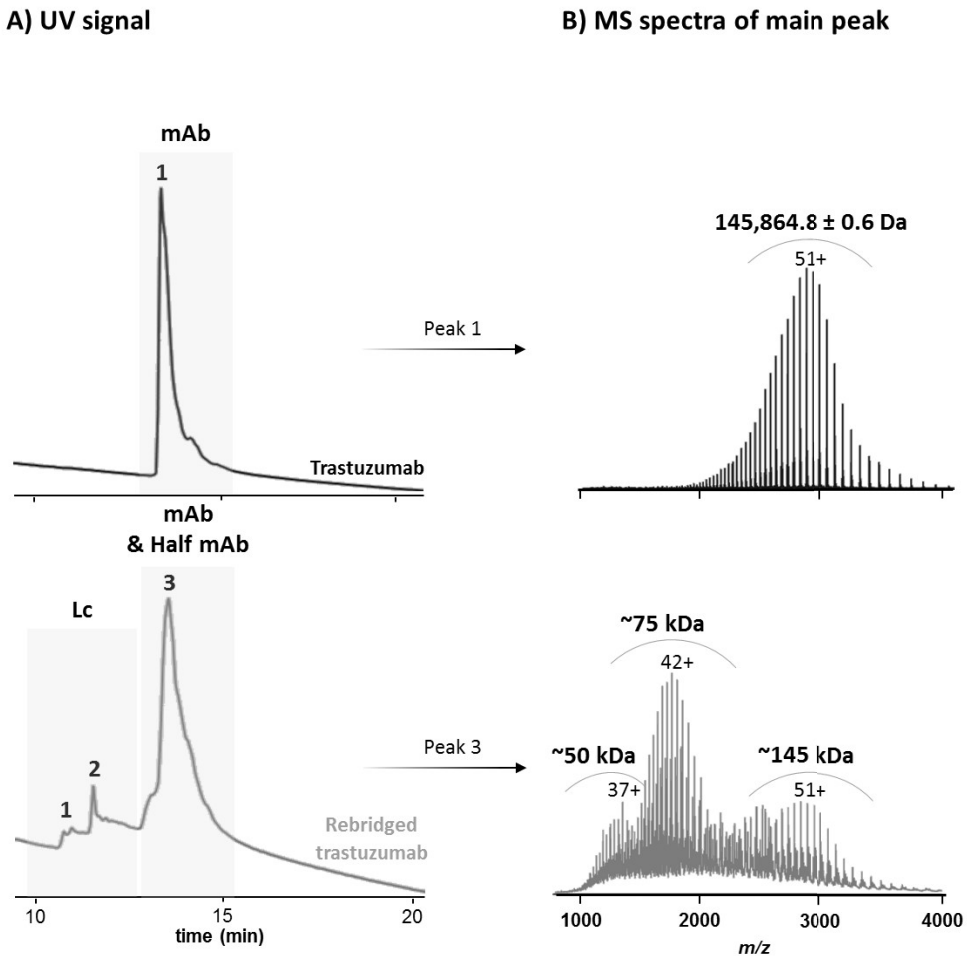


Figure 65: RPLC-MS experiments of unreacted trastuzumab (black) and rebridged trastuzumab (blue) **(A)** UV signal of the unreacted trastuzumab showing one major peak at 13.30 min while the profile of rebriged trastuzumab exhibits additional minor peaks before the main peak at 13.35 min. **(B)** The MS spectra of the most intense peak of unreacted trastuzumab revealing a mass corresponding to D0 species. In case of rebriged trastuzumab a mixture of overlapped species is observed corresponding to trastuzumab lacking the light chain (~50 kDa), half trastuzumab (~75 kDa) and intact trastuzumab (~145 kDa) species. The exact masses are noted in **Table 7**.

Table 12: Peak identification for untreated intact trastuzumab and rebridged trastuzumab based on RPLC-MS analysis.

	Peak	Observed rt(min)	Species	Observed mass (Da)	Expected mass (Da)	Mass accuracy (ppm)
Trastuzumab	1	13.30	mAb D0	145,864.8 ± 0.6	145,865.2	3
Rebridged trastuzumab	1	10.75	Lc	23,439.9 ± 0.1	23,439.0	35
	2	11.53	Lc+457 Da	23,896.9 ± 0.1	23,896.0	35
		12.15	2Lc D0	46,877.9 ± 0.4	46,878.1	5
			2Lc D1	47,430.2 ± 0.4	47,428.1	45
	3	13.35	mAb/2 D2-(LC+457 Da)	50,143.0 ± 0.6	/	/
			mAb/2 D1	73,488.6 ± 1.0	73,486.6	27
			mAb/2 D2	74,041.0 ± 0.3	74,038.6	32
			mAb D1-2Lc	99,545.0 ± 0.8	99,545.1	1
			mAb D1-(Lc+457 Da)	123,070.4 ± 0.8	/	/
			mAb D1	146,419.9 ± 6.9	146,417.2	18
			mAb D2	146,974.5 ± 1.2	146,969.2	36
			mAb D3	147,528.3 ± 2.3	147,521.2	48
			mAb D4	148,082.5 ± 1.3	148,073.2	63

3) Conclusions

Indeed, RPLC-MS experiments reveal more information than SEC-nMS regarding the conjugated Fab and trastuzumab samples, highlighting the co-existence of covalently rebriged and free species. However, the co-elution of several species due to their similar hydrophobicity makes it challenging to estimate relatively the quantity of side-products, hampering the straightforward monitoring of the reaction conditions.

3.3 Hybrid dSEC-MS for rebridged mAb-based formats characterization

Bearing in mind that the covalently rebridged and the free species could be distinguished only under denaturing conditions, and taking advantage of SEC capabilities to separate those species based on their hydrodynamic volume, we aimed at performing an adapted method based on SEC-MS coupling in denaturing conditions (dSEC-MS).

1) State of the art and objectives

Historically, the first studies to report the hyphenation of SEC to MS aimed at characterizing antibodies following their on-line buffer exchange for fast and precise mass measurement [303-304]. In another perspective, a study reported the benefits of SEC-MS for the characterization of reduced mAbs to obtain direct and precise mass measurement of the mAb subunits [305]. In fact, these experiments were conducted under classical denaturing conditions using acidic organic solvents to achieve a good separation along with ionization of the separated species in the gas phase. Long columns were used employed to increase the separation efficiency and resolution. However, these conditions are not really effective for high throughput experiments due to the time-consuming analysis and therefore need to be optimized for the purpose of our analysis.

2) Optimization of LC and MS parameters

First and foremost, the choice of the SEC column has an important role of reducing the analysis while keeping an effective separation. As the retention time decreases proportionally to the column length, shorter columns might be better to increase the throughput of analysis, as well as the resolution is sufficient. Therefore, some parameters could be compromised such as using lower particle size (sub µm) and bio-inert columns for example. In this context, we evaluated two SEC columns using Fab#A as a reference: the AdvanceBioSEC3 300Å 2.5 µm 4.6 x 150 mm column (Agilent) and the bio-inert Maxpeak Protein SEC 250Å 2.5 µm 4.6 x 150 mm (Waters). For both experiments, the voltage cone and desolvation temperature were fixed at 60V and 330 °C. Mobile phase was 20% ACN + 0.1 TFA + 0.1% FA, in isocratic mode for 15 min at a flowrate of 100 mL/min.

Although in both cases two main regions are detected, the UV profile was rather different. The peaks observed when using the AdvanceBioSEC3 column were larger (FWHM = 0.635 min versus FWHM = 0.190 min for the Maxpeak, for the major peak). The Maxpeak provides better separation for the two peak regions allowing a resolution of 5.8 (versus Rs = 0.73 for the AdvanceBioSEC3). Moreover, the major peak is significantly asymmetrical when using the AdvanceBioSEC3 column (As = 3.9 versus 1.7 for the MaxPeak). For the second region, the observed peaks are large in both cases due to the overlapping of two different species with similar size (~25 kDa) (**Figure 66A).**

The MS deconvolution confirmed the identity of each species, as it is shown in **Figure 66B**. The first peak corresponds to a mixture of D0 and D1 species of rebriged Fab in addition to a species corresponding to over-digested Fab (Fab-KTH, depicted as a star in **Figure 66B**). The second regions indeed correlates with overlapping species that could be separated thanks to the XIC signal. The species correspond to unconjugated Lc, and also to Lc and Fd conjugated to fragments of the payload (explained in **Figure 64**). Overall, the interactions between the analytes and the stationary phase of the Maxpeak Protein SEC column seem to be reduced, which allow us to select this column for further dSEC-MS analyses.

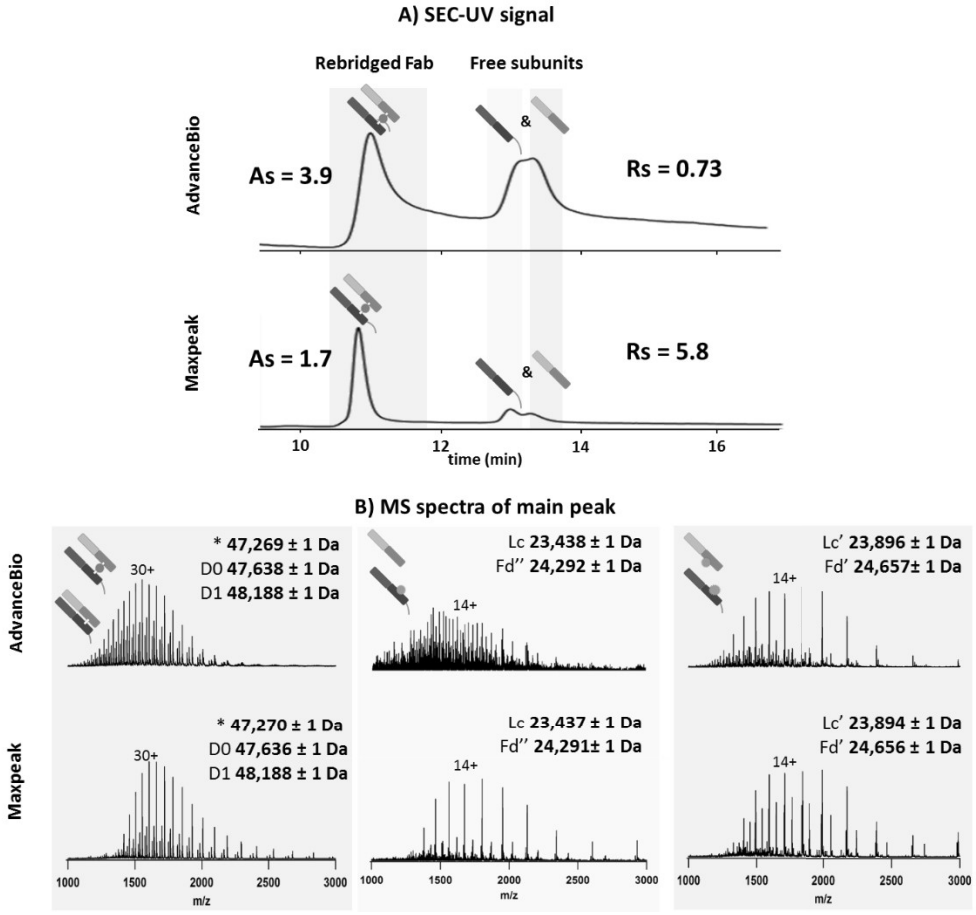


Figure 66: Comparison of Maxpeak and Advancebio SEC columns upon dSEC-MS analysis of rebriged Fab#A fragment. (**A**) SEC-UV signal of Fab#A sample using Advance BioSEC3 column from Agilent (**top**) and Maxpeak protein SEC column from Waters (**bottom**). A decrease in peak asymmetry (As = 1.7) and increase of resolution (Rs = 5.8) are noticed when switching to the

Maxpeak protein SEC column. Both columns exhibit three different distribution with one main peak and two over-lapping species eluting later. (**B**) MS spectra of the identified species in the three regions observed in the UV signal. The first region (**grey**) corresponds to a mixture of Fab lacking the KTH amino acids depicted as a star, unconjugated Fab (D0) and Fab with one payload (D1). The two overlapped regions (**green and blue**) correspond to a mixture of unconjugated and conjugated Lc and Fd species to fragments of the payload. The fragment with a mass of 93 Da is depicted as a green star while the fragment with a mass of 457 Da is represented in orange star.

3) Analysis of Fab and mAb samples

Performances and benefits of our optimized dSEC-MS method were obvious in case of comparing the two Fab fragments resulting from different reaction conditions: Fab#A and Fab#B, as well as for the rebridged trastuzumab analysis.

For the reference Fab sample, one sharp peak (FWHM = 0.19 min) is observed at 10.88 min on the dSEC chromatogram corresponding to the mass of the Fab fragment (47,638.9 ± 0.1 Da) co-eluting with Fab-KTH species (47,273.3 ± 0.3 Da). Interestingly, for Fab#A and Fab#B, three different peaks were detected according to dSEC-UV signal (versus only one peak observed in SEC-nMS, **Figure 67A**). The major peak for both samples (10.88 min) corresponds to co-elution of covalently rebridged Fab D1 species (48,190.8 ± 0.3 Da for both, Fab#A and Fab#B) and D0 species (47,638.6 ± 0.3 Da and 47,638.8 ± 0.3 Da for Fab#A and Fab#B, respectively). Fab-KTH species were also observed in this region (47,272.0 ± 0.4 and 47,272.0 ± 0.6 Da for Fab#A and Fab#B, respectively). The dSEC separation shows two partially resolved (Rs= 5.8 and Rs= 5.5 for Fab#A and Fab#B, respectively) additional peaks for both samples: peak 2 (~13.08 min) and peak 3 (~13.37 min) corresponding to Fd and Lc species, respectively. The deconvoluted mass spectra for peak 2 pinpoints the co-elution of both unmodified Lc (23,439.5 ± 0.2 Da and 23,439.6 ± 0.2 Da for Fab#A and Fab#B, respectively) and Fd+93 Da fragment (24,292.3 ± 0.5 Da and 24,292.6 ± 0.5 Da for Fab#A and Fab#B, respectively) species. The peak 3 shows the presence of half-rebridged species namely: Lc+457 Da (23,896.6 ± 0.4 Da and 23,896.7 ± 0.4 Da for Fab#A and Fab#B, respectively) and Fd+457 Da (24,657.6 ± 0.3 Da and 24,657.7 ± 0.5 Da for Fab#A and Fab#B, respectively) (**Figure 67B, C,** D). The experimental masses and the relative intensities corresponding to the Fab conjugates are overall in line with the masses reported with SEC-nMS, leading to a very similar avDAR value (0.8 ± 0.0 and 0.6 ± 0.1. for Fab#A and Fab#B, respectively). Similarly to SEC-nMS, relative intensities of chromatographic peaks of intact Fab versus non-rebridged fragments (such as Fd and Lc) could serve to easily approximate the ratio of covalent versus non-covalent rebridged Fab. Therefore, upon integration of the chromatographic peak area of the fragments signals, the signal corresponding to by-product species represents 11 ± 0% and 25 ± 2% of the total signal for Fab#A and Fab#B samples, respectively. However, as it can reasonably be assumed that Fab D0 and D1 species have similar ionization efficiencies, D1 species represent 75 ± 3% and 62 ± 2% of the Fab main peak area, respectively.

Performances and benefits of the dSEC-MS method were even more obvious for intact rebridged trastuzumab analysis. While one unique peak is observed for untreated trastuzumab, dSEC-UV shows

6 peaks corresponding to rebridging products of different sizes: intact mAb species (~150 kDa) are baseline resolved from half-mAbs (~75 kDa), HC (~50 kDa) and Lc (~25 kDa) species, which significantly improves the ease of data interpretation (**Figure 67E**). The most intense peak was assigned mainly to different intact rebridged trastuzumab species (from D1 to D4) with an avDAR of 2.5 ± 0.1 (**Figure 67F, G**), similar to the value calculated with the previous LC-MS methods. Of note, the chromatographic peak of rebridged trastuzumab is slightly larger compared to the reference trastuzumab (FWHM = 0.35 min versus 0.23 min for reference mAb). In this case, the peak broadening is caused by the co-elution of structures where either one (~123 kDa, mAb D1 lacking the fragment (Lc+457 Da)) or two Lcs (~99 kDa, mAb D1 lacking 2Lcs) have been released as a result of an incomplete rebridging reaction. This is consistent with the detection of free Lc species (11-14 min) (**Table 4**).

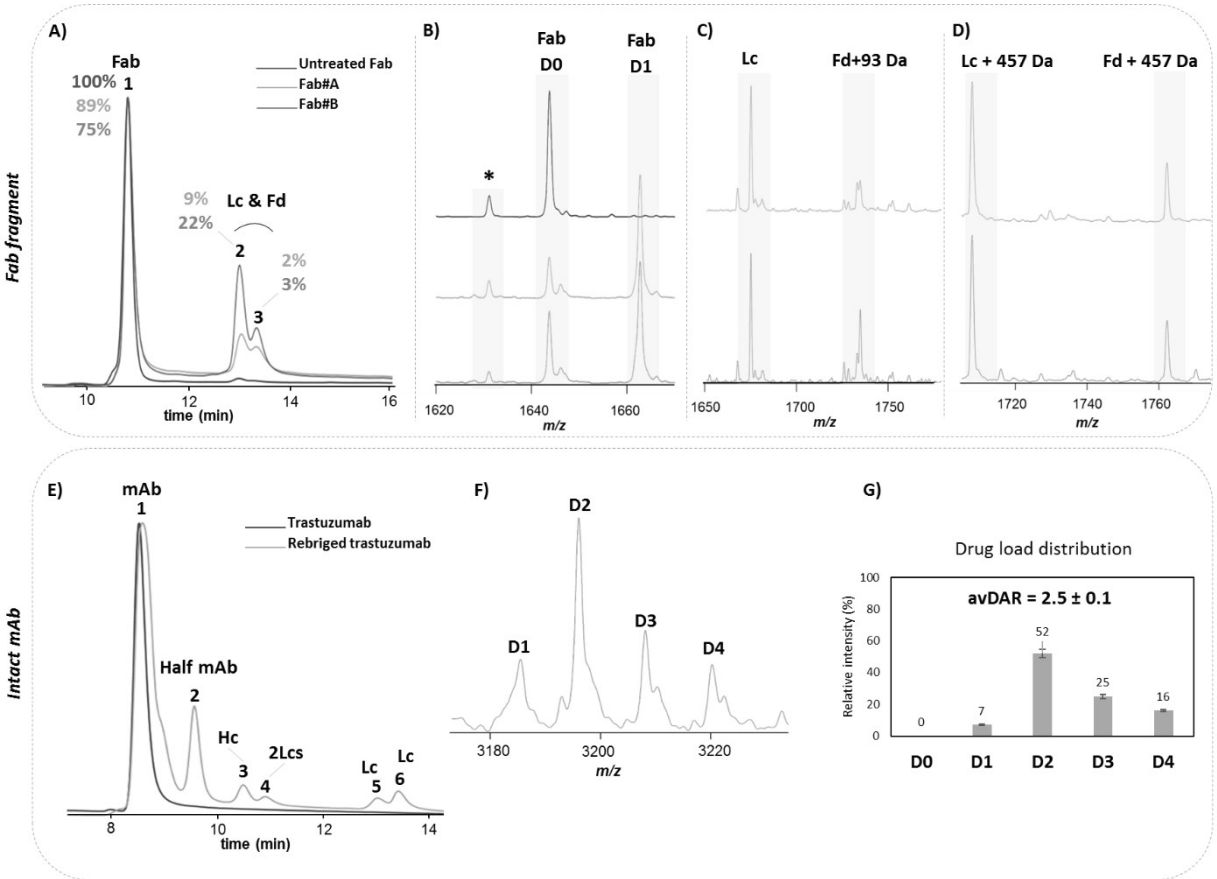


Figure 67: dSEC-MS analysis of reference Fab (blue), Fab#A (pink) and Fab#B (green). **(A)** Overlay of the UV-SEC signals of the three samples. The right side of the figure represents the corresponding mass spectra of each sample with **(B)** the zoom on the charge state 30+ for the Fab peak (~10.8 min), **(C)** zoom on the 14+ for the second peak (~12.9 min) and **(D)** is the zoom on 14+ for the last peak (~13.3 min). **(E)** SEC-UV profiles in denaturing conditions of naked trastuzumab (blue) and rebridged trastuzumab (pink). Peaks from 1 to 6 represent the different observed species with their corresponding masses detected further in the MS and summarized in table S3. **(F)** Mass spectrum of the intact rebridged mAb species zoomed on the 48+ and (g) represents the drug load distribution of the intact rebridged mAb.

However, in our dSEC-MS conditions, the peak representing the intact trastuzumab species is noticeably less complex to its counterpart in RPLC-MS analysis. This is due to non-overlapping half-mAb nor free Hc with intact mAb species (**Figure 68**).

Although this sample is very challenging due to its high heterogeneity, our method could clearly assign the avDAR along with the DLD of the full rebridged trastuzumab. More importantly, we could relatively quantify the covalent rebridging efficiency even if it is tricky to separate the mAb, the mAb-2Lc and the mAb-Lc, leading to monitor the reaction of intact trastuzumab rebridging.

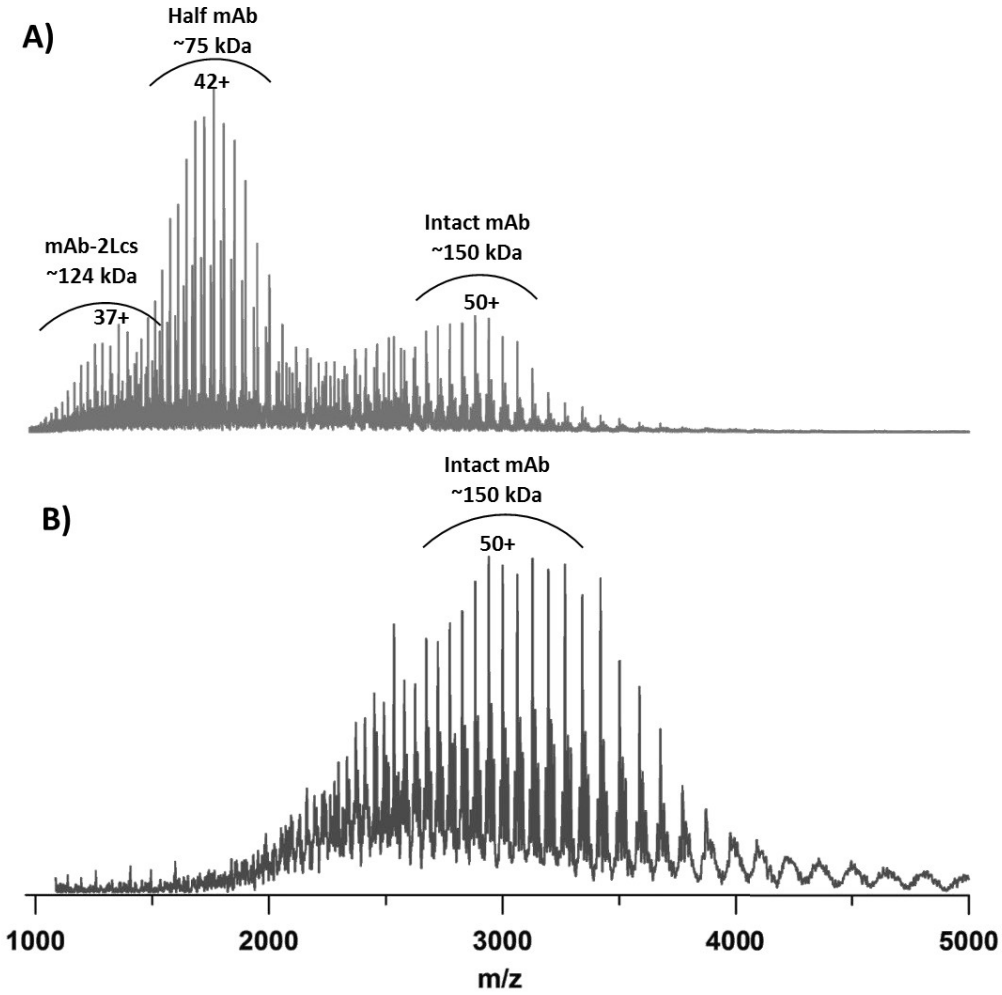


Figure 68: Overlay of MS spectra corresponding to intact trastuzumab signal after rebridging reaction, analyzed using (A) RPLC-MS showing the co-elution of 3 different species, and (B) dSEC-MS showing mostly one distribution corresponding to the intact mAb.

Table 13: Peak identification for untreated Fab, Fab#A and Fab#B, intact trastuzumab and rebridged trastuzumab based on dSEC-MS analysis. The retention time and identity of each peak in **Figure 11** is indicated here. The mass accuracy is calculated based on the error difference between the expected and the observed mass of each species.

Fab fragments						
	Peak	**Observed rt(min)**	**Species**	**Observed mass (Da)**	**Expected mass (Da)**	**Mass accuracy (ppm)**
Untreated Fab	1	10.88	Fab - KTH	47,273.3 ± 0.3	47,270.9	51
			Fab D0	47,638.9 ± 0.1	47,637.1	38
Fab#A	1	10.88	Fab - KTH	47,272.2 ± 0.4	47,270.9	27
			Fab D0	47,638.6 ± 0.3	47,637.1	32
			Fab D1	48,190.8 ± 0.3	48,189.1	35
	2	13.08	Lc	23,439.5 ± 0.2	23,439.0	22
			Fd + 93	24,292.3 ± 0.5	24,293.1	33
	3	13.37	Lc+ 457 Da	23,896.6 ± 0.4	23,896.0	24
			Fd + 457 Da	24,657.6 ± 0.3	24,657.1	19
Fab#B	1	10.88	Fab - KTH	47,272.0 ± 0.6	47,270.9	24
			Fab D0	47,638.8 ± 0.3	47,637.1	36
			Fab D1	48,190.8 ± 0.3	48,189.1	35
	2	13.08	Lc	23,439.6 ± 0.2	23,439.0	24
			Fd + 93	24,292.6 ± 0.5	24,293.1	20
	3	13.37	Lc + 457 Da	23,896.7 ± 0.4	23,896.0	28
			Fd + 457 Da	24,657.7 ± 0.5	24,657.1	26
Intact mAb						
	Peak	**Observed rt(min)**	**Species**	**Observed mass (Da)**	**Expected mass (Da)**	**Mass accuracy (ppm)**
Trastuzumab	1	8.50	mAb D0	145,867.2 ± 0.4	145,865.2	13
Rebridged trastuzumab	1	8.62	mAb D1	146,423.1 ± 3.9	146,417.2	41
			mAb D2	146,972.9 ± 0.8	146,969.2	25
			mAb D3	147,527.2 ± 1.1	147,521.2	41
			mAb D4	148,080.8 ± 2.0	148,073.2	51
			mAb D1-2Lc	99,543.1 ± 0.7	99,545.1	20
			mAb D1-(Lc+457 Da)	123,072.7 ± 1.5	/	/
	2	9.58	mAb/2 D1	73,486.5 ± 0.7	73,486.6	2
			mAb/2 D2	74,038.7 ± 0.5	74,038.6	1
	3	10.50	mAb/2 D2-(LC+457 Da)	50,140.9 ± 1.1	/	/
	4	11.01	2Lc D0	46,877.0 ± 0.6	46,878.1	24
			2Lc D1	47,429.1 ± 1.0	47,428.1	22
	5	13.10	Lc	23,439.3 ± 0.3	23,439.0	10
	6	13.52	Lc+457 Da	23,896.4 ± 0.3	23,896.0	16

4. Conclusions

In this chapter, we demonstrate the complementarity of LC-MS methods to decipher the challenges of rebriged mAb-based products characterization. In one hand, SEC-nMS allow a straightforward characterization of rebriged species providing an information about the avDAR and the DLD, yet fails

to distinguish between covalently rebriged and non-covalently maintained species. On the other hand, classical RPLC-MS in this case, completes the information by differentiating between fully rebriged mAb/Fab and rebridging by-products (mAb/2, Lc, Hc, Fd). However, the species with similar hydrophobicities tend to co-elute which hampers their identification. Consequently, our developed dSEC-MS method offers an all-in-one avDAR, DLD and rebridging efficiency assessment, thanks to the combination of SEC capabilities to separate size variants and denaturing conditions to maintain only strong covalent interactions. Moreover, with this study we demonstrate the benefits of cutting edge bio-inert SEC columns, along with the compatibility of the developed method with the benchtop LC-MS BioAccord system.

5. Scientific communications

Benazza R., Koutsoupetras I., Chaubet G., Hernandez-Alba O. and Cianférani S. **SEC-MS in denaturing conditions (dSEC-MS) for in-depth analysis of rebriged mAbs**. *Oral communication*. *Journées GDR ChemBio.* June **2023**, Strasbourg, France.

Benazza R., Hernandez-Alba O. and Cianférani S. **SEC-MS in Denaturing Conditions (dSEC-MS) for Rapid In-Depth Analysis of Rebridged Monoclonal Antibody-Based Formats**. *Article submitted to Talanta*, **2023**.

Koutsopetras I., A. K. Mishra, Benazza R., Hernandez-Alba O., Cianférani S., Chaubet G., Nicolai S. and Waser J.. Cysteine-**Cysteine Cross-Conjugation of both Peptides and Proteins with a Bifunctional Hypervalent Iodine-Electrophilic Reagent**. *ChemEurJ*, **2023**.

Conclusions of Part II

In the present part, the various uses of the benchtop BioAccord LC-MS system are discussed. This provides an overview of the benefits and limitations of this instrument through the characterization of newly developed mAb-based formats.

The coupling of non-denaturing liquid chromatography to native mass spectrometry has been widely used in the context of mAb-based formats characterization, mainly those stemming from optimal reaction conditions. However, a certain amount of expertise is required to perform those experiments, not to mention the time-consuming optimization process of the different LC and MS parameters. Therefore, monitoring the products stemming from a conjugation reaction of a specific mAb/protein in development is a real challenge in this case. The BioAccord LC-MS system is a self-calibrated and user-friendly benchtop system, which makes it convenient for analysis and data processing method automation. In a first chapter, I have demonstrated the possibility to perform SEC-nMS and CEX-nMS experiments through reference mAbs characterization. The SEC-nMS experiments allowed demonstrating the robustness, sensitivity and reproducibility of the instrument. Although the ability to perform CEX-nMS experiments is demonstrated, the reproducibility of this workflow could not be proven now.

Upon the high throughput mAbs characterization using the automated SEC-nMS method on the BioAccord, the same method was applied to characterize a large number of conjugates obtained through Ugi conjugation reaction. The results highlighted in the second chapter, allowed screening multiple reaction parameters to give the best avDAR and DLD values which highly suggests the implementation of this automated SEC-nMS workflow in biopharma laboratories. Albeit, when analyzing smaller proteins in native conditions such as Anticalins, the denatured form of the protein is observed due to the low pressure in the source that is a fixed parameter. This limitation could be tackled by inserting a modulating system at the primary pump level, to regulate the pressure in the interface region.

Bispecific antibodies (bsAb) are also mAb formats that draw the attention of biopharma companies for their use in cancer therapy; I wanted therefore to evaluate the adaptability of the BioAccord for these systems characterization through SEC-nMS analysis. In the third chapter, I showed one of the main drawbacks of the time of flight MS on the BioAccord which is the limited m/z range that reaches up to 7,000 m/z only, where we barely start to observe the first charge states of the bsAb 1:2 form. Later on, the same experiments on the Synapt G2 reveal the formation of 1:2 bsAb forms with hallmarks of free species and 1:1 bsAb forms. SEC-UV here is a powerful tool to quantify relatively each species and give an idea on the efficiency of the reaction, and native MS allow the precise mass measurement of the products while preserving the non-covalent interactions. Moreover, I have evaluated MP, which

requires far less quantity and time of analysis, while obtaining similar results from those of SEC-nMS. Albeit the mass resolution is lower than in mass spectrometry, in my opinion, it is the best technique to screen the reaction conditions to confirm the bsAb formation, before confirming its exact form with more resolutive MS experiments.

The final chapter emphasizes the ability of the BioAccord LC-MS system to perform method development while comparing with other methods, in a short amount of time. Herein, I presented a new method (dSEC-method), this time under denaturing conditions, suitable for characterizing rebridged mAb-based formats. The screening of different parameters mainly the choice of the column allowed selecting the Maxpeak Protein SEC sub-3µm column (Waters) for its bio-inertness benefits. Moreover, the comparison of this hybrid technique with methods used routinely in biopharma (namely SEC-nMS and RPLC-MS), allowed to demonstrate its benefits for the assessment of CQAs through precise mass measurement. **Table 14** resumes the beforehand mentioned pros and cons of the BioAccord LC-MS system.

Table 14: table resuming the benefits and limitations of the BioAccord LC-MS system.

Pros ✓	Cons ✗
User-friendly (training in hours)	Many parameters couldn't be changed
Self-calibrated (once a week)	Unfolding of native structures due to low interface region pressure (Pi= 2 mbar)
High throughput SEC-nMS experiments of 10 mAbs in less than 1 hour	Limited m/z range until 7,000 m/z
High throughput CEX-MS experiments of 10 mAbs in less than 3 hours	CEX-MS results couldn't be benchmarked using different mAbs
Adapted for SEC-MS in native and denaturing conditions	

In fact, a considerable part of my experiments in this part, focused on the implementation of high-throughput SEC-nMS for the characterization of ADCs and PDCs. Simply because the latter is the go-to method to allow a prompt buffer exchange followed by precise mass measurement of mAbs/proteins. This allows us to confirm whether the drug conjugation occurred or not through avDAR and DLD assessment. However, to determine the specific site of conjugation, it would be necessary to go further and characterize the primary structure of the ADCs/PDCs through fragments characterization.

Part III: Top- and middle-down MS approaches for mAb-based products characterization

Chapter 1: Implementation of a middle-down MS approach for mAb characterization

1. Analytical context

The common structural MS approaches (such as Native SEC-MS, **Part II**), used for the characterization of mAb-based formats [120-122], provide us with exceptional information about the different variants (charge, size, and hydrophobic), the conformational heterogeneity, the dynamics, and the interaction between these proteins [17, 19, 121, 123, 125-127, 306]. However, it is still necessary to investigate deeper the protein sequence to gain information at the primary structure about the different modifications, and the localization of cargo molecules in the case of ADCs for instance. For this purpose, bottom up MS (BU-MS) approaches (*i.e.*, peptide mapping) have been developed for the characterization of released peptides upon an extensive proteolysis of mAbs, and ADCs [116, 307]. These LC-MS assays are always required by the FDA to ensure the safety and the efficacy of biotherapeutics before their market release since those modifications can alter the physicochemical properties of the proteins, and then their safety and efficacy. The time-consuming sample preparation and the introduction of artifacts upon enzymatic digestion are considered bottleneck of BU-MS approaches. Additionally, the difficulty to correlate the information obtained at the peptide level with that obtained at the intact level is a major drawback, especially for ADCs characterization where BU-MS fails to provide valuable information such as the avDAR and the DLD. All these reasons along with the fragmentation of the labile modifications (glycans, payloads… etc.), were the main reasons to give rise to alternative techniques such as top-down MS (TD-MS) [22]. Since TD-MS relies on the fragmentation of intact proteins, it allows the correlation between the intact mass of the protein (MS1) and its fragmentation spectrum (MS2). Therefore, the different sequence variants identified and characterized from fragment ions can be straight linked with the intact structure of the protein giving access to the spatial information of sequence modifications. Nevertheless, MS/MS fragmentation of intact proteins is not trivial for several reasons. On the one hand, the fragmentation efficacy decreases with the size of the protein, being particularly challenging when proteins with greater MW than 30 kDa are studied [23, 242]. On the other hand, fragmentation spectra are extremely complex and overcrowded, with multiple overlapping and unassigned fragments. For all these reasons, TD-MS experiments require versatile MS platforms with high-resolution and mass accuracy capabilities to, at least, partially overcome these limitations. The latter is the only approach that preserves the structural protein characteristics while providing specific fragments that correlate with the intact mass, which makes it a suitable approach for mAb-based formats identification. Reducing mAbs to small ~25 kDa subunits using a digestion in mild-conditions (middle-down MS, MD-MS)[25, 30, 119, 208, 219] could undoubtedly help achieving efficient fragmentation (~50% sequence coverage of the subunits using Electron-transfer dissociation (ETD) and

~70% when using Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)), yet 100% of the sequence coverage is never achieved (**Figure 69**). The present chapter represents an attempt to reach higher sequence coverage in single runs of reference mAb (trastuzumab) using existing dissociation techniques on a state-of-the-art mass spectrometer: Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific).

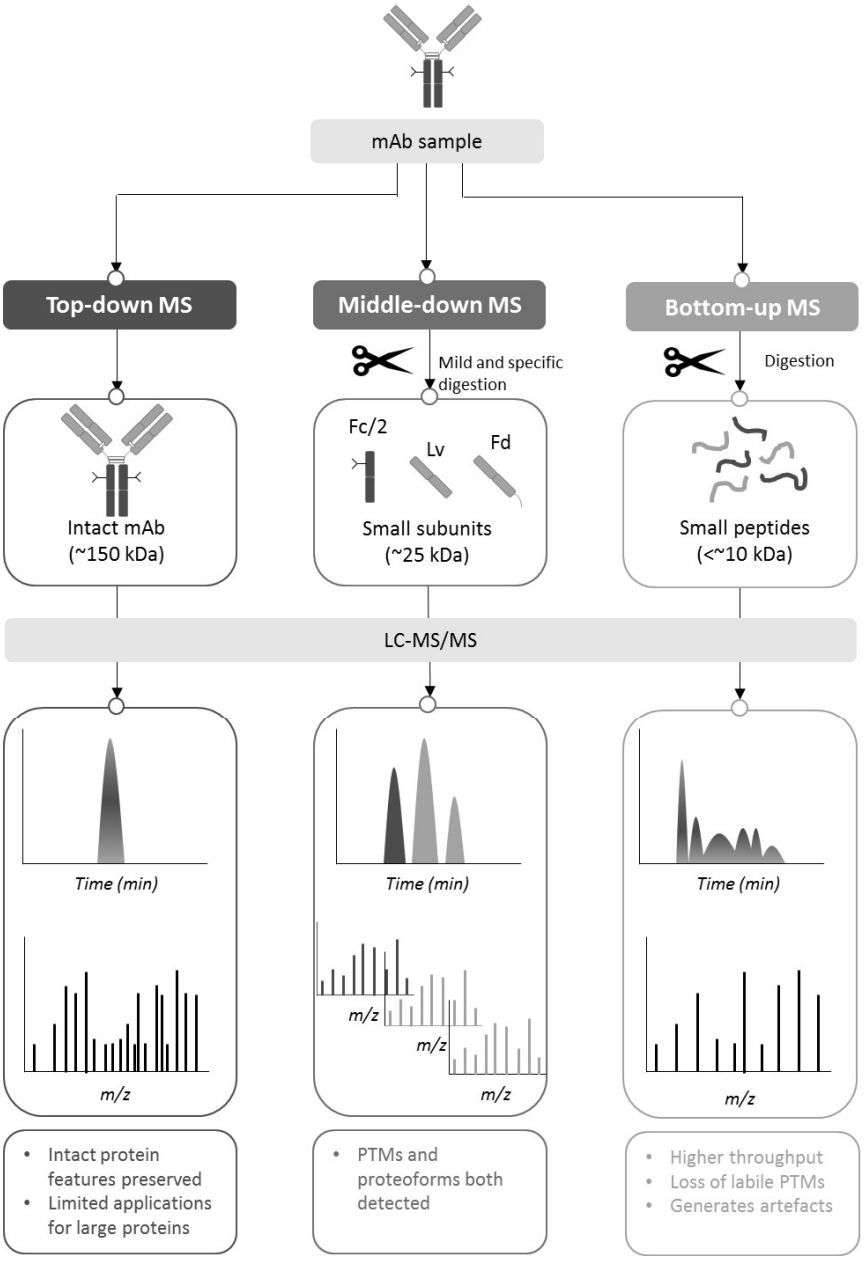


Figure 69: Representation of top-down (TD), middle-down (MD) and bottom-up (BU) mass spectrometry (MS) approaches for the characterization of monoclonal antibody (mAb). The first upper panels describe the no sample preparation required for TD-MS, the controlled and limited sample preparation performed for MD-MS analysis of small ~25 kDa subunits (Fc/2, Lc and Fd) or the extensive sample preparation needed for BU-MS experiments of digested small peptides. The middle panel depicts the LC-MS/MS analysis of the resulted samples and the lower panel cites the main advantages and/or limitations of each approach.

2. State of the art

The analysis of large proteins followed by their direct MS/MS fragmentation in the gas phase, to yield the molecular mass of the protein as well as the detection of PTMs, was first reported by Mclafferty *et al.* (1993). However, the first full article to refer to this technique as a "top down" MS approach, to distinguish it from bottom-up proteomics, was described in 1999 [20], and since then it gained tremendous interest in the MS field for protein characterization. These two steps techniques, resumed in a MS1 (or simply MS) scan to provide intact mass measurement followed by an MS2 (or MS^n) step to confirm the sequence, and characterize the sequence modifications of the selected precursor ion, is complicated, especially for "relatively" large proteins (greater than 30 kDa [23-24]). Due to the higher mass of large proteins, their degree of freedom increases as well and they can accommodate more internal energy before undergo fragmentation. For these cases, the use of different fragmentation parameters along with the use of complementary activation techniques is strongly recommended to improve the characterization of the primary structure of proteins. Although different activation techniques can be used to induce the fragmentation of molecular ions, the most popular fragmentations in TD-MS approaches are collisional activation methods (Collision-induced dissociation (CID) [227, 308] Higher-energy collisional dissociation (HCD) [309-310]) and electron-based dissociation techniques such as electron-transfer dissociation (ETD) [23, 34, 205, 207-208, 219] and Electron-capture dissociation (ECD) [27, 218, 223, 261, 311]). Photo-dissociation approaches are also used such as Ultraviolet Photodissociation (UVPD) with different wavelengths [35, 251, 312] and infrared multiphoton dissociation (IRMPD) [245, 313]. Electron-driven and photo-dissociation approaches are well known to provide extended protein backbone fragmentation with higher sequence coverage compared to slow-heating activation techniques such as CID or HCD. Furthermore, these techniques have been reported to maintain labile PTMs while improving the fragmentation of proteins with intra-chain disulfide bridges. More recently, the use of photon-based techniques has been reported as promising, in particular UVPD [35, 227, 247, 249-251, 312, 314-315] at 213 nm and 193 nm that generates a considerable amount of fragment ion types leading to a more comprehensive characterization of proteoforms [251].

MAbs and derived proteins have been also subjected to TD-MS experiments. In these particular cases, different fragmentation techniques have been included along with different experimental parameters in order to boost the overall sequence coverage. [26, 205, 208, 218-219, 222]. In fact, mAbs and ADCs are a real challenge to fragment due to their high MW, and presence of multiple disulfide bridges that narrows the fragmentation efficiency even with a combination of dissociation techniques. Thereby, more efficient results with almost complete sequence coverage have been reported by using middle-down MS strategies (MD-MS) [30, 219]. In this context, mAbs and ADCs are previously digested with enzymes that degrade the mAb scaffold near the hinge region [316-318]. Normally, after this digestion, a reduction

step is performed to avoid the presence of inter- and intra-chain disulfide bridges, and thus improving the fragmentation efficiency of the ~25 kDa subunits. Under these conditions, sequence coverages between 60-65% have been reported with electron-driven, photo-dissociation techniques, and combining different activation techniques, respectively [25, 28-30, 119, 208]. A recent study conducted by Dhenin *et al.* (2023) [30] highlighted the possibility of reaching near higher sequence coverage of mAb products with a combination of experiments, thanks to fine tuning of several parameters. Beyond the overall sequence coverage, TD-MS, and MD-MS experiments have proven their utility to localize the position of several mAbs PTMs such as glycoforms [25, 28, 128, 160, 207-208, 219, 228, 319]. In addition, some groups have focused their efforts in the characterization of disulfide bridges [129, 205, 207, 223] when dealing with mAbs or ADCs without any reduction step, either by managing to fragment the disulfide bridge and therefore determining the involved cysteines, or by detecting fragments containing two cysteines within a disulfide bridge upon UVPD fragmentation. Another main application of these approaches is to decipher the conjugation site of ADCs [28-29]. Here, the benefits of TD/MD-MS compared to BU approaches lies in the ability of proving the DLD and the avDAR, along with the localization of the drug site without the fragmentation of this latter. However, even at the subunit level (MD-MS), the maximum sequence coverage is never 100% (maximum ~74% using EThcD) [28, 219]. All the related TD/MD-MS literature illustrate the need of a thorough optimization step including different experimental parameters along with the use of alternative fragmentation techniques for an improved characterization of proteins. Hence, TD/MD-MS approaches ideally need to be developed on versatile, state-of-the-art mass spectrometry platforms with high-resolution capabilities.

3. Objectives

This chapter is dedicated to the optimizations of TD- and MD-MS approaches using the latest-generation Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific), which includes different fragmentation techniques (HCD, ETD, UVPD… etc.). In a first place, we aim at evaluating the instrument by optimizing the TD-MS parameters using reference proteins (Myoglobin, Carbonic anhydrase). Next, we intend for optimizing the MD-MS parameters for the characterization of a reference mAb sample (trastuzumab) with assessment of the N-glycosylation PTM.

4. Presentation of the Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific)

The Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) is one of the newest generation orbitrap instruments designed in 2019 for a large scale of applications including top-down proteomics. This mass spectrometer is composed of six different blocks (**Figure 70**):

- The ion source (Ion Max NG type) with adjustable heated electrospray ionization (HESI) probe for sample injection, with a section dedicated for fluoranthene anions (202.0 *m/z*) and

perfluoroperhydrophenanthrene (PPP, 623.9 *m/z*) ions for ETD and Proton-transfer charge reaction (PTCR), respectively.

- Ion optics with a series of lenses to direct ions and focalize their transmission to the quadrupole while preventing high-velocity clusters from entering.
- A quadrupole mass filter with ion transmission across 50-20,00 *m/z* range for efficient precursor selection with an isolation width of 0.4-1,200 *m/z*.
- An orbitrap (OT) mass analyzer with a resolving power of more than 500,000 FWHM and MS^n acquisition rate up to 40 Hz at 200 *m/z*.
- The ion-routine multipole (IRM) where HCD dissociation occurs with a variable pressure from 0.5 to 20 mTorr, this part ensures efficient ion transfer between the Orbitrap, the IRM and the linear ion trap (IT) analyzers.
- A modified dual pressure linear ion trap to enable MS^n for ion detection in both IT and OT mass analyzers. Multiple fragmentations can occur at this stage, including ETD, CID, EThcD and electron-transfer/collision-induced dissociation (ETciD) in addition to PTCR reaction that occurs in the high pressure cell. An extended front section is available where ETD and PTCR reactions happen. The low pressure cell is dedicated for UVPD activation with a 213 nm UVPD laser delivering 1.2 µJ/pulse at 2.5 kHz located at the back of the instrument.

The Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) could be easily switched to native conditions for the analysis of intact proteins, thanks to its capability to operate within lower IRM pressure values and larger *m/z* range.

All the different possibilities provided by the system open numerous options to conduct TD/MD-MS experiments, which means that the optimization of the experimental parameters will be critical for the characterization of proteins.

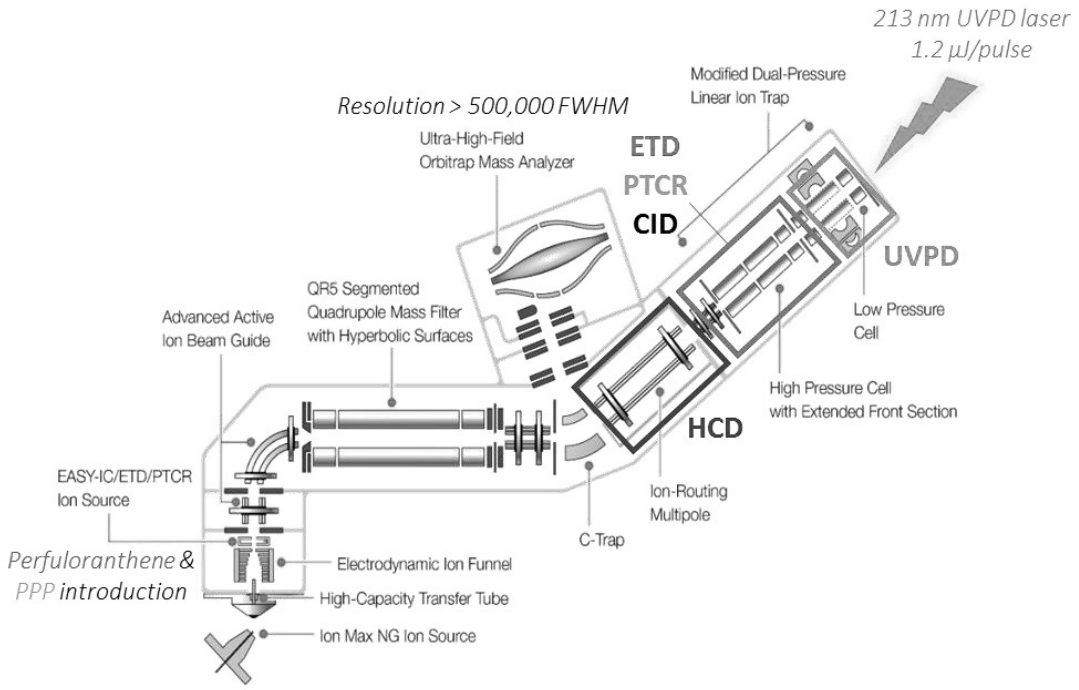


Figure 70: Schematic representation of the Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) used for the TD-MD/MS experiments. The scheme shows the ion path starting from the ion source, and passing through the ions optics to reach the quadrupole. The ions isolated in the quadrupole or in the ion trap are further fragmented in the ion-routine multipole when using HCD (blue) or in the dual linear trap when using ETD (red), PTCR (orange) or UVPD (green). The fragment ions and the precursor ions are detected further in the orbitrap for their mass analysis. The reagents used for ETD (red) and PTCR (orange) are depicted in the ion source with the corresponding color. The UVPD laser (green) is located at the back of the mass spectrometer.

5. Optimization of TD-MS parameters on reference proteins

In a first step, we aimed at evaluating the MS and MS^n parameters of the Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) using Myoglobin from equine heart as a reference protein. Due to its relatively small size (153 residues, 17 kDa) and its simplicity as it contains one single chain without any PTMS, it can be considered as the ideal protein for our TD-MS optimizations. Upon fine-tuning of MS parameters such as the resolution, the AGC target value, the maximum injection time… etc., the obtained MS spectrum revealed a distribution of more than 12 charge states allocated in the 600-2,000 *m/z* range (**Figure 71**).

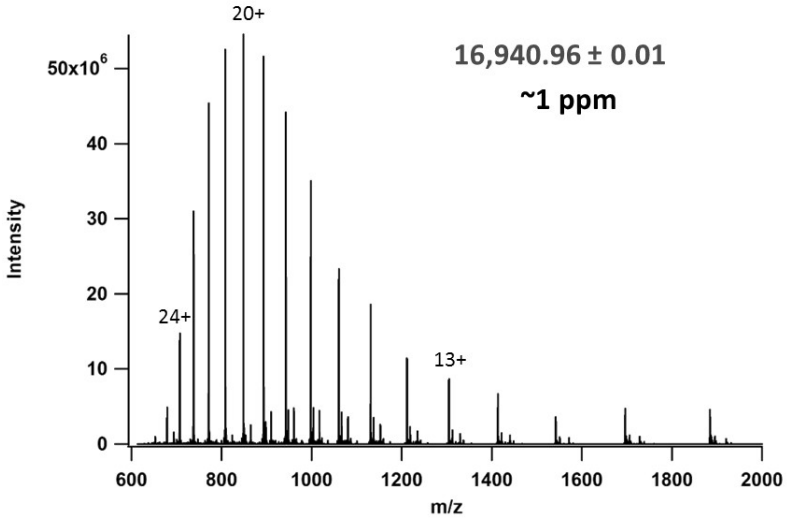


Figure 71: MS spectrum of Myoglobin obtained from direct infusion through an ESI source. The distribution is centered on the charge state 20+ which is the most intense ion, and the highly charged is the 24+.

5.1 Impact of the precursor ion

Taking into account that several fragmentation methods show a certain dependency on the charge and intensity of the precursor ion [227], we evaluated this parameter by selecting the most charged (24+, 707.33 *m/z*) and the most intense (20+, 848.50 *m/z*) precursor ions for ETD, HCD and UVPD fragmentation. **Figure 72** illustrates the MS/MS spectra stemming from both precursor ions fragmentation. HCD on the 24+ shows slightly more low abundant fragments compared to the 20+. On the other hand, ETD exhibited significantly different fragmentation patterns resulting in higher fragments number in the case of the 24+ fragmentation, whereas for the 20+ the charge reduction products were predominant in the MS/MS spectra. The inset illustrating UVPD fragmentation was very similar in the both cases. Unlike the other techniques, here the precursor ion remain very intense due to its incomplete depletion, resulting in low abundance product ions.

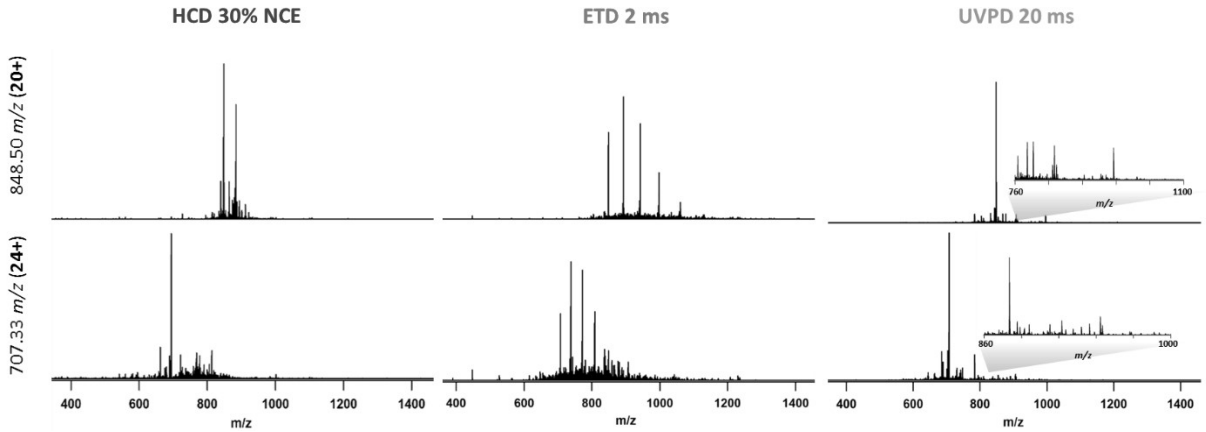


Figure 72: MS/MS spectra obtained upon TD-MS fragmentation of precursor ions 848.50 *m/z* charged 20+ (**upper panel**) and 707.33 *m/z* (**bottom panel**), using different fragmentation techniques: HCD (left, blue), ETD (middle, red) and UVPD (**right, green**).

The sequence coverage produced by HCD, ETD and UVPD varied also as the fragmentation patterns changed. Cleavage residue based on HCD was slightly enhanced when choosing the 20+ as a precursor ion, which increased the sequence coverage to 52% (*versus* 36% for the 24+), this shows a slight preference for the most intense precursor ion. Indeed, an intense precursor ion will have higher S/N ratio and will be result in distinguishable fragment ions products that are easily identified and thus contribute in increasing the sequence coverage. Interestingly, the sequence coverage upon ETD fragmentation jumped to 84% with a +13% increase when selecting the 24+ as a precursor ion, showing a clear preference for the highly charged ion, which corroborates the fact that electron-transfer is better achieved for highly protonated ions. Indeed, it was already demonstrated that the electron transfer cross section is proportional to the square of the charge [212, 233].

Unlike these typical fragmentations, UVPD exhibited no charge state dependence where the fragmentation remained similar in both cases (80% and 81% for 24+ and 20+, respectively), which highlights the statistical fragmentation this technique provides. Overall, our results pinpoint to the importance of the precursor ion's choice that is a crucial step before any efficient fragmentation.

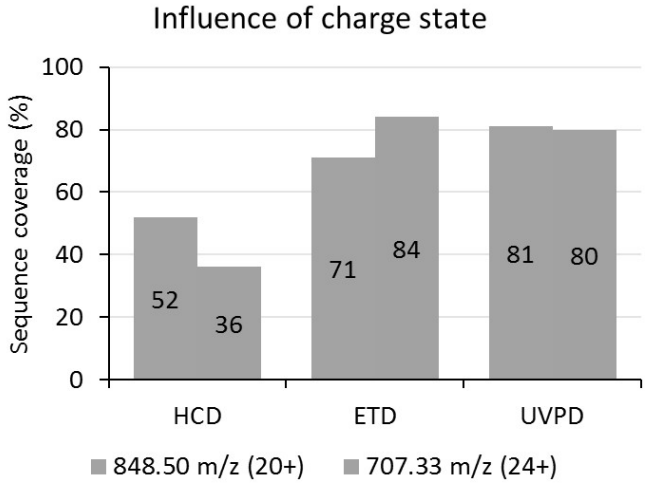


Figure 73: Sequence coverage obtained upon TD-MS fragmentation of precursor ions 20+ (**blue**) and 24+ (**red**) using different activation techniques: HCD, ETD and UVPD.

5.2 Impact of the fragmentation time/energy

Beyond the precursor ion's choice, the fragmentation energy and the activation time are crucial parameters to carefully consider before any TD-MS experiment. To emphasize this fact, we selected several dissociation energies for HCD (10% nce, 20% nce and 30% nce), reaction times for ETD (2ms, 4ms and 8ms) and irradiation times for UVPD (10ms, 20ms and 40ms) for the fragmentation of the 24+ charge state. As depicted in **Figure 74**, each technique exhibits significantly different fragmentation patterns, with a similar trend observed upon increasing the fragmentation time/energy. At lower HCD normalized collisional energy, the precursor ion remains highly abundant with no or very few

fragments observed, where increasing the energy increases the number of generated fragments. ETD at lower reaction times exhibits mainly highly abundant charge reduction products, with a non-neglected amount of low abundance fragment ions, which increased tremendously while increasing the reaction time. Unlike what the other techniques showed, in the case of UVPD the precursor ion shows less depletion even when increasing the irradiation time, which leads to generation of low abundant fragment ions compared to the precursor ion. Here, it is worth noting that the number of generated ions could not be directly aligned with the sequence coverage, as low mass and intensity fragments ions are more likely to be missed during deconvolution and matching with the amino acid sequence.

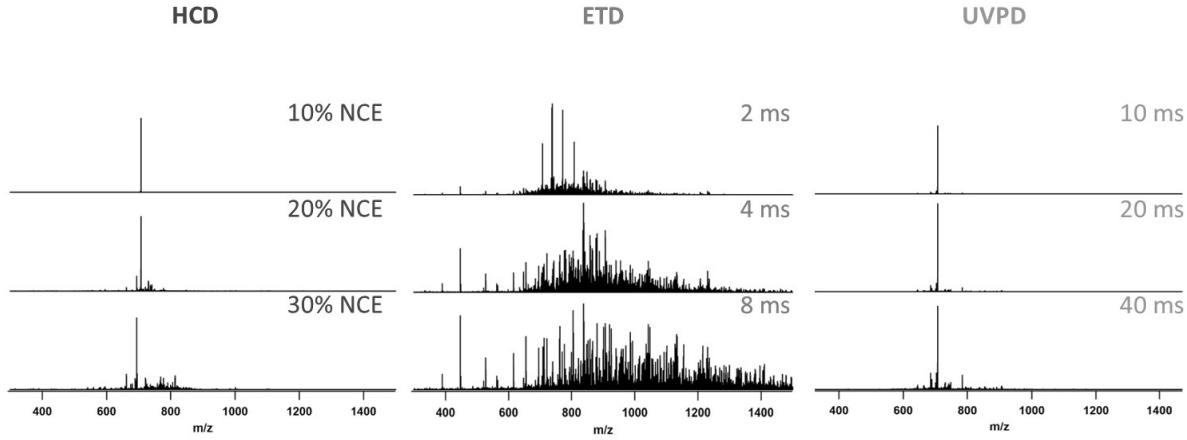


Figure 74: MS/MS spectra obtained upon TD-MS fragmentation of the 24+ charge state using different normalized collisional energies for HCD (**left**), different reaction times for ETD (**middle**) and different irradiation times for UVPD (**right**).

The sequence coverage results depicted in **Figure 75** show prominent variations upon the different activation techniques. HCD at lower energies exhibited less sequence coverage with the latter rising linearly with the dissociation energy (37% maximum sequence coverage with 20% nce). However, while low reaction times for ETD led to 84% sequence coverage, the extensive fragmentation that resulted in higher MS/MS product ions did not provide additional information. This could be due to the over-fragmentation of some ions of interest that lead to generating low abundance ions, which are difficult to deconvolute and identify or to generating internal fragments. Thanks to the different fragmentation and array of product ions obtained upon increasing the irradiation time for UVPD, we were able to rise the sequence coverage to 80% upon 40 ms irradiation. In fact, while UVPD exhibits no (or minimal) dependence on the charge state of the precursor ion, this particular technique shows significant variations upon changing the laser pulse. However, as mentioned above, in UVPD we do not expect complete precursor ion depletion thus upon increasing the irradiation time we will most probably induce secondary dissociations.

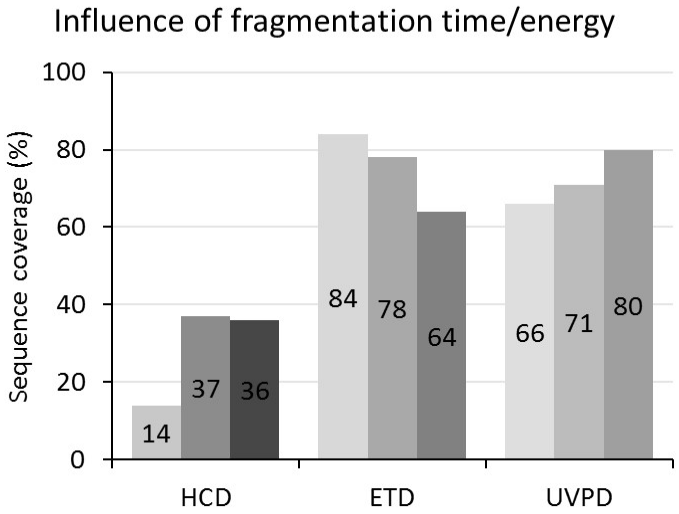


Figure 75: Sequence coverage of Myoglobin obtained upon the fragmentation of the 24+ using different energies for HCD (10% nce, 20% nce and 30% nce), different ETD reaction times (2ms, 4ms and 8ms) and different UVPD irradiation times (10ms, 20ms and 40ms).

5.3 Complementarity of the fragmentation techniques

The fragmentation patterns and the sequence coverage obtained upon each activation technique exhibit clear differences. Even if HCD provides very low cleavage residue, other MS/MS data obtained from ETD and UVPD compromise this. Indeed, after combining the results from the three experiments, the sequence coverage of Myoglobine rose to 99% (with one single missed cleavage), which highlights the complementarity of the different fragmentation techniques used. **Figure 76** emphasizes this deduction as it shows the distribution of the obtained fragment ions from each technique. While HCD provides fragments that are mainly located in the C-ter and N-ter regions of the protein, ETD exhibits a more rich fragment population located all along the protein backbone. Similarly to ETD, UVPD fragmentation results in fragments located along the protein sequence, with high abundant ones located in the central region with presence of multiple low abundance fragments.

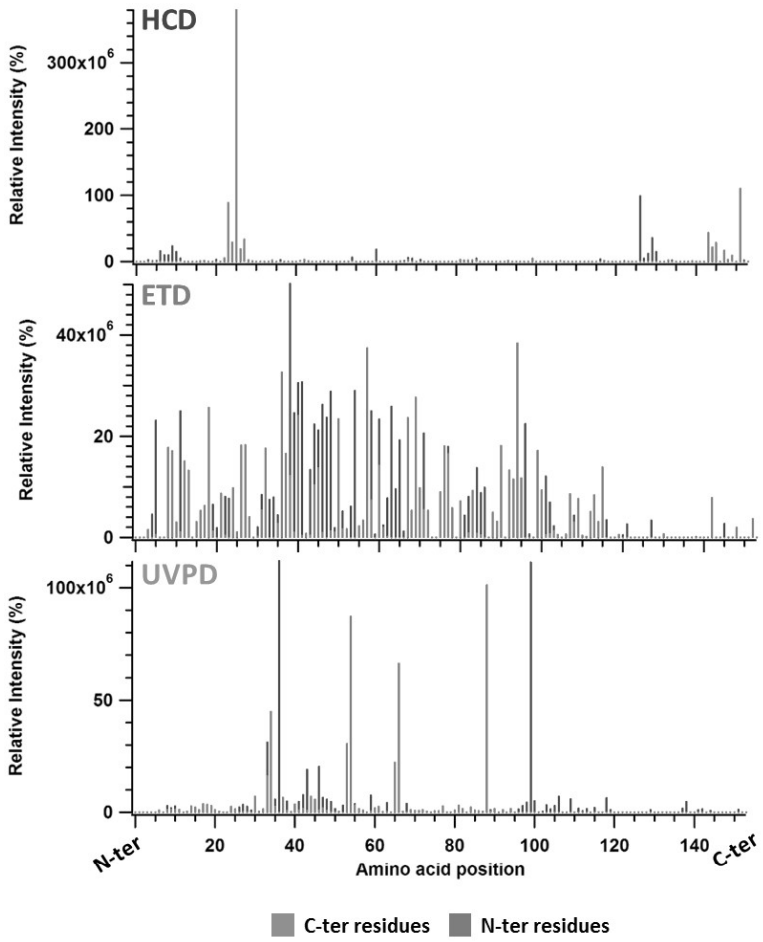


Figure 76: Distribution of C-terminal (red) and N-terminal (blue) residues along the Myoglobin amino acid sequence based upon different fragmentation techniques: HCD (top panel), ETD (middle panel) and UVPD (lower panel)

5.4 Top-down of larger molecular weight proteins

In fact, obtaining 99% sequence coverage is more likely for such a small protein as Myoglobin, due to the simplicity of the latter. However, larger molecular weight proteins are not expected to dissociate with similar efficiency. In this context, the same experiments were performed on Carbonic Anhydrase (29 kDa), which is a larger protein that contains 259 residues and one acetylation in its N-ter residue. As the graph in **Figure 77** shows, the difficulty of fragmentation of the largest protein are immediately highlighted. This difference is even more pronounced when using ETD and UVPD which means more reaction time is required for larger proteins. Of note, UVPD is expected to yield higher sequence coverage than HCD, however in this experiment we did not search for neutral losses and internal fragments that are often generated upon the fragmentation of large proteins. In addition, UVPD fragmentation efficiency is highly dependent on the laser condition. Overall, these results clearly show the impact of size and PTMs on the fragmentation efficiency, which means for our next mAb TD-MS experiments an extensive optimization of the MS/MS parameters is required to reach the highest sequence coverage along with PTMs assessment.

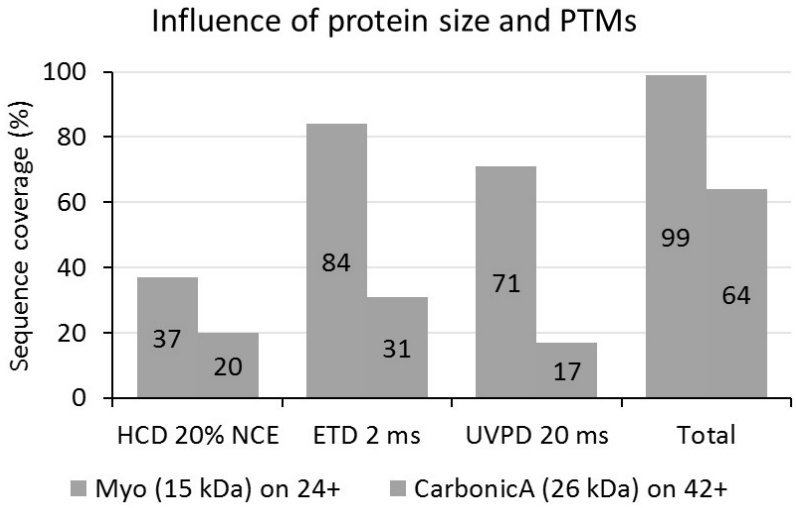


Figure 77: Sequence coverage obtained upon fragmentation of Myoglobin (blue) and Carbonic anhydrase (red) using HCD, ETD and UVPD in similar conditions. Total sequence coverage is depicted after combination of results obtained from the three experiments.

6. MD-MS workflow for a comprehensive characterization of trastuzumab

To downsize our intact trastuzumab of ~150 kDa to ~25 kDa subunits, we performed a minimal IdeS digestion followed by a DTT reduction prior to LC-MS analysis. One µg of the resulting sample was injected through a MAbPac RP 4 µm 1.0 x 1, mm column (Thermo Fisher Scientific) at a flowrate of 100 µL/min using a 10 min gradient. The RPLC step is mandatory in order to separate the different subunits based on their apparent hydrophobicity before to their injection into the mass spectrometer. The obtained chromatogram revealed three major species corresponding to the Fc/2, the Lc and the Fd subunits, respectively. The corresponding mass spectra revealed the masses of each species; the Fc/2 subunit exhibited a mass of 25,235.37 ± 0.01 Da for G0F/G0F glycoform centered on the 21+ charge state, in-line with this subunit bearing the N-glycosylation modification. A mass of 23,443.09 ± 0.01 Da was measured for the light chain with a distribution centered on the 20+ charge and the last species was attributed to the Fd subunit that showed a distribution centered on the 19+ with a mass of (25,383.57 ± 0.01 Da). Overall, our LC-MS method provides accurate mass measurements (~ 2.00 ppm) and confirmation of N-glycosylation of the Fc/2. (**Figure 78**)

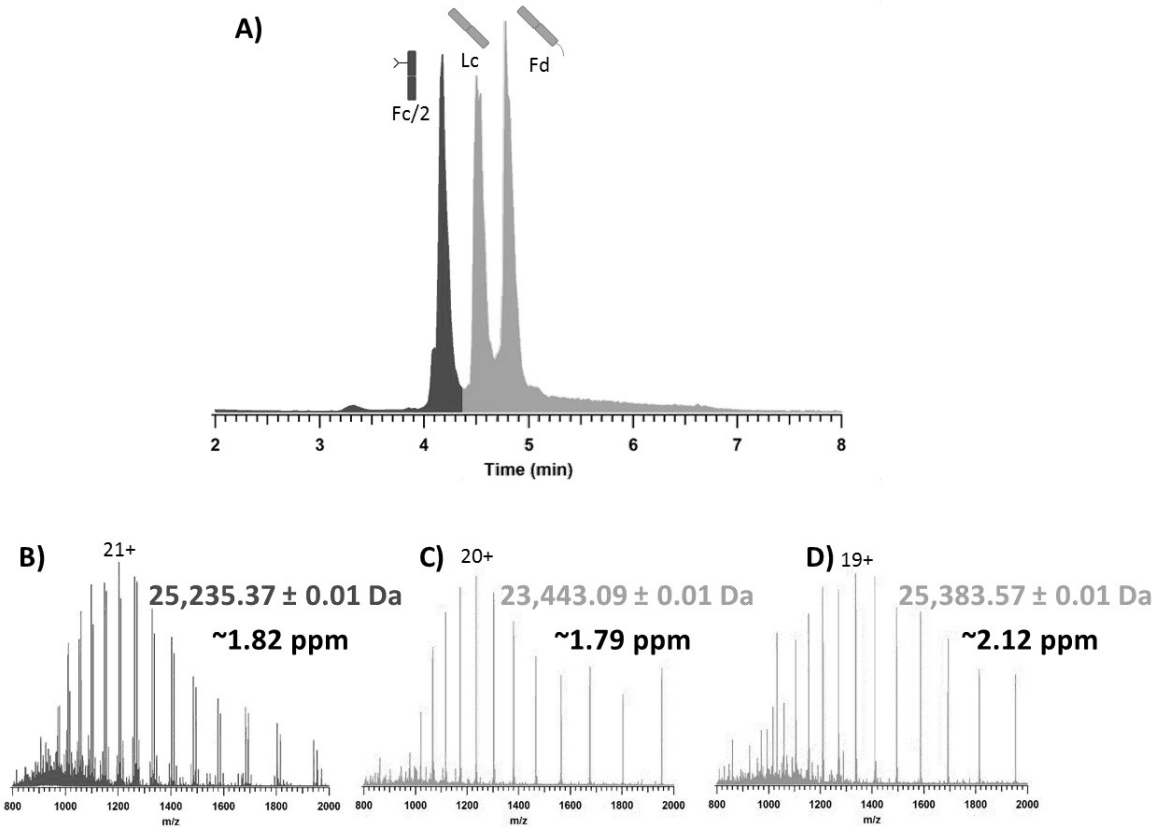


Figure 78: Middle-level RPLC-MS analysis of trastuzumab after Ides digestion and DTT reduction. (**A**) Total ion count (TIC) of the resulted subunits eluting in their apparent hydrophobicity order: Fc/2, Lc and Fd subunits (from the least to the most hydrophobic, respectively). The mass spectra and their corresponding measured masses are represented in the lower panel are obtained for (**B**) Fc/2 subunit centered on the 21+ charge state, (**C**) Lc subunit centered on the 20+ and (**D**) Fd subunit centered on the 19+. The depicted center charge states are the most intense.

With the aim of achieving maximum sequence coverage in single runs, along with the localization of the N-glycosylation site, several parameters were finely tuned namely:

- The fragmentation method along with dissociation energy or time
- In-source dissociation (iSD)
- Precursor ion choice, with or without multiplexing many precursor ions
- Size of the isolation window of the precursor ion

The resulted fragments were deconvoluted using Xtract algorithm *via* FreeStyle software v1.8.51.0, and the deconvoluted fragment ions were matched with the subunits sequences using ProSight Lite v1.4.

6.1 Fragmentation method

In view of the fact that the fragmentation methods available on the Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) provide each a variety of fragment ion types, we decided to focus on the evaluation of three main activation techniques namely HCD (providing *b/y* ions), ETD (providing *c/z* ions) and UVPD that exhibits a diversity of fragment ions among which *c/z*, *b/y* in addition to *a/x* ions.

a) Optimization of HCD normalized collisional energy (NCE)

Collisional dissociation-based methods are the most conventional techniques as they require relatively lower reaction times and are easily implemented on commercial mass spectrometers. In particular, HCD (20% NCE) outperformed classical CID (20% NCE) due to the variety of informative fragments it provides, thanks to the higher energy it applies in a short time, which access more dissociation pathways [313, 320]. To evaluate this technique on our instrument, we used a normalized collisional energy (NCE) by performing three experiments using 15% NCE, 20% NCE and 25% NCE. The results displayed in **Figure 79** show completely different fragmentation patterns when increasing the HCD NCE. While lower energies give fewer fragments, increasing the energy rises significantly the number of fragments. Interestingly, the three subunits respond in different ways as energy increases; the Fd seem to give rise to more fragments than the Fc/2 and the Lc at similar energies

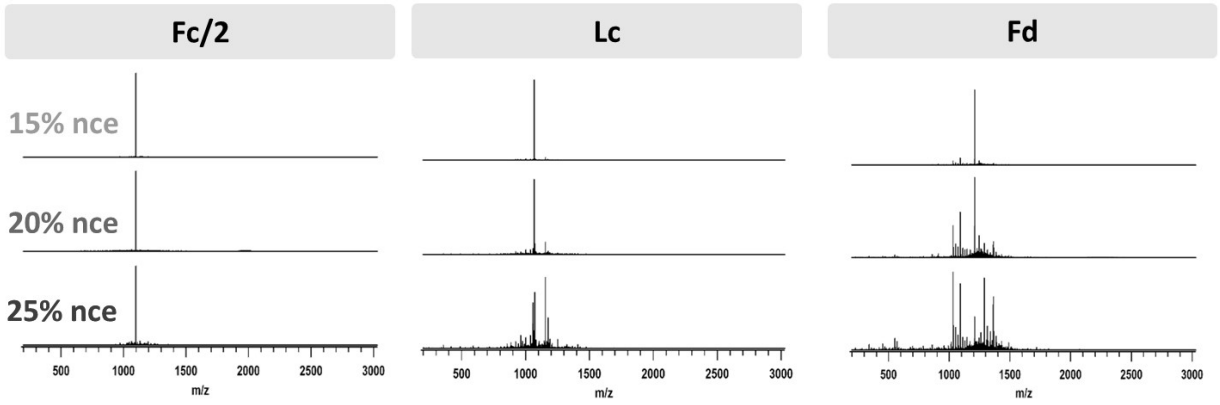


Figure 79: Optimization of HCD normalized collisional energy (nce). The Fc/2 (left), Lc (middle) and Fd (right) subunits are all subjected to different collisional energies: 15% nce (upper panel), 20% nce (middle pane) and 25% nce (lower panel). The experiments have been performed after selecting 4 most abundant precursor ions which namely the 20+, 21+, 22+ and 23+ for the Fc/2, 19+, 20+, 21+ and 22+ for the Lc and 18+, 19+, 20+ and 21+ for the Fd subunit. Each HCD energy shows a different fragmentation pattern with variable number of fragment ions.

In the present case, after matching the deconvoluted fragments to the subunits sequences, the results were linear. Increasing the energy to 25% NCE helped reaching the maximal sequence coverage values, with 13%, 26% and 21% for the Fc/2, the Lc and the Fd subunits, respectively (**Figure 80**). Of note, the localization of the N-glycosylation site was taken into consideration when assigning the obtained fragments to the Fc/2 amino acid sequence. This G0F glycan have already been studied in literature [321], and is expected to be located in the N61 position with an increment mass of +1,444.53 Da.

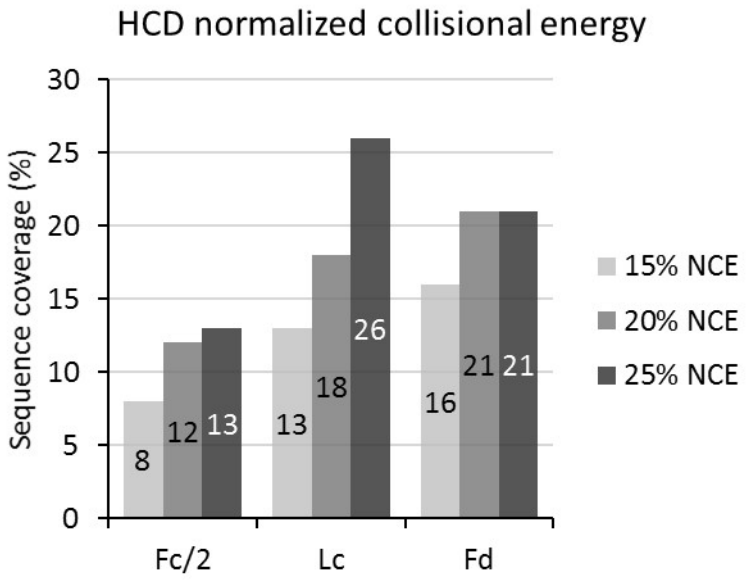


Figure 80: Impact of varying HCD normalized collisional energy (nce) on the sequence coverage of the trastuzumab subunits. The collision normalized energies are represented from lighter to darker blue for lower to higher reaction times, respectively. The sequence coverage value (%) is represented in the center of the bars. The bar plot reveal higher sequence coverage at 25% nce and conversely the lowest ones are observed at 15% nce.

b) Optimization of ETD reaction time

ETD has been reported as the go-to method for the dissociation of mAb subunits, as it often showed significant residue cleavage in short activation times (between 4-10 ms) by breaking the peptide backbone to generate complementary *c*- and *z*- type ions to those obtained from collisional techniques [28-29, 119, 222]. In this context, we evaluated three ETD ion-ion reaction times namely 2 ms, 4 ms and 10 ms for the fragmentation of the 4 most abundant precursor ions of each subunit (summarized in **Table 15**). **Figure 81** shows that the obtained MS/MS spectra at each reaction time were significantly different in term of fragmentation patterns, relative abundance of *c/z* fragment ions and background noise, with similar behavior of all three subunits contrary to what was observed upon HCD fragmentation. At lower reaction times (2 ms), few fragments were observed compared to higher reaction times (4 ms and 10 ms). At 2 ms, the most intense ions correspond to the charge-reduced species of the precursor ion. This phenomenon corroborates that few number of electrons have been transferred to the precursor ions, leading to the charge reduction of the overall net charge with a limited amount of fragment ions. This process is called ETnoD [322-323]. Conversely, when precursor ions from the different subunits are subjected to longer reaction times (10 ms) the charge-reduced species are no longer observed while the number of fragment ions appears to increase, leading to the conclusion that the ETD reaction is more efficient with longer reaction times, as expected. Those ions are the result of a series of re-fragmentation that reduces the size of fragments to smaller ones that

could probably not be matched with the subunits sequences. Similar behavior was noticed for the three subunits.

Table 15: The four selected precursor ions for each subunit (Fc/2, Lc and Fd) for ETD, HCD and UVPD fragmentations.

mAb subunit	Selected precursor ions
Fc/2	1,262.87 *m/z* (20+), 1,202.76 *m/z* (21+), 1,147.94 *m/z* (22+) and 1,098.05 *m/z* (23+)
Lc	1,234.77 *m/z* (19+), 1,173.06 *m/z* (20+), 1,117.22 *m/z* (21+) and 1,066.52 *m/z* (22+)
Fd	1,411.06 *m/z* (18+), 1,336.93 *m/z* (19+), 1,270.63 *m/z* (20+) and 1,209.63 *m/z* (21+)

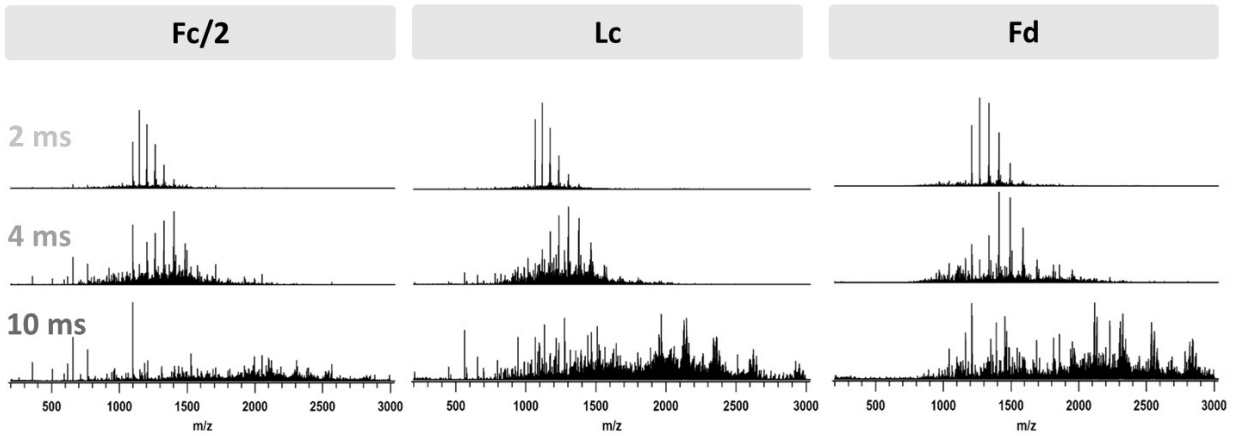


Figure 81: Optimization of ETD reaction time. The Fc/2 (left), Lc (middle) and Fd (right) subunits are all subjected to different electron-transfer reaction times namely: 2 ms (upper panel), 4 ms (middle pane) and 10 ms (lower panel). The experiments have been performed after selecting 4 most abundant precursor ions namely 20+, 21+, 22+ and 23+ for the Fc/2, 19+, 20+, 21+ and 22+ for the Lc and 18+, 19+, 20+ and 21+ for the Fd subunit. Each reaction time shows a different fragmentation pattern with different number of fragment ions.

The residue cleavage obtained upon matching the deconvoluted data to the subunits sequences, shows that the maximum sequence coverage (37%, 31% and 26% for Fc/2, Lc and Fd, respectively) was obtained at 4 ms (**Figure 82B**), with confirmation of the N-glycosylation site at the N61 position. The results corresponding to the longest reaction time (10 ms) did not provide an increase in the overall sequence coverage as expected, but a decrease even when comparing with the lowest dissociation time (2 ms). This is because the fragments were no longer informative as they are a result of over-fragmentation or internal fragments that are not assigned. However, at the lowest reaction time (2 ms) comparable sequence coverage than at 4 ms was obtained, especially in the case of the Lc subunit, which means that very few additional fragments from those deconvoluted, contributed to maximizing the sequence coverage while increasing the reaction time. The fragmentation map of the Fc/2 shown in **Figure 82A,** corroborates our hypothesis; when increasing the ETD reaction time we tend to generate smaller fragments located in the N-terminal region of the sequence. The latter fragments are

a result of consecutive dissociation of fragments of interest, which contribute to increasing the sequence coverage.

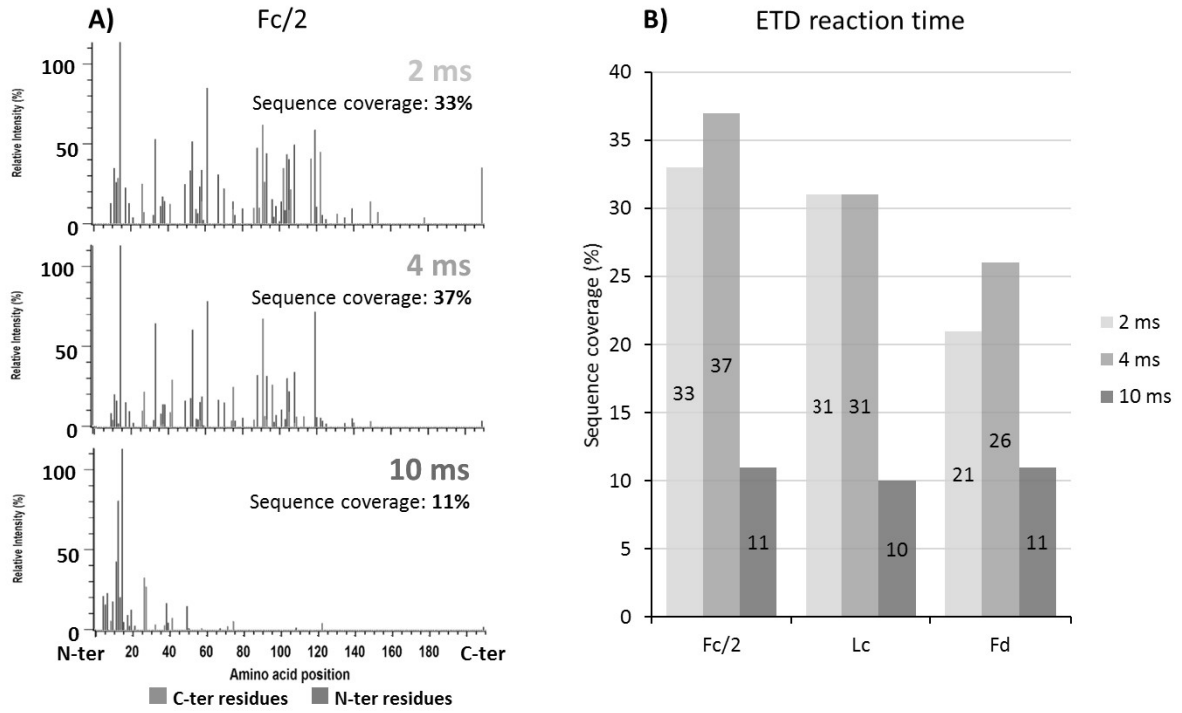


Figure 82: Impact of varying ETD reaction time on the sequence coverage of the trastuzumab subunits. (**A**) Impact of varying ETD reaction time on the distribution of fragments along the Fc/2 sequence. C-ter residues are represented in red and N-ter ones are in blue. (**B**) The reaction times are represented from lighter to darker red for lower to higher reaction times, respectively. The sequence coverage value (%) is represented in the center of the bars. The bar plot show higher sequence coverage at 4 ms reaction time, in contrast to 10 ms reaction time that exhibit lower sequence coverage.

Overall, these results clearly show that the optimization of ETD reaction time is at utmost importance to aim to increase the sequence coverage. The fragmentation trend showed that maximum sequence coverage is achieved between 2 and 4 ms, with 4 ms giving significantly higher sequence coverage for the Fc/2 and Fd subunit, along with assessment of the N-glycosylation site on the Fc/2 chain. In addition, the overall appearance of the MS/MS spectra should be interpreted carefully, as even when there is charge-reduction ions the sequence coverage is still significant. Altogether, these results enabled us to select 4 ms as an optimal reaction time for further optimizations.

c) Optimization of UVPD reaction time

With the same intention, we tested different reaction times for 213 nm UVPD to select the most suitable irradiation time. UVPD capabilities for covering more than 50% of proteins sequence have been demonstrated in several prior studies [28, 35, 222, 250, 312]. However, as each UVPD experiment shows a different cleavage propensity, many experiments are often combined to reach maximal sequence coverage values thanks to the combination of all obtained fragments. In our case, we wanted to understand more, which parameters affect this fragmentation starting by the UVPD reaction time.

Interestingly, the MS/MS spectra obtained upon UVPD of trastuzumab subunits were significantly different from the other techniques. Here, the precursor peaks remain intense despite the generation of a tremendous amount of fragment ions. This is because UVPD activates different types of fragmentation pathways resulting in many low-intense fragment ions compared to the precursor ion. This number of fragments rises with increasing the photoactivation time (from 346 generated fragments at 10 ms to 508 fragments at 20 ms). This boost is due to the several possible fragmentation pathways that UVPD enables, leading to a wide variety of fragment ions (*a/x*, *b/y* and *c/z*) (**Figure 83**).

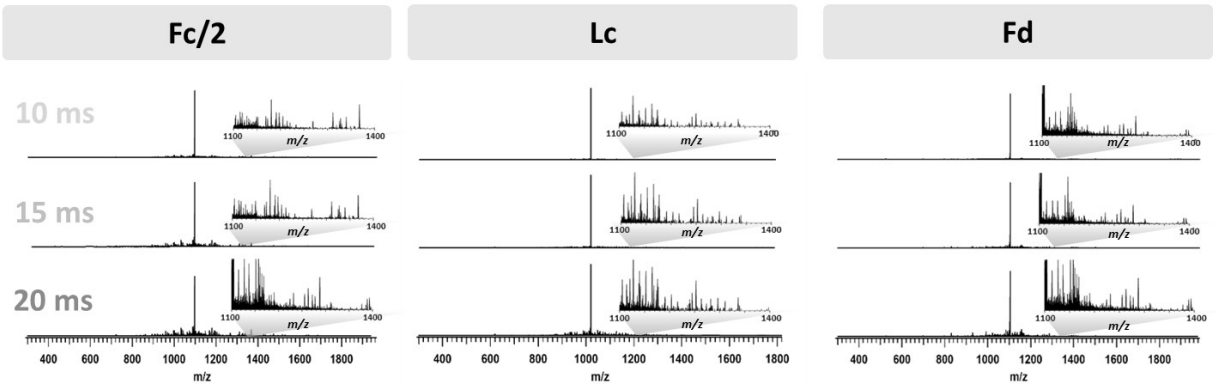


Figure 83: Optimization of UVPD activation time. The Fc/2 (left), Lc (middle) and Fd (right) subunits are all subjected to different UVPD photo-dissociation times: 10 ms (upper panel), 15 ms (middle pane) and 20 ms (lower panel). The experiments have been performed after selecting 4 most abundant precursor ions namely the 20+, 21+, 22+ and 23+ for the Fc/2, 19+, 20+, 21+ and 22+ for the Lc and 18+, 19+, 20+ and 21+ for the Fd subunit. Each UVPD activation time shows a different fragmentation pattern with variable number of fragment ions. Even at higher activation times, the MS/MS spectra show that the precursor ion is still most abundant.

As expected, increasing the irradiation time led to an improvement for the sequence coverage. The subunits behaved differently as for the Fc/2 the sequence coverage was plateaued leading to a maximum of 12%, considering the N-glycosylation is at the 61st amino acid position. In the case of the Lc, the maximum sequence coverage was obtained upon 15 ms UVPD with a value of 18%, while the Fd reached 29% of its cleavage residue at 20 ms UVPD (**Figure 84**). Surprisingly, these values were disappointing compared to the number of obtained UVPD fragment ions, simply because there is a considerable amount of the latter fragments that remained unexplained and thus do not contribute in confirming the amino acid sequence. In fact, at higher reaction times the number of explained fragments start to decrease (6%, 3% and 8% at 20 ms *versus* 9%, 12% and 9% at 15 ms for the Fc/2, the Lc and the Fd, respectively.). Similarly to ETD, these fragments are actually small internal ions that are a result of secondary dissociations of fragments of interest, which are challenging to assign by conventional softwares. In fact, in UVPD the fragment ions and the precursor ions are both irradiated at the same time, which increases the probability of over-fragmentation and thus the generation of internal fragments.

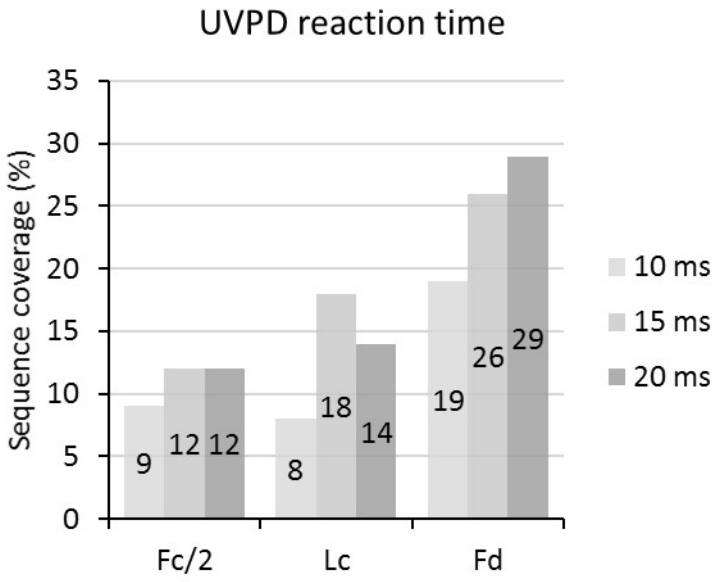


Figure 84: Impact of varying UVPD activation time on the sequence coverage of the trastuzumab subunits. The UVPD photo-dissociation times are represented from lighter to darker green for lower to higher reaction times, respectively. The sequence coverage value (%) is represented in the center of the bars. The bar plot reveal higher sequence coverage after 15 ms and 20 ms UVPD time and conversely the lowest sequence coverages are observed upon 10 ms UVPD.

d) Complementarity of the three fragmentation techniques (HCD,ETD and UVPD)

The generated fragmentation maps from each optimal condition (*i.e.* HCD 25% NCE, ETD 4 ms and UVPD 15 ms) showed different fragments location along the subunits sequence (**Figure 85**). Even if HCD gives low sequence coverage compared to ETD, it provided us with golden complementary pairs (up to 12 pair for the Fc/2) that helped confirming the amino acid sequence and thus increased the confidence in the N-glycosylation site assessment. These are pairs of *b/y* fragment ions that has been formed upon cleavage between the same pairs of amino acids. ETD gives the highest residue cleavage with fragments located all along the subunits backbones, as we previously observed for reference proteins, with a slight preference to cleave the N-terminal region of the mAb subunits. Despite the low sequence coverage obtained after 15 ms UVPD reaction, this technique provided us with unique fragments that were observed upon neither ETD nor HCD dissociation, which correspond to *a/x* fragment ions.

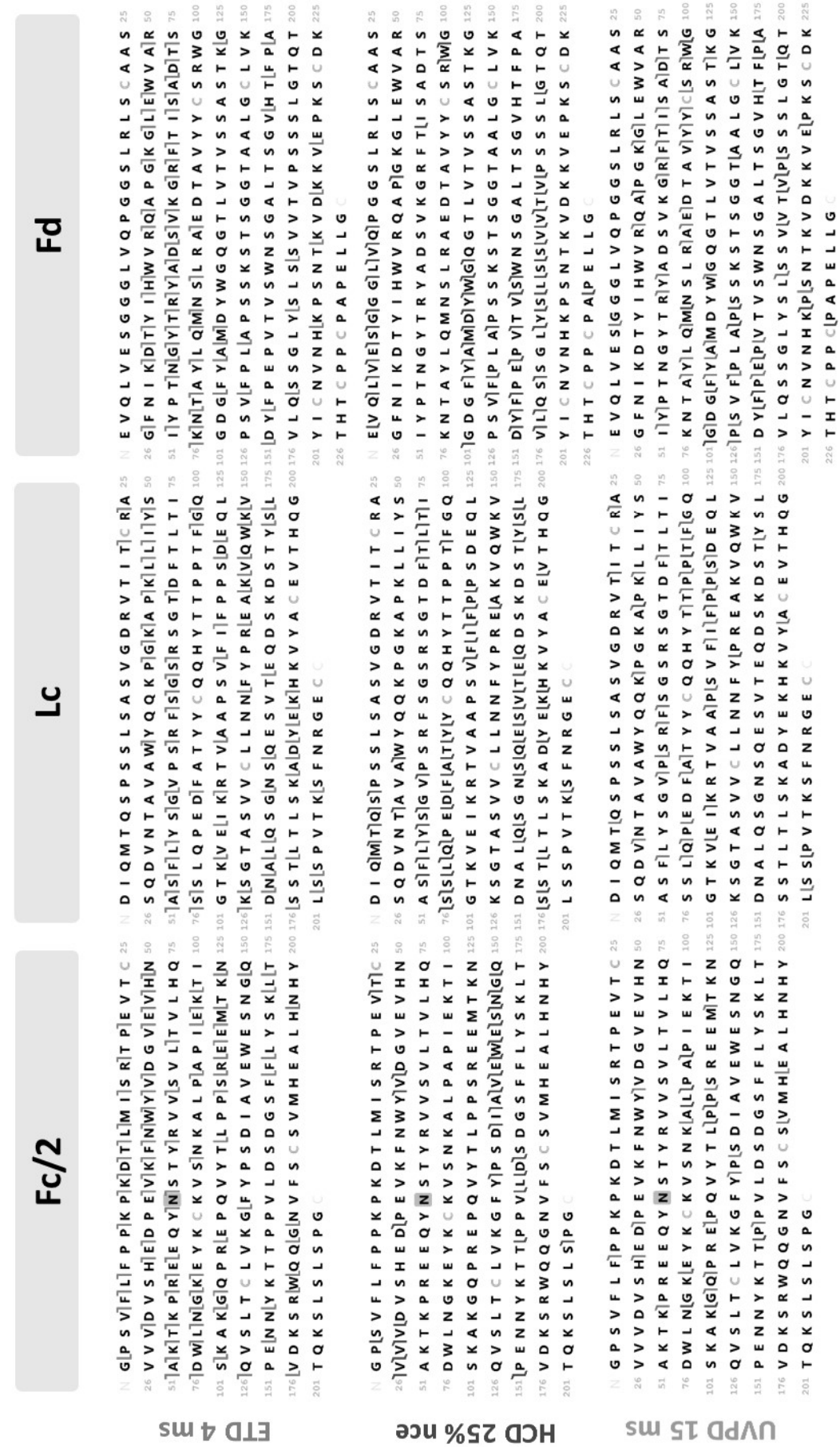


Figure 85: Fragmentation maps of Fc/2 (left), LC (middle) and Fd (right) subunits upon ETD 4 ms (top), HCD 25% nce (middle) and 15 ms 213 nm UVPD (bottom) fragmentations. Each type of fragment ions is depicted with a different color, *c/z* type ions obtained from ETD are in red, *b/y* ions obtained from HCD are in blue and finally *a/x* provided upon UVPD activation are in green. The N-glycosylation site is highlighted with an orange square.

Figure 86 demonstrates clearly that each technique contributes with different fragment patterns. Taking the Fc/2 as an example, 67, 20 and 14 fragments were obtained with ETD, HCD and UVPD individually, respectively. While many fragments were shared between two or the three techniques (3 fragments between ETD and HCD, 3 fragments between HCD and UVPD, 7 fragments between ETD and UVPD and one fragment between ETD, HCD and UVPD). This complementarity is highlighted by the combination of the fragments from the three experiments that rises the sequence coverage to 58%, 65% and 56% for the Fc/2, the Lc and the Fd subunits, respectively. This is less than what was obtained in a similar study aiming at fragmenting a mAb sample using a combination of different MS/MS runs (60-80% sequence coverage of the mAb subunits) [29]. Despite the fact that, in our study, we combine only three (~8min) LC-MS/MS runs which allow us to get a good subunits characterization rapidly compared to previous studies, we believe that we could still reach higher sequence coverage by explaining more the obtained fragments. Here, the results were obtained through the explanation of less than 50% of the obtained fragments, for each experiment. This could be due to many reasons: i) First, we only search for predominant ions, whereas neutral losses are not considered. ii) Second, many fragments could be attributed to the cleavage of the same pair of amino acids, therefore even if they increase the confidence in the cleavage site attribution, they do not contribute in enhancing the sequence coverage. iii) Finally, many generated fragments might be internal fragments, which are computationally difficult to assign.

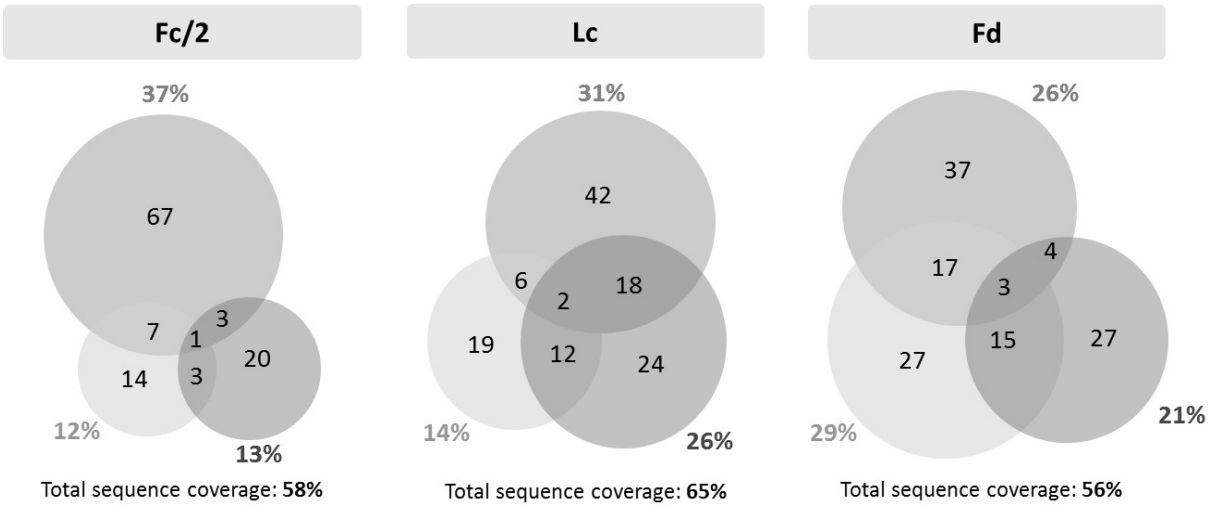


Figure 86: Venn diagram of trastuzumab subunits (Fc/2, Lc and Fd) fragmentation upon HCD (blue), ETD (red) and UVPD (green). The number of obtained fragments from each technique is indicated inside the circles and the sequence coverage (%) is depicted outside the circles with the corresponding color of each technique. The total sequence coverage indicated for each subunits is obtained upon combination of the fragment ions from each technique. Each LC-MS/MS run was performed in monoplicates.

Moreover, the combination of the fragments obtained upon the three activation techniques (HCD, ETD and UVPD) contributed highly in increasing the confidence in the assessment of the N-glycosylation

site. **Figure 87** show the sequence coverage and the number of specific fragments to the modification site, obtained upon variation of the N-glycosylation site over 10 possible asparagine sites (N40, N50, N61, N79, N89, N125, N148, N153, N154 and N185). As expected, the highest sequence coverage (58%) was observed when considering the N-glycosylation occurred at the N61 position. Conversely, the sequence coverage of the Fc/2 subunit significantly decreased when the N-glycosylation was assumed to occur at different asparagine residues. Additionally, the number of fragments specific to the modification site was significantly higher when considering the N-glycosylation occurred at the 61st position (51 fragments) compared to other sites (28-47 fragments).

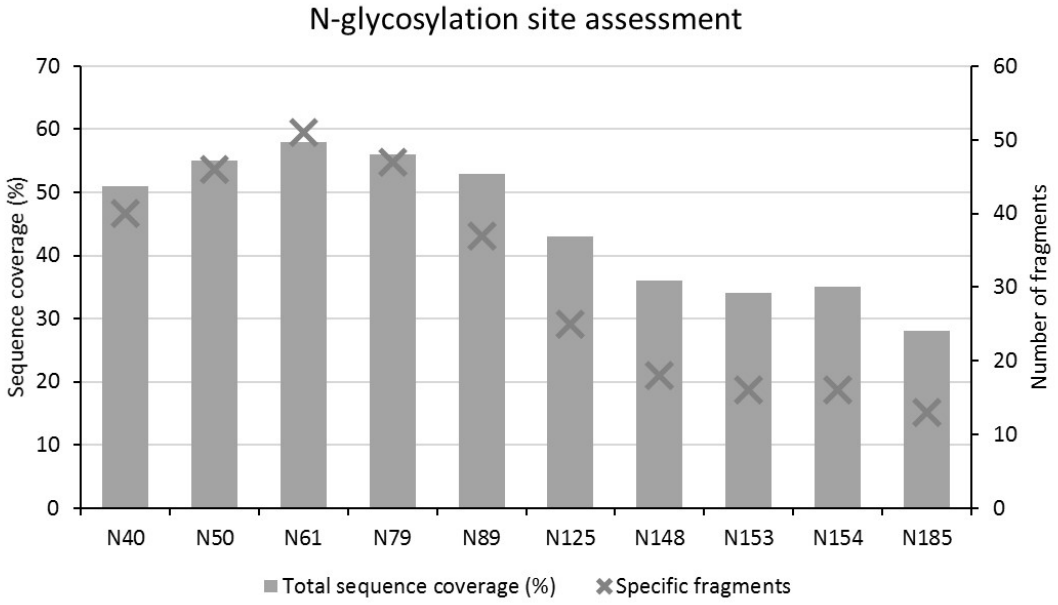


Figure 87: Total sequence coverage (%) of the Fc/2 subunit, and the number of specific fragments to the N-glycosylation site upon combination of HCD, ETD and UVPD fragments. The sequence coverage and the number of specific fragments were obtained upon variation of the glycosylation sites over 10 different asparagine sites (N40, N50, N61, N79, N89, N125, N148, N153, N154 and N185)

In summary, the combination of fragments upon HCD, ETD and UVPD activations shows a clear benefit of increasing the sequence coverage, which allow to increase the confidence in fragment ion assignment and thus to localize the modification different modifications (N-glycosylation).

6.2 Choice of precursor(s) ion(s)

The fragmentation efficiency is not only affected by the time or the energy dedicated for the activation reactions, in fact, one of the most important parameters to take into consideration is the choice of the precursor ion. Different studies showed that multiplexing several abundant charge states or selecting a large isolation window could remarkably increase the sequence coverage of mAb subunits [28, 30]. Our previous results performed on four consecutive charge states with an isolation window of 150 *m/z* showed overall good sequence coverage. We thought to maximize this sequence coverage by evaluating different precursor ion possibilities, either by selecting one abundant precursor ion, which

is well charged and intense enough to avoid bad quality MS/MS spectra, or 2 multiplexed ions in comparison to those fragmented beforehand, the precursor ions selected are summarized in **Table 16**.

Table 16: Summary of the selected precursor ions for ETD, HCD and UVPD fragmentation. For each subunit we selected 4 multiplexed ions, 2 multiplexed ions or one unique ions (from top to bottom).

mAb subunit	Selected precursor ions
Fc/2	1,262.87 *m/z* (20+), 1,202.76 *m/z* (21+), 1,147.94 *m/z* (22+) and 1,098.05 *m/z* (23+)
	1,147.94 *m/z* (22+) and 1,098.05 *m/z* (23+)
	1098.05 *m/z* (23+)
Lc	1,234.77 *m/z* (19+), 1,173.06 *m/z* (20+), 1,117.22 *m/z* (21+) and 1,066.52 *m/z* (22+)
	1,117.22 *m/z* (21+) and 1,066.52 *m/z* (22+)
	1,066.52 *m/z* (22+)
Fd	1,411.06 *m/z* (18+), 1,336.93 *m/z* (19+), 1,270.63 *m/z* (20+) and 1,209.63 *m/z* (21+)
	1,270.63 *m/z* (20+) and 1,209.63 *m/z* (21+)
	1,209.63 *m/z* (21+)

As depicted in **Figure 88**, each experiment led to a different residue cleavage. While multiplexing 2 or 4 ions led to lower sequence coverage, it seems that focusing the fragmentation on one ion at a time is the suitable option to achieve higher values. Taking as an example the UVPD fragmentation, we were able to reach up to 32% sequence coverage for the Fd subunit, conversely to 11% obtained *via* fragmenting 4 multiplexed ions. In some cases, similar results were obtained while multiplexing 2 ions or fragmenting only one (Fc/2 with 17% sequence coverage for both). Similar behavior was highlighted with ETD and HCD fragmentation, with only one slight difference at the Fc/2 subunit (maximum reached when multiplexing 2 ions, 15% sequence coverage). In fact, the width of the chromatographic peak limits the choice of the number of scans that can be acquired for each ion fragmentation. For example, when selecting one unique ion we are able to acquire up to 20 MS^2 scans, therefore we can average more spectra and obtain additional fragment ions with increasing S/N ratio. Whereas when selecting four ions, the total number of scans drops to only 5 scans per (10 scans/ion for 2 multiplexed ions) leading to a decrease of the averaged spectra with low number of fragment ions. Consequently, the relative intensity of some fragment ions could be biased leading to their miss-identification. Therefore, selecting one precursor ion for its fragmentation seems to be the most appropriate strategy to be sure to include the maximum fragment ions obtained.

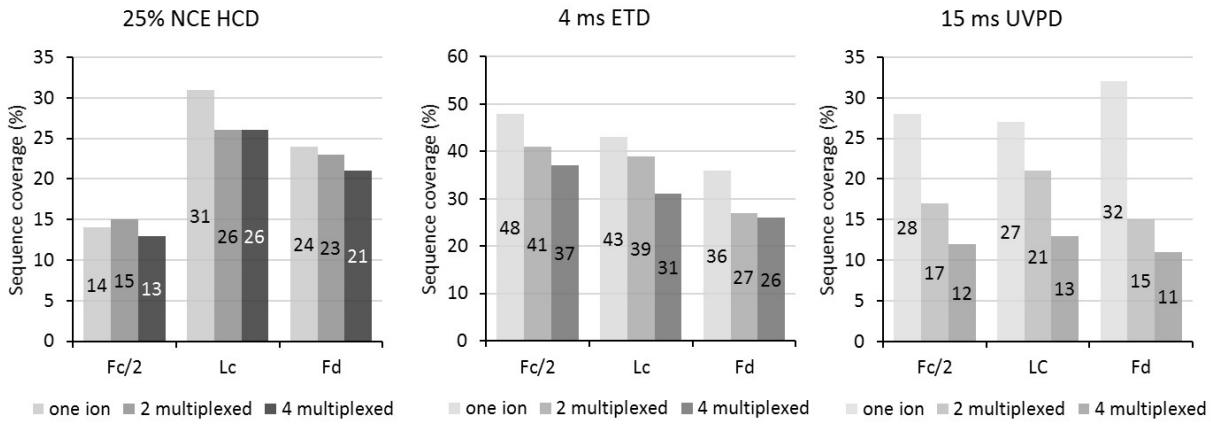


Figure 88: Impact of choosing one, 2 or 4 precursor ions for MS/MS fragmentation using HCD, ETD and UVPD. Residue cleavage obtained for trastuzumab subunits upon 25% nce HCD (blue), upon 4 ms ETD (red) and upon 15 ms 213 nm UVPD (green). The gradient of color represent the number of multiplexed ions, from lighter to darker when choosing from one to 4 precursor ions. The best results are obtained when selecting only one precursor ion. For the fragmentation of a unique precursor ion we selected the 23+, 22+ and 21+ for the Fc/2, the Lc and the Fd, respectively. For 2 multiplexed ions we added to the previously selected ions, the 22+, 21+ and 20+ for the Fc/2, the Lc and the Fd, respectively.

Moreover, to understand better the propensity of precursor ions fragmentation, we compared the highly charged precursor ion (previously compared with the 2 and 4 multiplexed ions) with the most intense one. The results revealed in the **Figure 89** an overall decrease of sequence coverage when attempting to fragment the most intense ions in all cases. Despite the fact that the most intense ions seem easily accessible in term of abundance and isotopic distribution, increasing the sequence coverage does not only rely on the intensity of the precursor ion. As expected, ETD is the most impacted by the charge state with the electron transfer that increases proportionally to the charge of the precursor ion. Moreover, there are multiple sites accessible for protonation, mainly resulting from cleavage of covalent bonds additionally to disruptions of weak non-covalent interactions. HCD, on the other hand, depends on the mobility of protons, thus the more protons available, the higher the probability of triggering fragmentation at various sites. In contrast, UVPD fragmentation exhibits less dependence on the precursor ion charge state, unlike HCD and ETD. Overall, the fragmentation of the most charged precursor ion was the best in our case, and yielded to 57%, 65% and 56% total sequence coverage for the Fc/2, the Lc and the Fd subunits, respectively after the combination of the three techniques.

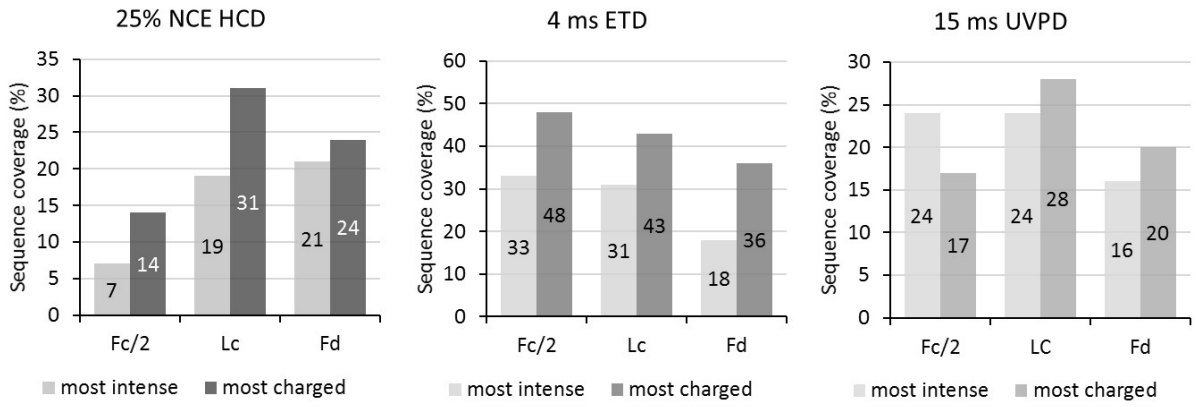


Figure 89: Impact of intensity and charge state of the selected precursor ion for MS/MS fragmentation using HCD, ETD and UVPD. Residue cleavage obtained for trastuzumab subunits upon 25% nce HCD (blue), upon 4 ms ETD (red) and upon 15 ms 213 nm UVPD (green). The lighter colors represent the most intense precursor ions (20+, 19+ and 18+ for the Fc/2, the Lc and the Fd, respectively.), while the darkest color depict the highly charged precursor ions (23+, 22+ and 21+ for the Fc/2, the Lc and the Fd, respectively.). The best results were obtained when selecting the most charged precursor ion with a decent abundance.

These results pinpoint to the fact that the number of precursor ions to be selected is highly dependent on the protein structure (hydrophobicity, size, number of modifications…). As every subunits releases a different charge state envelope (See **Figure 3**), the most convenient precursor ion could be either selected alone, or could be a part of the 2 or 4 multiplexed ions. Moreover, the charge state is a crucial parameter to be considered, in our case the used fragmentation techniques were leaning to the choice of the most charged ions for ETD and HCD to be fragmented instead of the most intense ones, with comparable results upon UVPD fragmentation. Thus, we recommend highly assessing this parameter during TD-MS experiments.

6.3 Isolation window width impact (IW)

In addition to all the parameters evaluated previously, the choice of the isolation window must be assessed. All the previous experiments were carried out using a 2 *m/z* window for the precursor ion's isolation, thus we wanted to evaluate a wider isolation window of 150 *m/z* to understand its impact. In fact, when using a wider isolation window, we are indirectly selecting several ions instead of one, thus the same trend than multiplexing could be observed. This suggestion is supported by the results shown in **Figure 90**. On the one hand a higher sequence coverage is observed when using a narrow isolation window of 2 *m/z*. On the other hand, broadening the window caused a significant drop in sequence coverage especially upon HCD (-16% and -17% for the Lc and Fd subunits, respectively) and UVPD (-12% and -24% for the Lc and Fd subunits, respectively) fragmentation, with an exception for the Fc/2 that increased for HCD by +7% and that decreased slightly for UVPD (-3%) upon choosing 150 *m/z* as an isolation window value. For ETD, the benefits of choosing a narrower window was also highlighted whereas a drop of -10%, -7% and -7% was noted for the Fc/2, the Lc and the Fd subunits,

respectively. Indeed, these results corroborates our hypothesis of the wider isolation window behaving as the multiplexing effect of several precursor ions.

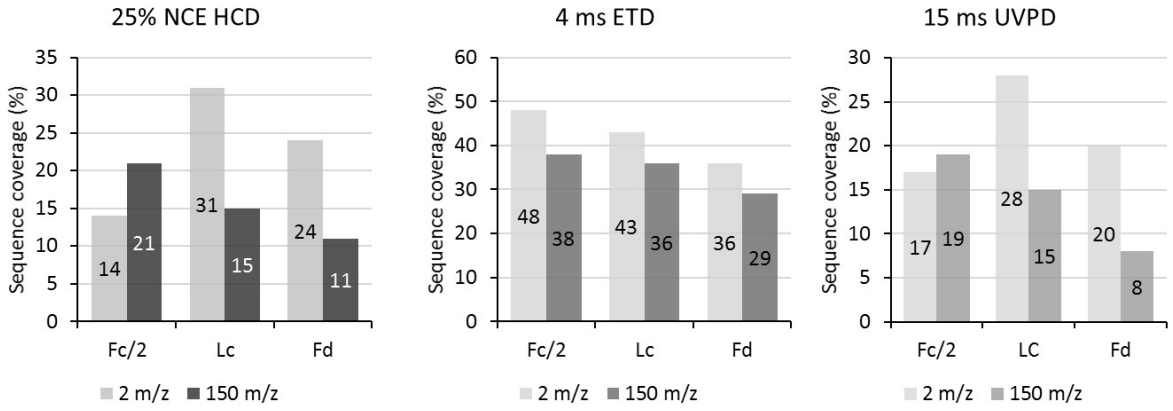


Figure 90: Comparison of MS/MS fragmentation upon 25% nce HCD, 4 ms ETD or 15 ms UVPD when selecting narrow (2 *m/z*) *versus* wide (150 *m/z*) isolation window. The lightest colors represent narrow window that was actually used for the previous optimization experiments, whereas the darkest colors represent the wider isolation window. This graph highlights the importance of choosing narrow isolation window to increase the sequence coverage, with one exception for the Fc/2 fragmentation upon HCD.

Overall, we showed here that using a narrow isolation window can increase the specificity of ions fragmentation and thus yield in higher sequence coverage.

6.4 Final sequence coverage outcome upon optimization of several parameters

The total residue cleavage obtained upon optimization of several parameters allowed to characterize from 71-78% sequence coverage of all subunits instead of 56-65% before optimizations (**Figure 91**). This highlights the importance of taking into consideration the several key parameters listed in **Table 17** namely; the choice of the precursor ion, the fragmentation time or energy and the choice of the most suitable isolation window of the fragmented ion.

Table 17; Name and value of the optimal MS/MS parameters leading to the highest sequence coverage (71-78%) of the three subunits.

Parameter	Optimized value
Precursor ion	Fc: 1,098.05 *m/z* (23+) Lc: 1,066.52 *m/z* (22+) Fd: 1,209.63 *m/z* (21+)
HCD energy	25% NCE
ETD reaction time	4 ms
UVPD irradiation time	15 ms
Isolation window width	2 *m/z*

Nonetheless, despite this considerable increase in sequence coverage, nearly 90% of the obtained fragments remained unassigned with theoretical fragments from the targeted sequences (**Figure 91**). In more details, only 281, 353 and 336 fragments were explained out of 2,119, 2,480 and 2,332 obtained for the Fc/2, the Lc and the Fd subunits, respectively. The latter challenge is often noted in TD-MS experiments and could be due to the several reasons we mentioned beforehand in **paragraph (d).**

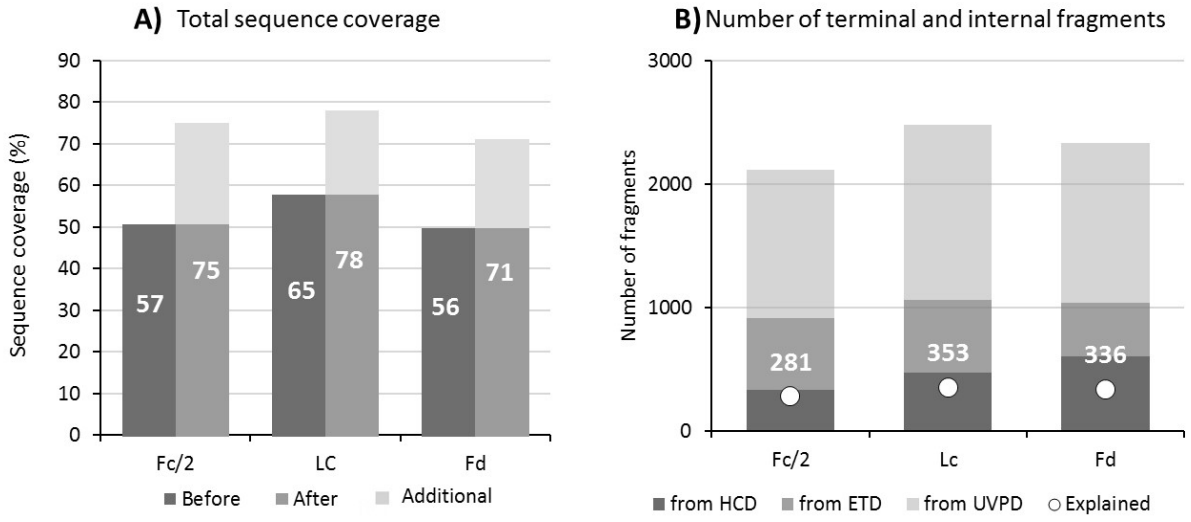


Figure 91: Final outcome of cumulated residue cleavage and number of fragments obtained with the optimal conditions: HCD 25% nce, ETD 4 ms and 213 nm UVPD 15 ms with an additional SID at 10V and an isolation window of 2 *m/z*. The highly charged precursor ions were selected for MS/MS fragmentation namely 23+, 22+ and 21+ for the Fc/2, the Lc and the Fd, respectively. (**A**) Total sequence coverage when combining the fragment ions obtained from the three individual experiments (bars in light red), compared with the residue cleavage obtained before optimizations (bars in darker red). The additional residue cleavage is depicted in pink. (**B**) The number of obtained fragments (black axis) *versus* the matched fragments with the subunits sequence (red axis) obtained in the optimal conditions. The fragments obtained from HCD, ETD and UVPD are represented in blue, red and green, respectively.

In the case of all fragmentation techniques, we generally search for the predominant ions without looking for neutral losses or fragments upon side chains dissociation. Consequently, there may be internal fragments that remain unassigned, or fragments that are miss-deconvoluted due to their overlapping isotopic profiles. In particular, in the case of ETD activation, the charge-reduced species are highly abundant with overall good isotopic distribution and therefore, could hamper the detection of their neighboring low abundant *m/z* and their assignment (**Figure 92**).

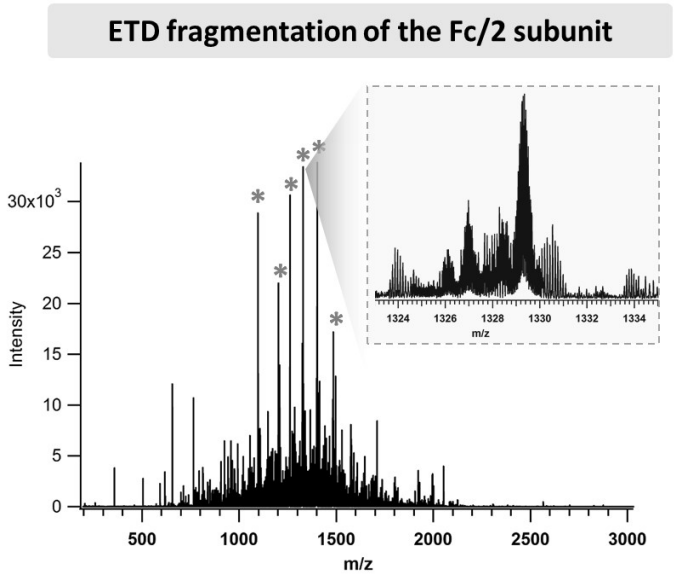


Figure 92: MS/MS spectra obtained after 4 ms ETD fragmentation of the Fc/2 subunit upon selection of the 23+ charge state. The fragment ions depicted with a red star represent the most abundant ions that result from a charge reduction of the precursor ion. A zoom on the 1329.24 *m/z* (19+) shows an overlapping distribution of the most abundant fragment ions with neighboring less abundant ions, which hinders the identification of the latter ones during deconvolution process.

One of the main reasons also, remains in failing to assign all the obtained fragments resulting from the different fragmentations pathways, from the 90% obtained ions that were not matched with the subunits sequences, more than 50% corresponded to fragment ions obtained *via* UVPD which are shown in green color in **Figure 91**. UVPD generates a plethora of *a/x*, *b/y* and *c/z* fragment ions which are located throughout the whole amino acid sequence of the proteins, some of those fragment ions do not contain the N-ter nor the C-ter residues of the subunits, thus are not assigned by conventional TD-MS softwares. Assigning those internal fragments can undoubtedly facilitates their identification in the amino acid sequence.

7. Addressing the challenges of limited total sequence coverage

To tackle the limitations mentioned above, some solutions are proposed. First, we aimed at reducing spectral congestion by using proton transfer charge reduction (PTCR), to enable the identification of more fragment ions species. Next, we took advantage of a recently developed TD-MS software; ClipsMS [229], dedicated for internal fragments assignment, in addition to terminal ones.

7.1 Benefits of PTCR in MS^3

In order to evaluate the impact of PTCR for providing a more thorough identification of the obtained fragments, we first combined it with ETD fragmentation. **Figure 93** illustrates the workflow that we followed on the Fc/2 as an example. After a first MS scan that provides the accurate mass measurement of the subunit, we performed an ETD fragmentation using the optimized MS^2 parameters (see **Table 17**). Among the product ion population generated by ETD, we isolated the *m/z*

1108.95 that clearly overlapped with several other fragment ions, and subjected it to PTCR MS^3 using an isolation window of 200 *m/z*. Of note, at the MS^3 level, we are only reducing the charge of the selected fragments ions to decomplexify the MS^2 spectra in this window, and no further fragmentation is induced.

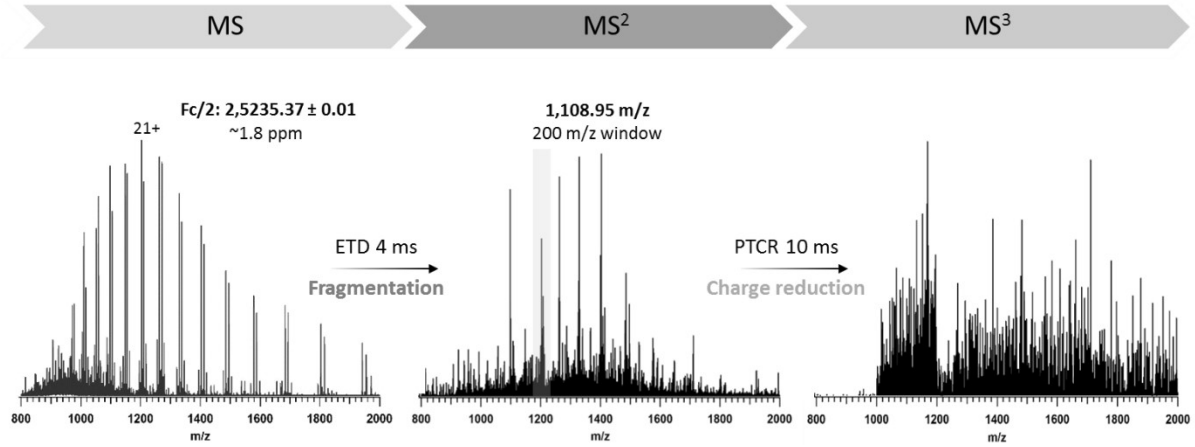


Figure 93: PTCR reaction after ETD fragmentation of the Fc/2 subunit. The Fc/2 is first analyzed at the MS level for accurate mass measurement, then the selected precursor ion (21+ in this case) is subjected to ETD fragmentations in the optimized conditions (4 ms here). A product ion (1108.95 *m/z*) is selected using a window of 200 *m/z* and subjected to PTCR ion-ion reaction for charge reduction to identify more overlapping species in the selected window.

Interestingly, the fragment ions obtained upon PTCR MS^3 at the selected window allowed reaching a sequence coverage of 42% for the Fc/2, these results summed with those upon only ETD MS^2 fragmentation increased the overall sequence coverage up to 56%. Similar results were observed for the other subunits rising the sequence coverage to 50% and 47% for the Lc and the Fd, respectively (**Figure 94**).

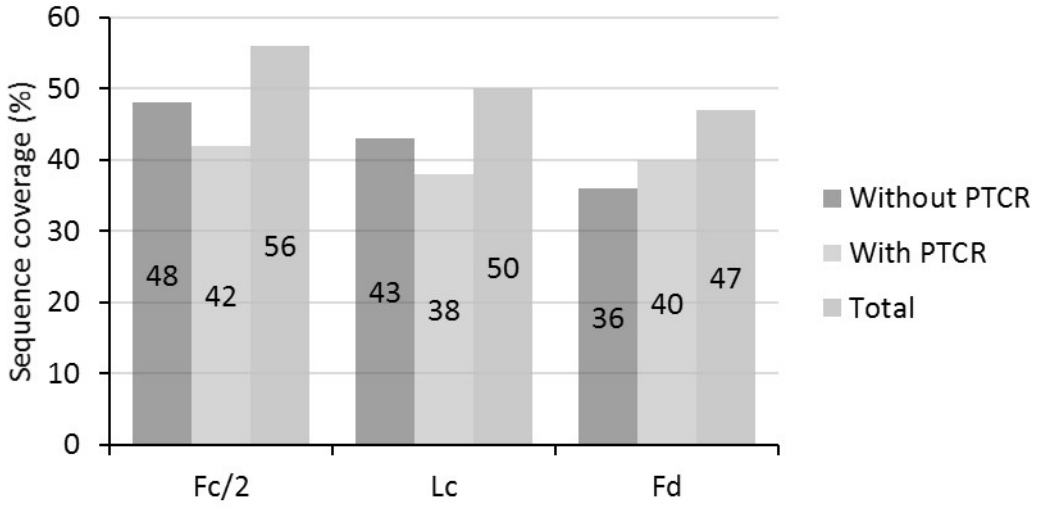


Figure 94: Sequence coverage obtained upon 4 ms ETD MS^2 reaction before and after PTCR MS^3 reaction. The residue cleavage of the Fc/2, Lc and Fd subunits before PTCR is represented in red, and the one obtained upon ETD 4 ms x PTCR 10 ms is represented in orange. Cumulative sequence coverage of the two experiments is represented in pink.

In light of this considerable increase, we compared the fragment ions obtained from both experiments. Several new *c*- and *z*- type fragment ions were observed when applying PTCR MS^3, these unique fragments account for 27, 22 and 32 for the Fc/2, the Lc and the Fd subunits, respectively (**Figure 95**).

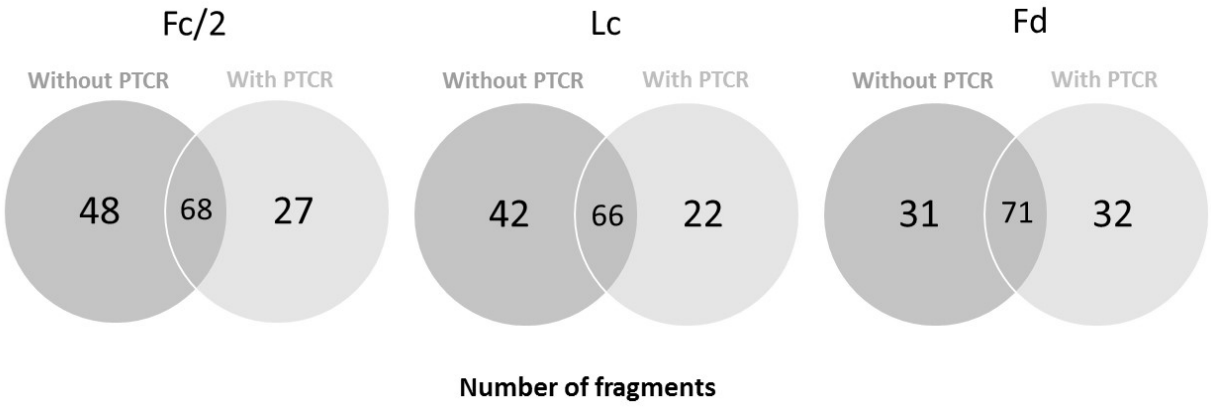


Figure 95: Venn diagrams of the number of obtained fragments *via* ETD 4 ms with and without PTCR 10 ms for the Fc/2, the Lc and the Fd subunits. These diagrams show that a considerable number of unique fragments are obtained only when using PTCR which pinpoint to the importance of using this reaction for more fragments identification.

Moreover, the fragmentation maps showed that the gained fragment ions were located all over the central portion of the subunits sequences, leading to a homogeneous and complementary distribution. It is worth noting that a large number of ions were accounted for one, as they belonged to complementary ion pairs and thus do not contribute in increasing the sequence coverage. The number of those golden pairs was nearly doubled when subjecting the Fc/2 subunit to PTCR reaction (from 13 to 24), which helped increasing the confidence in the N-glycosylation site assessment (**Figure 96**). The same behavior was noticed upon Lc and Fd subunits PTCR reaction where the complementary ion pairs were accounted for 21 for both Lc and Fd subunits instead of only 8 and 14 for the Lc and Fd, respectively without PTCR.

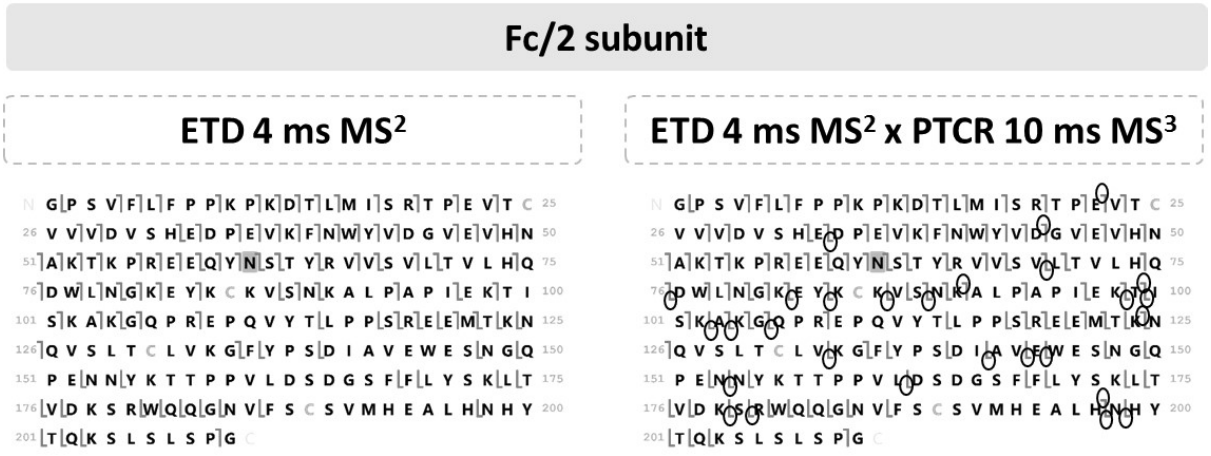


Figure 96: Fragmentation maps obtained upon ETD 4 ms fragmentation of the Fc/2 upon selection of the 23+ charge state (left). Fragmentation of the same ion followed by PTCR MS3 reaction on the 1108,95 *m/z* for 10 ms with an isolation window of 200 *m/z* (right). The fragment ions obtained after PTCR are depicted with a black circle.

The benefits of applying PTCR at the MS^3 level was also highlighted after HCD fragmentation. The spectral simplification at the same selected region using similar isolation window and similar reaction time (200 *m/z*, 10 ms), allowed to increase the sequence coverage up to 18%, 34% and 32% instead of only 14%, 31% and 24% for the Fc/2, the Lc and the Fd subunits, respectively. This increase is owed to the identification of new ions through PTCR reaction, accounted for 10, 12 and 21 new *b*- and *y*-fragment ions with no prominent increase in the complementary ion pairs (15 when applying PTCR *versus* 14 without PTCR, for the Fc/2). Despite the overall limited sequence coverage obtained upon HCD fragmentation, the benefit of PTCR could not be neglected as it increases the confident in confirming the amino acid sequence, and thus confirm the presence of the modification site of the Fc/2 (**Figure 97**).

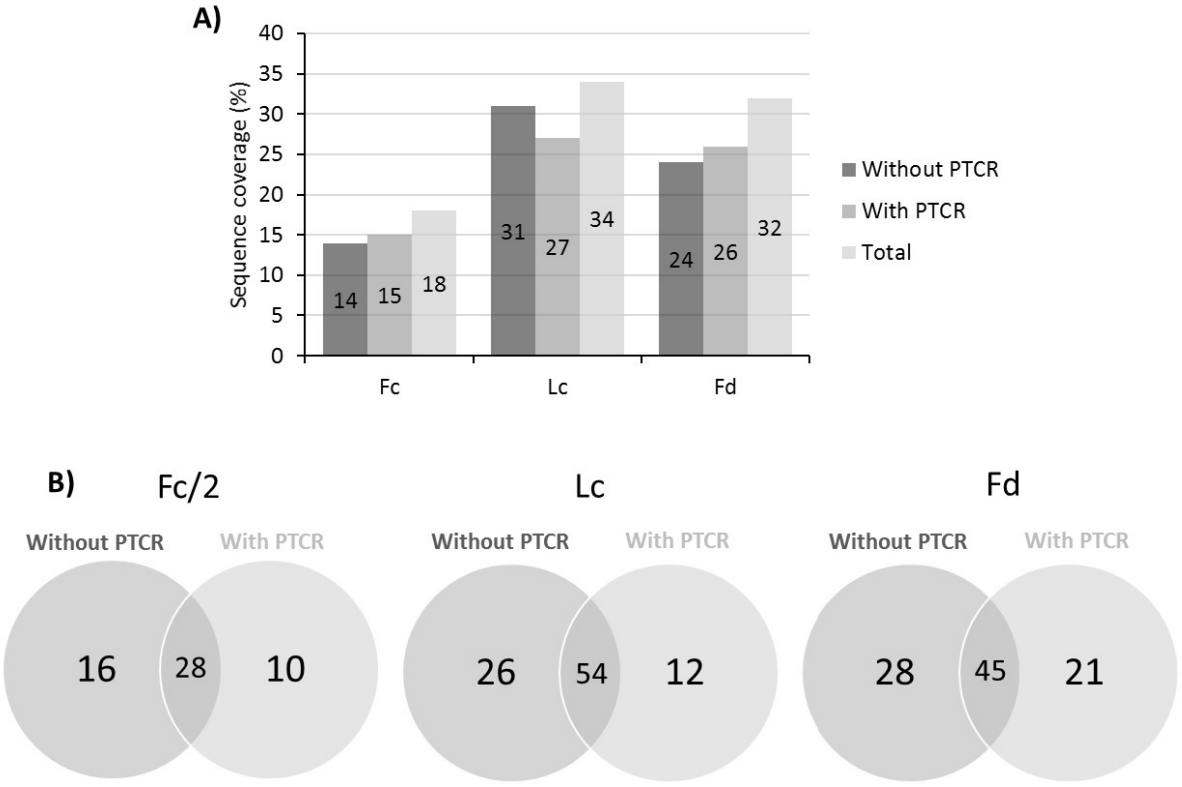


Figure 97: Results obtained upon using PTCR reaction in combination with HCD fragmentation. (**A**) Sequence coverage obtained upon 25% nce HCD MS2 reaction before and after PTCR MS3 reaction. The residue cleavage of the Fc/2, Lc and Fd subunits before PTCR is represented in blue, and the one obtained upon HCD 25% nce x PTCR 10 ms are represented in orange. Cumulative sequence coverage of the two experiments is represented in purple. (**B**) Venn diagrams of obtained fragments *via* HCD 25% nce with and without PTCR 10 ms for the Fc/2, the Lc and the Fd subunits. These diagrams show that a considerable number of unique fragments are obtained only when using PTCR which pinpoint to the importance of using this reaction for more fragments identification.

In light of the results shown above, we can clearly observe the considerable impact of PTCR ion-ion reaction to reduce the spectral congestion. The latter gives rise to identifying new fragments, by simply reducing the overlap between existing fragment ions, which improves their identification. Herein, a narrow window of 200 *m/z* for 10 ms was sufficient to allow an increase of 8-11% of the sequence

coverage when applying PTCR after ETD fragmentation, whereas a lower increase of 3-8% was observed when HCD outcome is subjected to the same reaction. In fact, PTCR does not enhance the ion dissociation but it facilitates product ion identification through spectral de-congestion. Thus, as the resulted spectra from ETD MS^2 spectra is more complex compared to the one upon HCD fragmentation, fragment ion overlapping is more likely. For these reasons, we tend to reduce the overlapping of the different isotopic envelopes in the selected *m/z* region and therefore more fragment ions are identified. Moreover, the number of fragments upon ETDxPTCR contributed more in increasing the confidence in the N-glycosylation site assessment, and accounted for 27 overall fragments instead of only 8 fragments in the case of HCDxPTCR.

Nevertheless, it is worth noting that several studies reported the use of larger isolation windows or a combination of several narrow windows, which tends to increase significantly the sequence coverage of large intact proteins [225], however, it was noticed that when applying longer ion-ion reaction times, several informative product ions are lost due to excessive charge reduction or complete deprotonation. Thus, our promising results open doors to more optimizations of this parameter when analyzing mAb-based formats.

7.2 Added value of internal fragments assignment

Internal fragments are a result of two cleavage events that generate *ax*, *ay*, *az*, *bx*, *by*, *bz*, *cx*, *cy* and *cz*, depending on the used activation technique. Contrary to terminal fragments that contain the N- or C-terminus of the protein, internal ones do not include neither protein termini, and have often remained unassigned due to the focus of classical TD-MS softwares on assigning only terminal fragments. Here, we evaluate a recently developed software, ClipsMS [229], intended for internal fragment assignment, in additional to terminal ones. ClipsMS has been used to assign internal fragments obtained upon collisional and electron-transfer dissociations [129], as well as in combination with UVPD fragmentation [249], using mass tolerance of 2 ppm generally to avoid, or at least mitigate, false positives assignment [229, 324]. Therefore, to understand more the unidentified regions of the subunits sequences, we processed the obtained data through ClipsMS. The consideration of internal fragments upon the fragmentation of trastuzumab subunits using HCD, ETD and UVPD fragmentations, showed directly an increase in sequence coverage (**Figure 98**). In the case of HCD fragmentation, the sequence coverage was slightly enhanced and reached up to 25-49% for all subunits. As previously demonstrated, HCD provides the lowest fragmentation propensity due to the restrained cleavage in the N- and C-terminus sides of proteins and mAb subunits. Because collision-based dissociations are based on less energetic fragmentations pathways, therefore there is a lower chance for this technique to provide more internal fragments than ETD and UVPD. For ETD, the maximum sequence coverage reached was 65-72% for the trastuzumab subunits after internal fragments inclusion, which accounts for x1.5-1.8 times the

fragments matched when considering only terminal fragments. The formation of internal fragments upon electron-driven dissociations of <30 kDa proteins was already highlighted in a previous study [325], thus our results are in good agreement with literature. As expected, UVPD generated a significant number of internal fragments that allowed yielding up to 88-94% sequence coverage of the trastuzumab subunits, which account for more than the double of the sequence coverage when considering only terminal fragments. UVPD takes multiple fragmentation pathways that lead to a higher number of fragment types (*c/z*, *b/y*, *a/x*) compared to other techniques (only *b/y* for HCD and *c/z* for ETD), as explained previously. Thus, the probability of generating internal fragments with UVPD is even higher compared to HCD and ETD, namely *ax*, *ay*, *az*, *bx*, *by*, *bz*, *cx*, *cy*, and *cz* ion types. Consequently, the internal fragments number is significantly higher in the case of UVPD fragmentation and allow to drastically increase the sequence coverage of the three subunits which reached nearly.

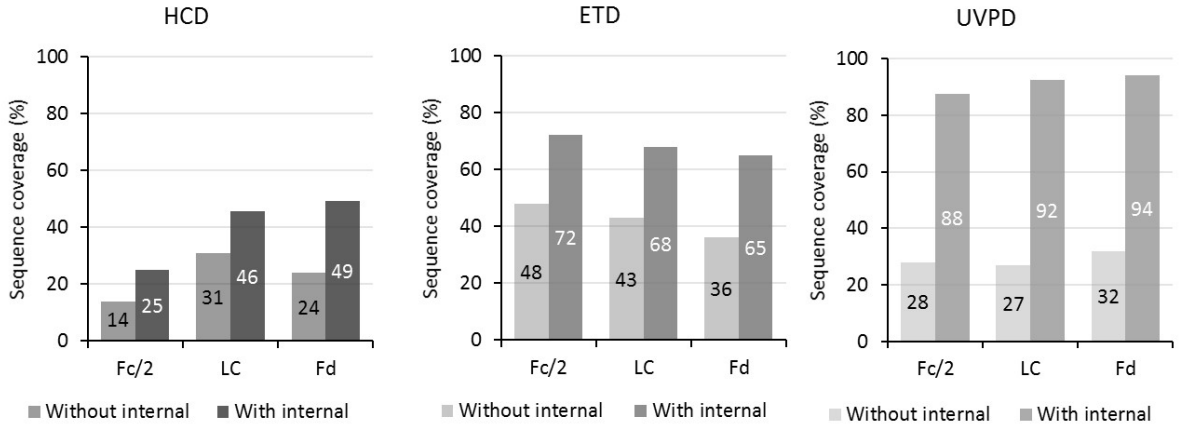


Figure 98: Sequence coverage of the different trastuzumab subunits before and after considering internal fragments for HCD (blue), ETD (red) and UVPD (green). Significant increase is observed especially upon UVPD fragmentation.

Beyond enhancing the sequence coverage, some of the assigned internal fragments could increase the confidence in the N-glycosylation site assignment, expected at the N61 position, prior confirmed upon terminal fragments consideration. Indeed, inclusion of internal fragments allowed increasing the confidence in the N-glycosylation site localization, as many fragments contained the N-glycosylation (**Figure 99)**. These fragments are represented in the fragment location maps obtained *via* HCD, ETD and UVPD in **1**, and accounted for 8, 13 and 113 internal fragments upon HCD, ETD and UVPD fragmentation, respectively. The number of internal fragments specific to the N-glycosylation site increased linearly with the sequence coverage upon the variation of fragmentation technique choice. These results emphasizes our hypothesis of UVPD generating more internal fragments compared to other techniques (HCD and ETD), which was also demonstrated by Brodbelt's group [324].

Overall, the inclusion of internal fragments upon combination of all techniques, allowed reaching 100% of the sequence coverage of all subunits with assignment of up to 38% of the MS/MS spectra.

Moreover, it increased the number of specific fragments to the N-glycosylation site, which accounted for 197 fragments upon combination of all three fragmentation techniques (28, 32 and 137 upon HCD, ETD and UVPD fragmentation).

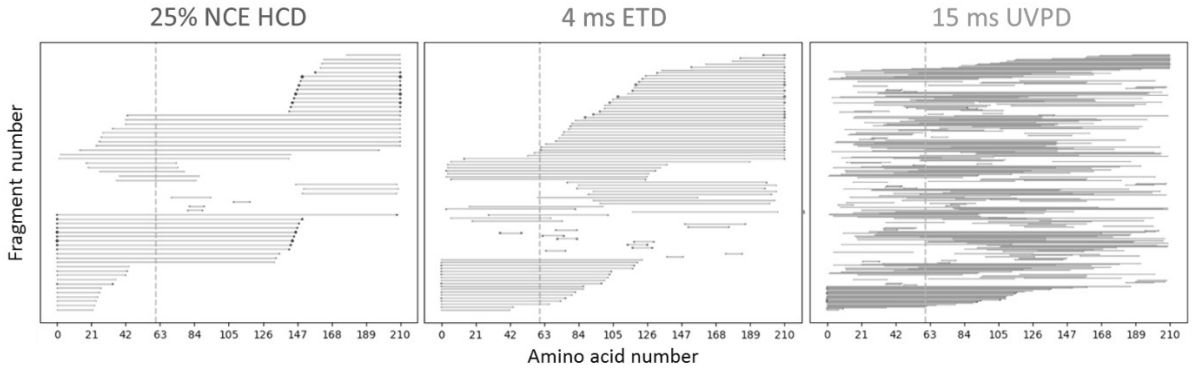


Figure 99: Fragment location map generated *via* Fc/2 subunit fragmentation upon 25% nce HCD (left), 4 ms ETD (middle) and 15 ms UVPD (right). Grey lines represent the fragment ions from their start to their end positions. Vertical orange dashes indicate the N-glycosylation site position. The blue, red and green dots represent the relative abundance of the fragment ions, with N- and C-terminus ions being more intense than the internal (central region) ones.

Nevertheless, the inclusion of internal fragments can increase the risk of false-positive identification, which requires a caution when considering those fragments [324]. In fact, we chose stringent deconvolution parameters (S/N ratio and fit factor) to avoid the risk of assigning bad quality peaks, therefore we increase the matching probabilities and thus the risk of assigning false positives is increased. Indeed, the **Figure 99** highlights the intensity of the terminal and internal matched fragments, depicted by dots on the extremities of the fragments, represented by grey lines from the start to the end amino acid. The most intense dots correspond the most intense fragment ions, and vice versa. As observed in the maps, the terminal fragments covering the N- and C-terminus are clearly more intense than the internal ones in the case of HCD and ETD fragmentation. Upon UVPD, the majority of the fragments represent lower intensities, as expected due to the low depletion of the precursor ion that remains very intense and the high number of fragments generated, prior discussed in this chapter.

In the case of UVPD in particular, the high number of internal fragments either could be due to the multiple fragmentation pathways, or could be a result of mis-matching due to the complexity of the spectra or matching several fragments with the same mass to multiple amino acid association possibilities. To explore this risk of false positives, we searched for internal fragments using a scrambled sequence of the trastuzumab subunits. Sequence coverage plots obtained for trastuzumab Fc/2, Lc and Fd subunits, and the scrambled sequences are displayed in **Figure 100** based on HCD, ETD and UVPD.

First, to understand more the nature of the internal fragments matched upon sequence scrambling, we searched for generated fragments from HCD, where a low sequence coverage of 9%, 17% and 27%

was observed for the scrambled sequences of the Fc/2, the Lc and the Fd subunits, respectively, instead of 25%, 46% and 49% for the real sequences. This decrease is due to the drastically loss of terminal fragments that did not match with the scrambled sequences, as expected. Similarly, for ETD a decrease of 48-57% was observed upon sequence scrambling of all subunits, highlighting the reduced number of terminal fragments and few internal fragments. However, for UVPD the sequence coverage of the scrambled subunits sequences was surprisingly comparable with the one obtained for the real sequences. Only 1-8% decrease was observed for the three subunits, which means that a high number of fragments is still miss-matched. Those fragments are mainly internal (176, 288 and 212 internal fragments for the Fc/2, Lc and Fd subunits, respectively.), and could be a result of a series of amino acids that have similar masses than those from the real sequences. These results were supported by studies from Brodbelt's group that observed similar behavior upon sequence scrambling when considering a high variety of fragments [324]. Indeed, as UVPD is the technique that affords multiple fragmentation pathways leading to a wide variety of internal fragments, the risk of matching false-positives is even greater than for other techniques.

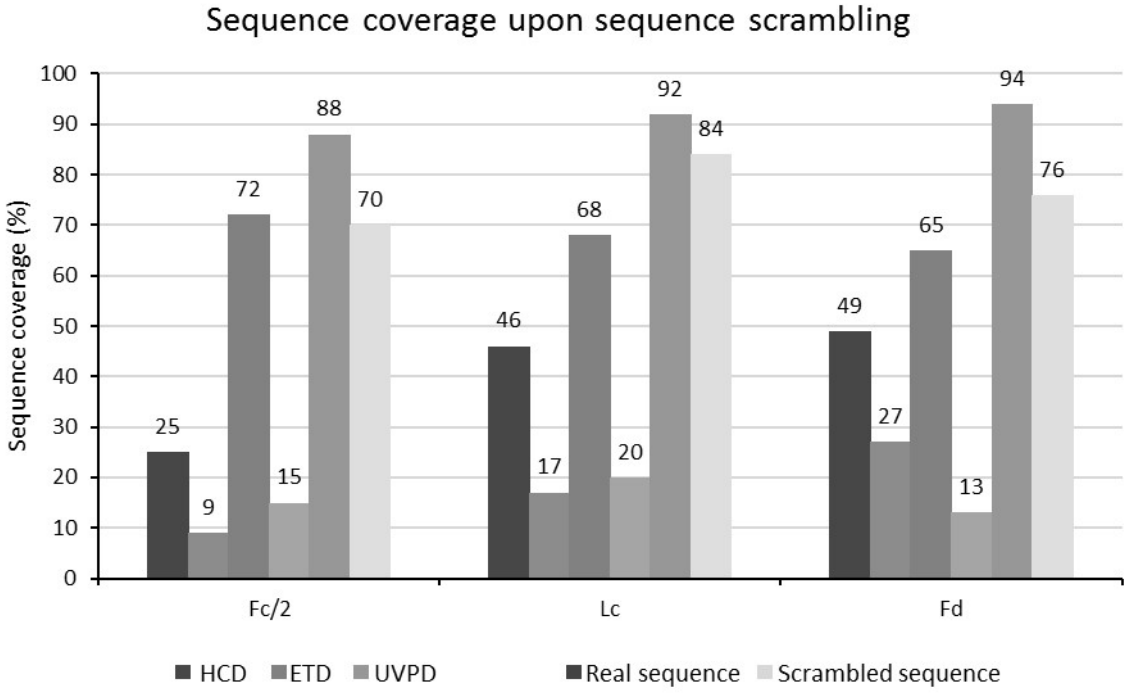


Figure 100: Sequence coverage upon three experiments (HCD 25% NCE, 4 ms ETD and 15 ms UVPD) with consideration of internal fragments. The sequence coverage obtained when using the real sequence is represented in darker colors (blue for HCD, red for ETD and green for UVPD), while the data obtained from the scrambled sequence are in lighter colors.

Moreover, among the matched fragments upon sequence scrambling, several internal fragments contained the N-glycosylation at the N61 position of the Fc/2 subunit. In more details, 9 fragments upon HCD, 6 fragments upon ETD and 102 fragments upon UVPD were containing the N-glycosylation modification when using the scrambled sequence *versus* 28, 32 and 137 fragments for HCD, ETD and

UVPD, respectively when using the real sequence. This shows clearly that many false-positives are assigned, especially upon UVPD fragmentation that generates many internal fragments. Therefore, when it comes to assessment of modification sites, it is worth noting that internal fragments can be tricky and can easily lead to mis-assignment (**Figure 101**).

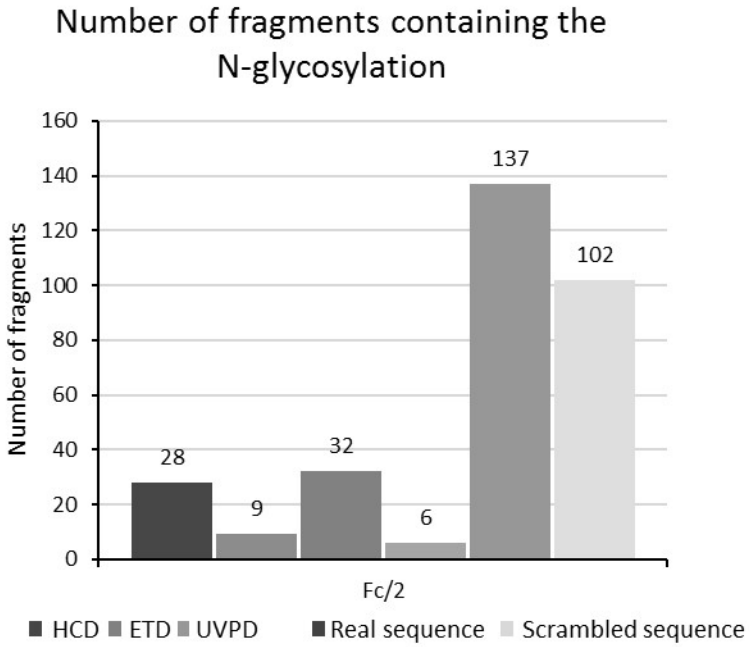


Figure 101: Number of obtained internal fragments that contain the N-glycosylation modification upon three experiments (HCD 25% NCE, 4 ms ETD and 15 ms UVPD). The number of internal fragments obtained when using the real sequence is represented in darker colors (blue for HCD, red for ETD and green for UVPD), while the data obtained from the scrambled sequence are in lighter colors.

In summary, our results pinpoint to the possibility of maximizing the sequence coverage of trastuzumab subunits upon inclusion of internal fragments, usually not assigned with classical TD-MS softwares. However, even when using low mass tolerance (2 ppm), the probability of matching false-positives is increased when matching internal fragments, especially in the case of UVPD. The increase risk of false-positives is mainly due to absence of scoring systems that reflect the confidence of sequence or modification site assignment; therefore, more caution should be taken when considering those internal fragments.

8. Conclusions

In this chapter, we aimed at developing a TD-MS workflow on the latest generation Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) for the characterization of reference proteins (Myoglobin 17 kDa and Carbonic anhydrase 26 kDa) as well as a reference mAb (trastuzumab, ~150 kDa). Our optimizations pinpointed that the achievement of high sequence coverage of a small size protein or a

challenging mAb along with PTMs localization, is possible when taking into consideration the variety of tunable MS^n parameters.

In summary, our optimizations open doors for the primary structure characterization of more complex mAb-based formats. Among the most challenging structural information to obtain is the conjugation sites on ADCs for instance, which could possibly be provided thanks to fine instrumental tuning or combination of several fragmentation techniques, as showed in a recent study [129]. This ability of TD- and MD-MS approaches to provide information on the primary sequence of mAb-based products, could be valuable for the comprehensive characterization of multiple therapeutics formats, which will be discussed in the next chapter.

9. Scientific communications

Benazza R., Beaumal C., Hernandez-Alba O. and Cianférani S. **Optimization of MD-MS analysis and data treatment to improve the characterization of mAbs.** *Oral presentation*. *INSERM*. January **2022**, *Online.*

Chapter 2: Full characterization of a single domain antibody-drug conjugate (sdADC) using top down MS

1. Analytical context

To advance treatment and diagnostic options in cancer therapy, biopharmaceutical companies focus on the development of new antibodies-formats to broaden the immunotherapy library [1]. Among those entities, single-domain antibodies (sdAb) also called VHHs (heavy chain variable domain) gained an increased interest in the last decade, thanks to their small size, fast generation, their easy tissue penetration as well as their high specificity and affinity compared to their antibody counterparts [93-95, 97, 99, 326]. Known as nanobodies, sdAbs are the small variable domain that binds to antigens, derived from camelid heavy-chain antibodies (HCAbs) that are deprived from the light chain contrary to conventional antibodies [93] (**Figure 102**). Similarly to antibodies, nanobodies could be conjugated to fluorescent molecules to generate chromobodies and fluobodies that are widely used for live-cell imaging of their endogenous antigens [327-328]. Moreover, to improve their efficacy and potency, sdAbs could be conjugated to cytotoxic payloads paving the way for the next generation of chemodrug therapies [96]. Consequently, robust and straightforward analytical strategies have to be developed for the comprehensive characterization of these sdAb conjugates (single-domain antibody drug conjugates, sdADCs). Commonly, the biochemists tend to lean towards SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) analysis to assess the production and the conjugation efficiency of these entities [326, 329]. Albeit this technique offers relatively prompt and approximative molecular weight determination, it still suffers from lower resolution, lower sensitivity and lower mass accuracy compared to mass spectrometry approaches. Furthermore, the localization of the exact site of conjugation (for sdADCs) is unattainable through a simple SDS-PAGE gel. In this context, we aimed at evaluating our TD-MS workflow in complementarity to SEC-nMS approach, to fully characterize an anti-EGFR sdAb [330] with its sdADC analog, prepared during a TACT secondment at Cotton's team (Almac discovery, UK). Owing to their small size, sdAbs represent an ideal candidate for TD-MS approaches.

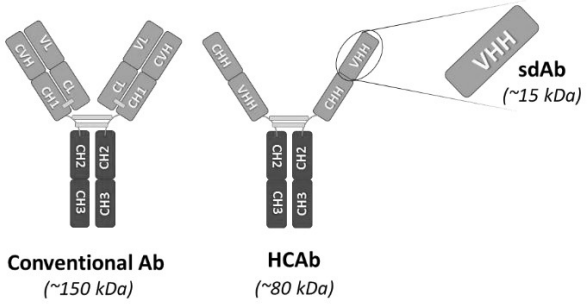


Figure 102: Schematic representation of single-domain antibodies (sdAb, VHH or nanobody) in comparison to heavy-chain only antibodies (HCAbs) and conventional antibodies (Abs).

2. State of the art

In terms of recently developed approaches aiming at nanobodies characterization, a limited number of studies focused on the characterization of their primary structure. A noteworthy study conducted by Resemann *et al.* (2010), discussed a valuable approach of elucidating the primary structure of ~14 kDa nanobodies, using in-source Decay (ISD) with Matrix-Assisted Laser Desorption Ionization ToF (MALDI-TOF) for top-down analysis. This study represents a significant milestone in the field of mass spectrometry as it was the first to elucidate a complete primary structure characterization of proteins with MW of >13 kDa [204]. More recently, Macias *et al.* (2022) conducted a study focused on the characterization of nanobody-antigen complexes upon the combination of native mass spectrometry and UVPD dissociation. Indeed, the comparison of UVPD-generated fragments from free nanobodies and nanobodies included in the sdAb-antigen complex, facilitated the precise assignment of nanobody paratopes location [248]. Nonetheless, no fully comprehensive TD-MS characterization of nanobodies and nanobody-conjugates was documented yet.

3. Objectives

The first step of this study would consist of the generation of the site-specific fluorescent anti-EGFR nanobody (~15 kDa) using expressed protein ligation technology (EPL) [331-333], followed by its complete characterization using our cutting-edge MS-based approaches:

- Native MS (nMS) to assess homogeneity, drug load distribution (DLD) and average drug-to-antibody ratio (avDAR),
- Top-down MS (TD-MS) using a combination of complementary fragmentation techniques (HCD, ETD and UVPD) to validate the primary structure of the protein, provide the precise localization of the conjugation site and assess the disulfide bond integrity in the nanobodies scaffolds.

4. Production of anti-EGFR sdAb/sdADC

Anti-EGFR nanobodies are proteins that target the epidermal growth factor receptor (EGFR or Her1) [330, 334-335], which is over-expressed in a large number of cancer cells. By targeting EGFR [330, 336], these nanobodies aim to block the growth signals that promote cell proliferation, thereby helping to slow disease progression and improve treatment outcomes in cancer patients. Herein, we aimed at producing an anti-EGFR nanobody, followed by its site-specific conjugation to a fluorescent molecule to mimic commercially available nanobody conjugates.

4.1 Anti-EGFR nanobody expression in *Escherichia coli (E. coli)*

In a first place, we transformed competent BL21 *E. coli* cells with the suppressor plasmid pTXeGFR, the latter containing the nanobody sequence in Luria-Bertani (LB) media. BL21 are chemically made competent through a $CaCl_2$ or $MgCl_2$ salt treatments followed by a heat-shock step, which allows the plasmid entry. Next, we streaked the stock cells in 100 µl/ml ampicillin (AMP)-containing agar plates that were incubated at 37°C overnight to induce cells transformation. Here, the use of AMP antibiotic was necessary to prevent any bacterial growth. After the successful transformation of colonies on the agar petri plate, a single colony was transferred to a 100 ml solution supplemented with LB media and AMP, which were incubated overnight at 37°C at 180 RPM. The pre-inoculum was then transferred to bigger flasks (3 flasks of 1L) and were incubated again at 37°C under 180 RPM for 2h30 for larger scale cultivation. When the optical density (OD600) reached 0.9 we stopped cells growth by a cold shock (4°C, 2h) before inducing their expression using 1 ml of a 500 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) solution at 25°C to the supernatant, overnight (**Figure 103**).

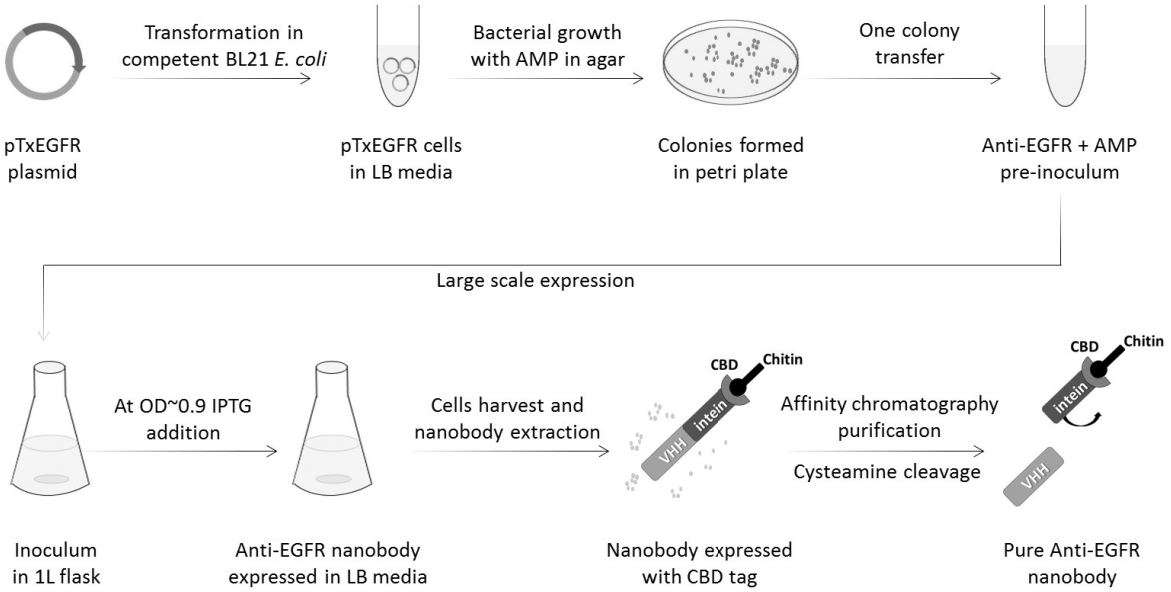


Figure 103: Bacterial expression anti-EGFR nanobody. The competent BL21 E. coli cells are transformed with the pTxEGFR plasmid containing the gene of the anti-EGFR nanobody, and are streaked on AMP agar plates for cells formation. A single colony was transferred to prepare a pre-inoculum that was then transferred to 1L flasks for large scale expression in LB media. At OD~0.9 the IPTG is added to the inoculum to induce protein expression. Protein expression resulted in a nanobody-intein-CBD complex that was purified with intein-CBD affinity chromatography with cysteamine cleavage to release the nanobody.

Briefly, IPTG binds to the lac repressor (LacI) on the BL21 E. coli cells, which inactivates its ability to prevent target genes transcription [337]. This allows RNA polymerase to transcribe the gene of interest and produce the desired nanobody. In absence of IPTG, the LacI binds to lac operon (LacO) which blocks the T7 promoter from initiating the transcription (**Figure 104**). The final step consisted in harvesting

the cells to extract the nanobody prior to nanobody purification using CBD affinity chromatography [338].

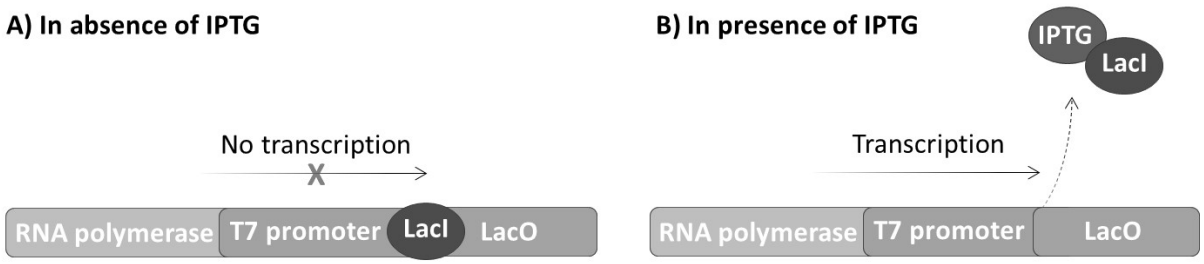


Figure 104: Schematic representation of the nanobody gene transcription before (**A**) and after (**B**) addition of IPTG that induces protein expression in *E. coli* cells. In absence of IPTG LacI binds to LacO which prevents the promoter recognized by the RNA polymerase to start the gene expression. When IPTG is present, it binds to lacI and releases the latter, which allows the promoter to initiate the transcription.

4.2 Purification by affinity chromatography

As mentioned above, we aimed in this study at using the expressed protein ligation (EPL) approach to produce our site-specific conjugated nanobody. The principle of EPL consists in labelling proteins through their C-terminal side upon thiol-mediated cleavage of the intein domain [332]. Therefore, a purification using chitin binding domain (CBD) affinity chromatography followed by a cysteamine cleavage (200 mM cysteamine, 16h, rT°) was necessary to eliminate impurities and facilitate the thiol-mediated cleavage [339]. In contrary to other intein cleavage methods that use dithiothreitol (DTT) [340], here the disulfide bond contained within our protein will be maintained thus the conjugation is expected to occur on the C-terminal thiol. **Figure 105** shows the obtained anti-EGFR nanobody sequence along with a SDS-PAGE analysis used to monitor the nanobody purification. For the expressed nanobody with the CBD tag (lane 1 and 2, **Figure 105**), the SDS-PAGE revealed a spreading of bands which reflects the heterogeneity of the protein. After washes using lysis and cleavage buffers, the drained fractions exhibited a single band around ~40 kDa (lane 5-12, **Figure 105**) that means the protein is relatively pure, in line with the nanobody attached to the CBD tag *via* its intein. The same mass is observed after inducing cysteamine cleavage (lane 13, **Figure 105**) which indicates that no reaction has occurred yet. After 16 hours of reaction, two distinct visible bands were observed corresponding to nanobody at ~15 kDa and CBD at ~30 kDa (labelled i and ii in **Figure 105**, respectively). This confirms the occurring of the cleavage of the nanobody-CBD complex that released a homogeneous nanobody after few washes (lane 17, **Figure 105**).

A)

Protein	Sequence	Expected mass (Da)
Anti-EGFR nanobody	MAVQLVESGGGSVQAGGSLRLTCAASGRTSRSYGMG WFRQAPGKEREFVSGISWRGDSTGYADSVKGRFTISRD NAKNTVDLQMNSLKPEDTAIYYCAAAAGSTWYGTLYEY DYWGQGTQVTVSSG	13,521.89
Chitin binding domain (CBD)	CITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDR HGNPVLADRLFHSGEHPVYTVRTVEGLRVTGTANHPLL SLVDVAGVPTLLWKLIDEIKPGDYAVIQRSAFSVDSAGFA RGKPEFAPTTYTVGVPGLVRFLEAHHRDPDAQAIADELT DGRFYYAKVASVTDAGVQPVYSLRVDTADHAFITNGFV SHATGLTGLNSGLTTNPGVSAWQVNTAYTAGQLVTYN GKTYKSLQPHTSLAGWEPSNVPALWQLQ	27,810.24

B)

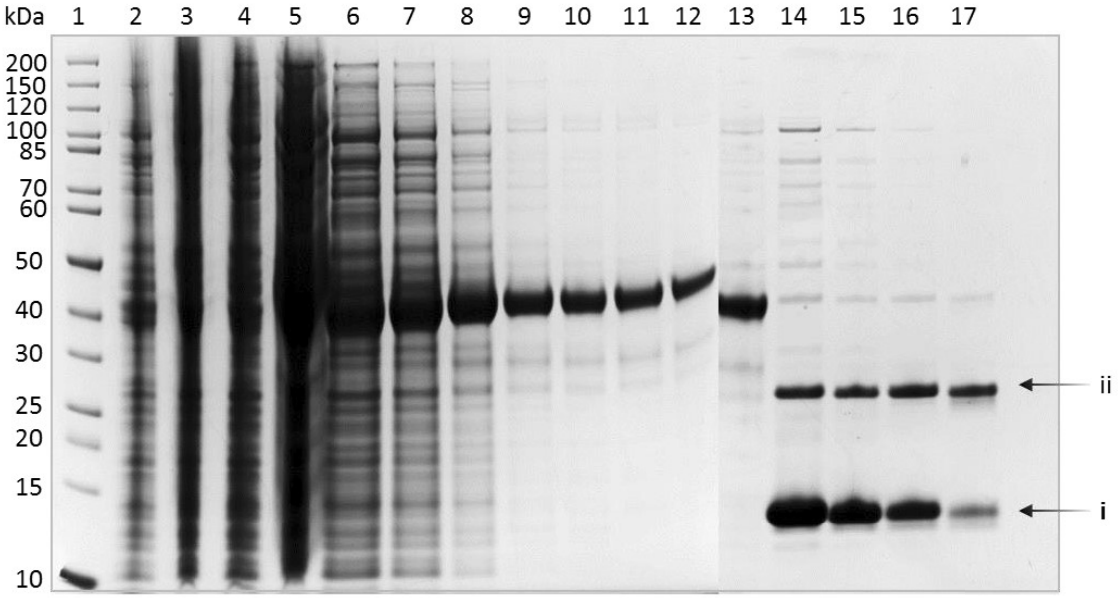


Figure 105: Purification of anti-EGFR nanobody by intein-CBD affinity chromatography. (**A**) anti-EGFR nanobody and chitin binding domain amino acid sequences with the expected theoretical masses. (**B**) SDS-PAGE gel stained with coomassie blue following the different steps of the nanobody expression and purification. Lane 1 corresponds to the PageRuler marker (Thermo Scientific) band. The IPTG induced cells, post-spin cells and unbound cell lysate showed heterogenic spreading bands (lane 2, lane 3 and lane 4, respectively). Lanes 5-8 represent the protein after lysis buffer wash while lanes 9-12 corresponds to the protein after cleavage buffer wash. After cysteamine cleavage, the mixture was analyzed at t0 (lane 13) and after 16h (lane 14). Two distinct bands were observed at (lane 15-17) after additional cleavage buffer wash, which correspond to (**i**) nanobody with cysteamine and (**ii**) CBD alone.

To separate our nanobody fraction from the CBD, we performed SEC with a Superdex 75 Increase 3.2/300 column (Cytiva, Germany) in PBS, using an AKTA system (Cytiva, Germany). The SEC-UV signal at 280 nm showed one major peak that could correspond to the nanobody species, due to no observed change in the conductivity (**Figure 106**). Consequently the fraction within this peak range were collected and showed a single band on Coomassie blue stained SDS-PAGE with a molecular weight of ~15 kDa (lane 3, 4 and 5 in **Figure 106**), in line with the nanobody bearing a cysteamine modification (theoritical mass: 13,580.89 Da), which implies a loss of one H_2O molecule as a result of a peptide bond between cysteamine and the glycine at the C-terminus.

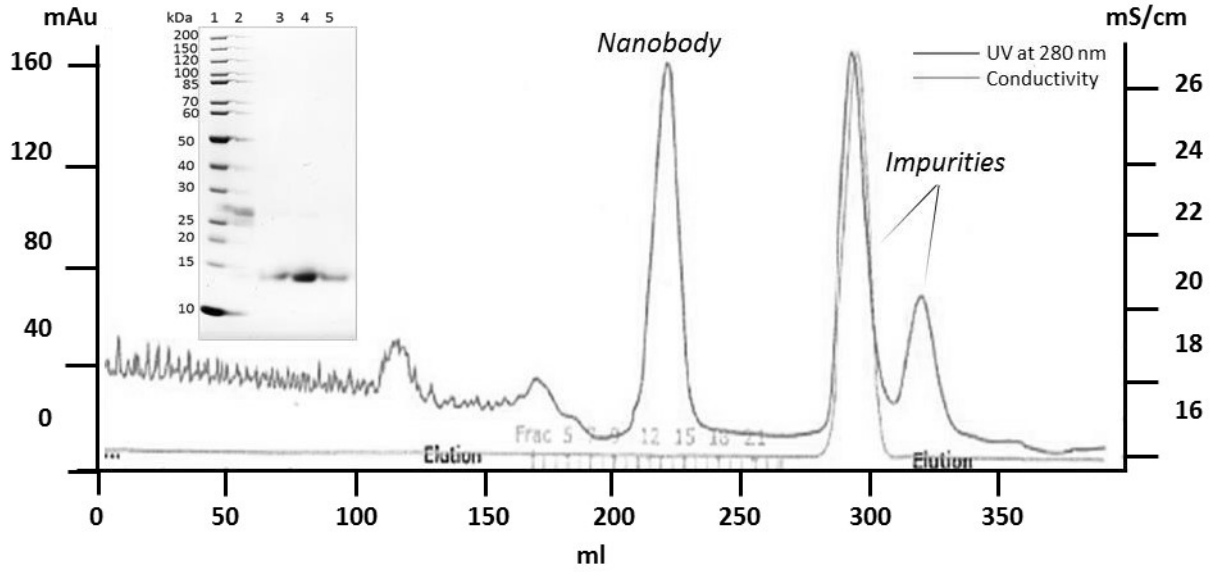


Figure 106: Purification of the obtained nanobody upon cysteamine cleavage using a Superdex 75 Increase 3.2/300 column (Cytiva, Germany) in PBS on an AKTA system (Cytiva, Germany). The UV signal at 280 nm is represented in blue line while the conductivity signal is represented in orange line. The eluted nanobody fractions were analysed on SDS-PAGE gel. The SDS-PAGE gel stained with Coomassie blue shows bands around ~15 kDa for nanobody fractions (lane 3, 4 and 5). Lane 1 corresponds to the PageRuler unstained marker (Thermo Scientific) and lane 2 corresponds to the impurities fractions.

In summary, the production of an anti-EGFR sdAb with a cysteamine at the C-terminal was achieved, based on SDS-PAGE confirmation, thus we next aimed at conjugating this nanobody with a fluorescent molecule to mimic a nanobody drug-conjugate (sdADC).

4.3 Site specific bioconjugation of anti-EGFR nanobody

In order to generate our site-specific sdADC, we aimed at covalently conjugating the resulted nanobody with a commercially available fluorescent dye; Alexa Fluor 488 (AF488, C30H25N4O12S2Na, 720 Da), through a maleimide reaction. We reacted 4 mg of anti-EGFR nanobody with 4x molar excess of AF488 reconstituted in 10 mM N,N-Dimethylformamide (DMF), that we left for 30 min in the dark at room temperature. Briefly, the fluorescent molecule that contains a maleimide functional group will react readily with the thiol group of the C-terminal cysteamine and form a thioether bond. Thanks to the high specificity and selectivity of this reaction, the disulfide bond within the nanobody will be maintained and thus an addition of an increment mass of 757 Da is expected exclusively on the C-terminal of the amino acid sequence (**Figure 107**). The latter mass corresponds to the nanobody mass after the addition of one cysteamine (C2H7NS, + 77 Da), through maleimide-thiol reaction inducing a loss of a H_2O molecule (- 18 Da), followed by a covalent linking of the AF488 molecule in absence of sodium clusters (C30H25N4O12S2, + 698 Da).

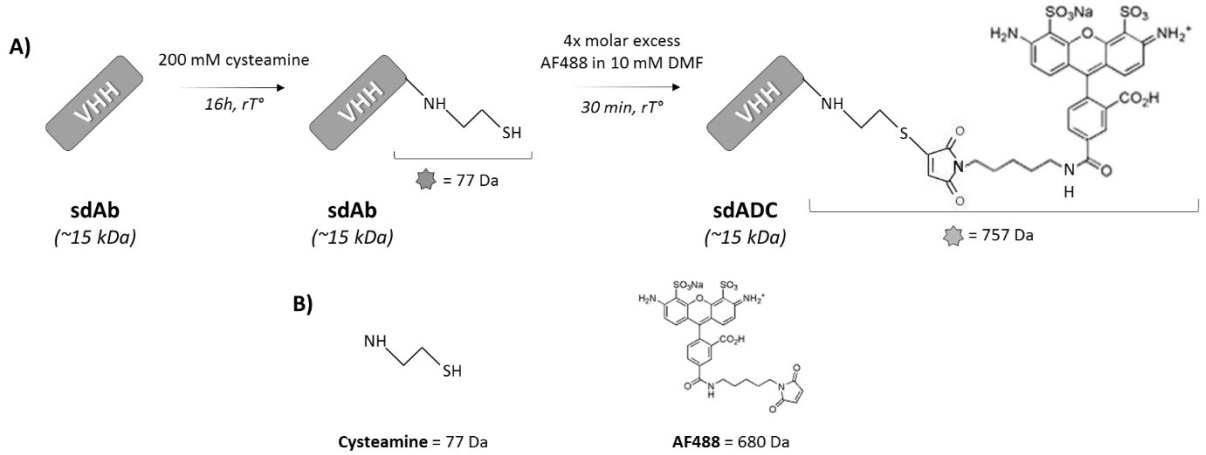


Figure 107: Labeling of anti-EGFR with Alexa Fluor 488 (AF488) dye following a cysteamine addition. (**A**) Schematic representation of the anti-EGFR nanobody cleavage by cysteamine followed by a labelling with AF488 through a maleimide reaction with the cysteamine's free thiol. (**B**) Chemical formulas and masses of cysteamine and AF488 before reaction. The cysteamine addition (+77 Da) will be depicted in a blue star in this manuscript and the payload addition (cysteamine + AF488) will be depicted in an orange star during the next sections.

The SEC-UV chromatogram (**Figure 108A**) shows the profile obtained after loading 3 times the reacted mixture onto a HiTrap 5 ml desalting column, using the AKTA system (Cytiva, Germany). Each change in conductivity corresponds to the signal of the unbound AF488, consequently the fractions that corresponded to the sdADC species were collected and analysed on SDS-PAGE to confirm the mass of the latter (**Figure 108B**). The observed masses on the coomassie blue stained SDS-PAGE reveal a mass around ~15 kDa corresponding to the sdADC. Moreover, a UV stained SDS-PAGE was run to confirm the fluorescent molecule labelling, which revealed a band around ~15 kDa, as expected. The sdAb and sdADC were both analyzed by UV/VIS to determine the degree of labelling (DOL). Taking advantage of a characteristic 450 nm absorbance for the fluorescent molecule attached on the sdADC led to achieve a DOL of 95% for the sdADC, suggesting a labelling of one AF488 molecule.

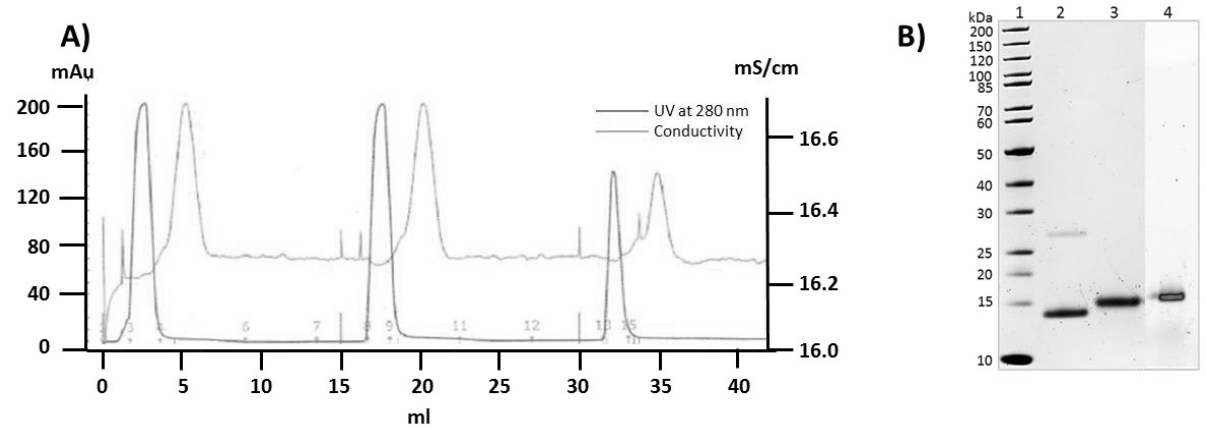


Figure 108: Purification of labelled anti-EGFR nanobody with AF488. (**A**) SEC-UV profile of the anti-EGFR labelled nanobody loaded three times and showing three peaks corresponding to the nanobody (blue line). Conductivity changes pinpoint to the unbound AF488 signal (orange line). The fractions corresponding to the nanobody were analyzed by SDS-PAGE. (**B**) SDS-PAGE gel of anti-EGFR nanobody with cysteamine (lane 2) and after AF488 conjugation (lane 3) both stained in coomassie blue. Lane 4 represents the conjugated nanobody visualized in UV which confirms the fluorescence of the nanobody.

Overall, those preliminary results obtained at Almac laboratory (Edinburgh, UK) highlight the successful formation of the reference (sdAb with cysteamine) and the conjugated nanobody (sdADC with AF488). Albeit SDS-PAGE revealed approximative masses of the obtained products, this technique still suffers from low resolution, low sensitivity and low mass accuracy. In addition, the ability of AF488 to access the cysteine residues within the disulfide bond that remains a possibility if the reaction conditions are not optimal could not be confirmed by a simple SD-PAGE analysis. Consequently, to pinpoint the exact site of conjugation with high resolution, our MS toolbox appear to be a promising solution.

5. Characterization of anti-EGFR sdAb products

For more accurate mass measurement, we took advantage of our MS toolbox to; first confirm the sdADC labelling with one AF488 molecule using SEC-nMS, before going further with TD-MS experiments to determine the specific site of conjugation and to assess the disulfide bond preservation.

5.1 Intact mass analysis and average degree of conjugation (avDAR) determination using SEC-nMS

For more sensitive, resolutive and accurate intact mass analysis of our final sdADC, we performed size exclusion chromatography (SEC) hyphenated to mass spectrometry (MS) in native conditions (SEC-nMS)[121-122, 124] which is a suitable approach for the assessment of DLD, D0 and avDAR for mAb-based formats, as previously demonstrated (**Part II**). This experimental set-up allows providing a more precise experimental mass along with a reliable relative quantitation of all the different species. The SEC-UV chromatogram (**Figure 109A**) reveals a major peak at ~3.88 min corresponding to homogeneous sdADC species. The mass spectrum corresponding to the main chromatographic peak exhibits mainly one population with an experimental mass of 14,145.7 ± 0.1 Da that can be assigned to the sdAb conjugated to one AF488 molecule, lacking the first methionine of the amino acid sequence. Indeed, this phenomenon of N-terminal methionine excision (NME) is well known for *E. coli* proteins that have alanine at position 2 [341], therefore this PTM was expected.

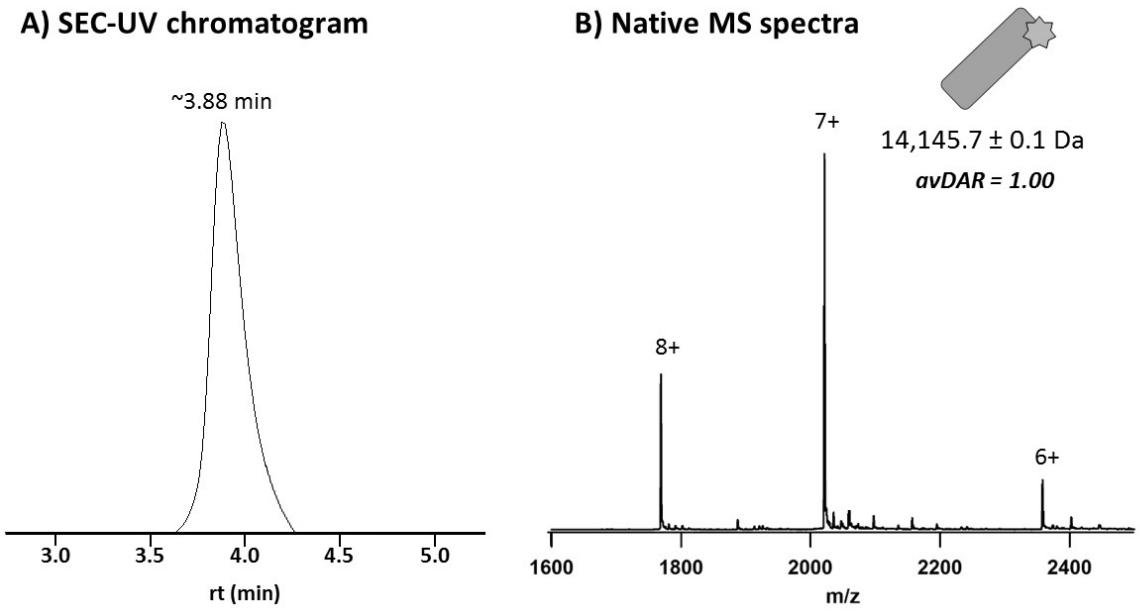


Figure 109: SEC-nMS analysis of anti-EGFR sdADC. **(A)** SEC-UV chromatogram of intact sdADC showing one major species at ~3.88 min with minor species at ~4.50 min. **(B)** Mass spectra of the main peak corresponding to the conjugated nanobody centered on the charge state 7+ with avDAR value calculated based on the MS intensities. The species depicted as a star corresponds to a fragment of AF488 conjugated to the nanobody.

In summary, SEC-nMS results allowed to confirm the conjugation of the sdADC, leading to an average drug-to-antibody ratio (avDAR) of 1 (**Figure 109B**). Additionally, the experimental mass of the D1 population corresponds to the sdADC with a disulfide bridge, suggesting that the integrity of the linkage between both cysteine residues has been maintained during the conjugation reaction. However, to determine the specific site of conjugation we need to investigate deeper into the amino acid sequence through TD-MS analysis.

5.2 Optimization of LC-MS parameters upon TD-MS analysis of the sdAb

As demonstrated in the previous chapter (**Part III, chapter 1**), top-down MS (TD-MS)[28-33] is a valuable approach for the primary sequence characterization of antibody (Abs)-based formats, through their direct fragmentation at the intact level, without prior sample preparation. Consequently, we focused at first adapting our TD-MS workflow to characterize the primary sequence of the sdAb and confirm the cysteamine addition on the C-terminus. Then, we aimed at determining the specific site of the sdADC conjugation (AF488), and to investigate the preservation of the disulfide bridge during the conjugation reaction.

Taking into consideration that efficiency of most fragmentation techniques depends not only on the mass [23], but also on the number of disulfide bridges of the protein of interest and the conjugation site [250, 342-343], the primary sequence confirmation of the sdAb was carried out upon reduction of the cysteine residues (10 mM TCEP for 30 min, rT°). Subsequently, the reduced sdAb was analyzed by LC-

TD-MS (See **Experimental part**) using different activation techniques (HCD, ETD, and UVPD). The reduction of the disulfide bridge within the sdAb consequently altered the structure leading to a higher number of residues exposed to protonation (**Figure 110**). Indeed, upon disulfide bridge reduction, a mass difference of 2 Da was observed; *i.e.* a mass of 13,449.1 ± 0.4 Da *versus* 13,447.2 ± 0.8 Da for the unreduced sdAb.

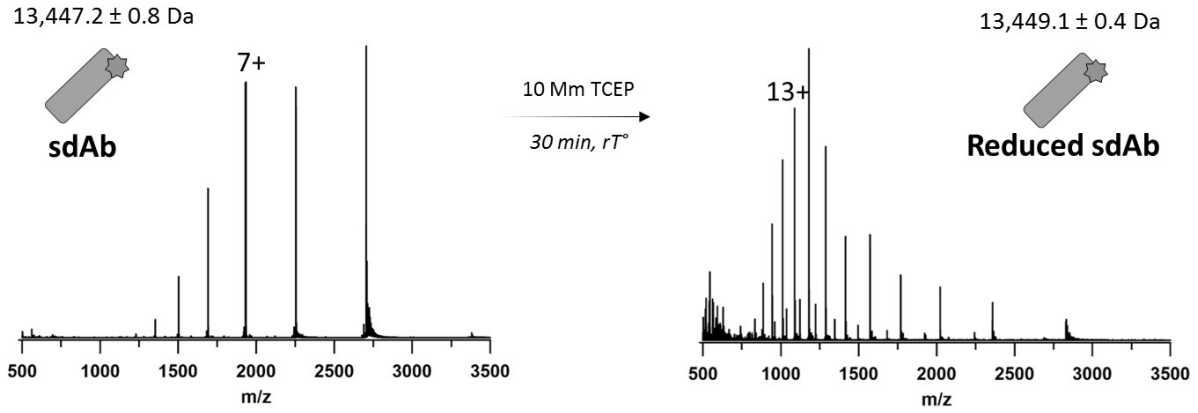


Figure 110: RPLC-MS analysis of anti-EGFR sdAb before and after 10 mM TCEP reduction to unfold the proteins and release additional charge states. The nanobody is depicted by a pink square and the addition of one cysteamine is depicted by a blue star corroborating the addition of 59 Da.

a) Choice of precursor ion

The importance of the ion's charge state prior to every TD-MS experiment was already highlighted in the previous chapter, thus we wanted to evaluate this parameter using the reference sdAb. Three high charge states (1,121.78 *m/z* (12+), 1035.49 *m/z* (13+) and 961.67 *m/z* (14+)) were selected for HCD fragmentation using 30% NCE. The resulted MS/MS spectra from each ion showed significantly different fragmentation patterns. Although the fragmentation of the 14+ ion generated the largest number of fragments, some ions were not informative due to further fragmentation of product ions (internal fragments). The ion charged 13+ generated more informative fragments leading to a maximum sequence coverage of 28% (*versus* only 20% and 27% for 12+ and 14+, respectively) (**Figure 111**).

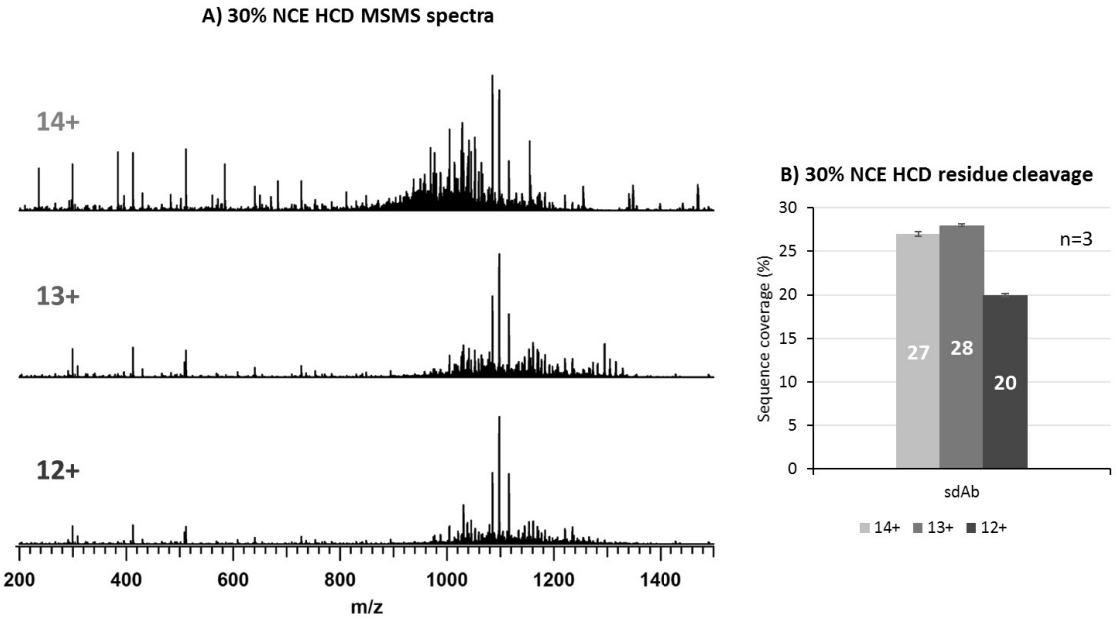


Figure 111: Optimization of the choice of the precursor ion upon 30% nce HCD fragmentation of the reference sdAb. (**A**) MS/MS spectra with the selected precursor ion (left) showing different fragmentation patterns. (**B**) Sequence coverage based on the MS/MS spectra deconvolution.

b) Optimization of the energy (HCD) and reaction time (ETD, UVPD) for sdAb fragmentation

Upon the choice of the precursor ion, the fragmentation technique used could provide significantly different information on the primary structure. We first used HCD [33, 309] as it is the reference and more characterized techniques of fragmentation. We evaluated three different NCE values for the sdAb fragmentation. **Figure 112** shows the obtained spectra upon 10%, 20% and 30% NCE HCD, where a clear difference in the fragmentation patterns is observed. In fact, at lower energies, the precursor ion remains intact because the peptide backbone is barely broken, where higher collision energies appear to re-fragment some ions of interest leading to the observation of fragments at lower *m/z* that probably will not be informative. Indeed, HCD at 20% nce exhibited the largest number of fragment ions leading to a maximum sequence coverage of 30% *versus* only 23% and 28% at 10% nce HCD and 30% nce HCD, respectively.

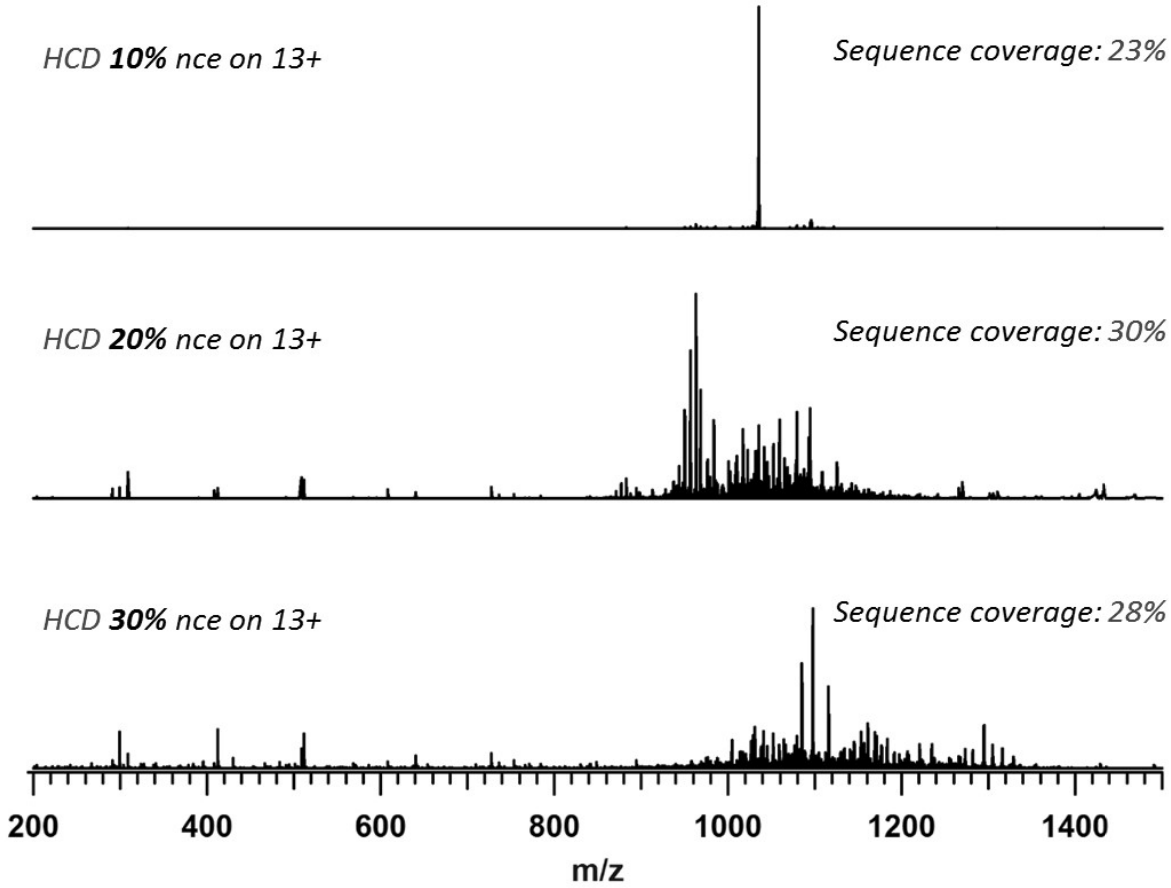


Figure 112: Optimization of the NCE of HCD upon the fragmentation of the precursor ion 13+ of the reference sdAb. The selected nce HCD values are represented in the left side of the spectra and the obtained sequence coverage are at the right.

Despite the limited sequence coverage upon HCD fragmentation, this method provided 13 diagnostic *y*-fragment ions of the position of the cysteamine molecule at the C-terminal side (**Figure 113**) which strengthens our conclusion that the conjugation was performed in the C-terminal side of the sdAb.

Figure 113: Fragmentation map of sdAb upon 20% NCE HCD confirming the cysteamine modification occurred on C-terminal position. The specific fragments to the modification site are depicted with a red circle.

Next, in order to maximize the sequence coverage and gain more confidence in the modification site assessment, we used both ETD [23, 205, 207-208, 222, 242] and UVPD [28, 35, 227, 247-248, 250-251, 312] for the fragmentation of the 13+ charged ion, to reach higher backbone fragmentation. Similarly to HCD optimizations, we evaluated three activation times for ETD (2ms, 4ms and 10ms) and three dissociation

times for UVPD (10ms, 15ms and 25ms) upon the fragmentation of the 13+. At lower reaction times, very few fragmentation pathways are accessed resulting in lower sequence coverage (33% at 2ms ETD and 59% for 10ms UVPD). When increasing the reaction time for both techniques, additional fragment ions are observed, resulting in different fragmentation profiles. However, when the reaction time reaches its maximum the fragments can undergo further fragmentation, leading to lower sequence coverage for both experiments (23% and 58% for 10ms ETD and 20ms UVPD, respectively). The results reached here are similar to the conclusions drawn in the case of HCD, which allowed obtaining a maximum sequence coverage of 47% and 64% for ETD and UVPD, respectively at the optimal conditions (**Figure 114**).

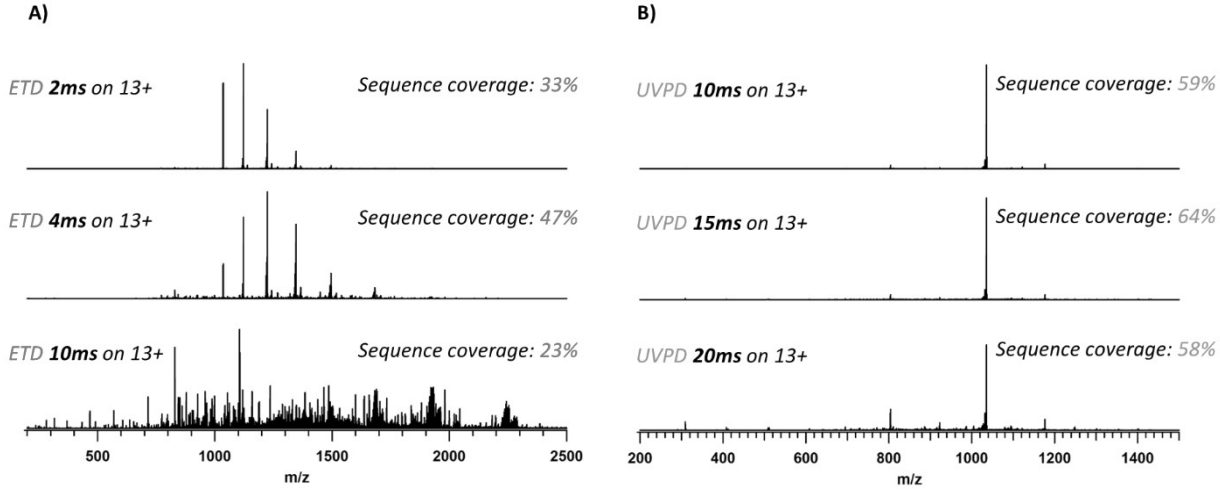


Figure 114: Optimization of **(A)** ETD and **(B)** UVPD reaction time upon the fragmentation of the precursor ion 13+ of the reference sdAb. The selected nce HCD values are represented in the left side of the spectra and the obtained sequence coverage are at right.

c) Complementarity of fragmentation techniques (HCD, ETD and UVPD)

Upon combination of the results obtained *via* HCD, ETD and UVPD fragmentation of the sdAb, we were able to reach up to 88% of the sequence coverage (**Figure 115**). The obtained fragments from each technique, cover different regions of the sequence, with *c/z* fragment ions covering mainly the interior of the amino acid sequence, *b/y* fragment ions located mainly in the N- and C-terminus sides, and *a/x* fragment ions located all along the amino acid sequence (**Figure 115A**). This complementarity is highlighted by the Venn diagram represented in **Figure 115B**. This diagram shows that there are many fragments unique to every techniques (28, 18 and 5 fragments upon HCD, ETD and UVPD alone, respectively), while many fragments are shared between at least two techniques (47 between ETD and HCD, 4 between ETD and HCD and 15 between HCD and UVPD). In addition the three techniques share 3 cleavage sites which highlights the complementarity of the latter dissociations.

More importantly, UVPD shares many fragments with both other techniques (ETD and HCD) as it uses additional similar dissociations pathways than ETD and HCD resulting in more fragments. The complementarity of these techniques is also highlighted by the sequence coverage that reached up to

88% when combining fragments from the three experiments. This sequence coverage was only 30%, 49% and 60% when using individually HCD, ETD and UVPD, respectively.

In more details, UVPD allowed reaching high sequence coverage when used individually, which is mainly due to the small size of the sdAb sequence which is composed of only 125 residues (~15 kDa), and to the reduction of the disulfide bridge that allowed reaching the entire backbone of the nanobody. HCD provided specific fragments that confirm the conjugation site (13 *b* type diagnostic ions that contained the cysteamine modification at C-terminus), despite its low sequence coverage. Finally, ETD provided fragments in the interior region of the sequence coverage, which allowed to increase the confidence in the primary structure and the modification site assessment of the nanobody.

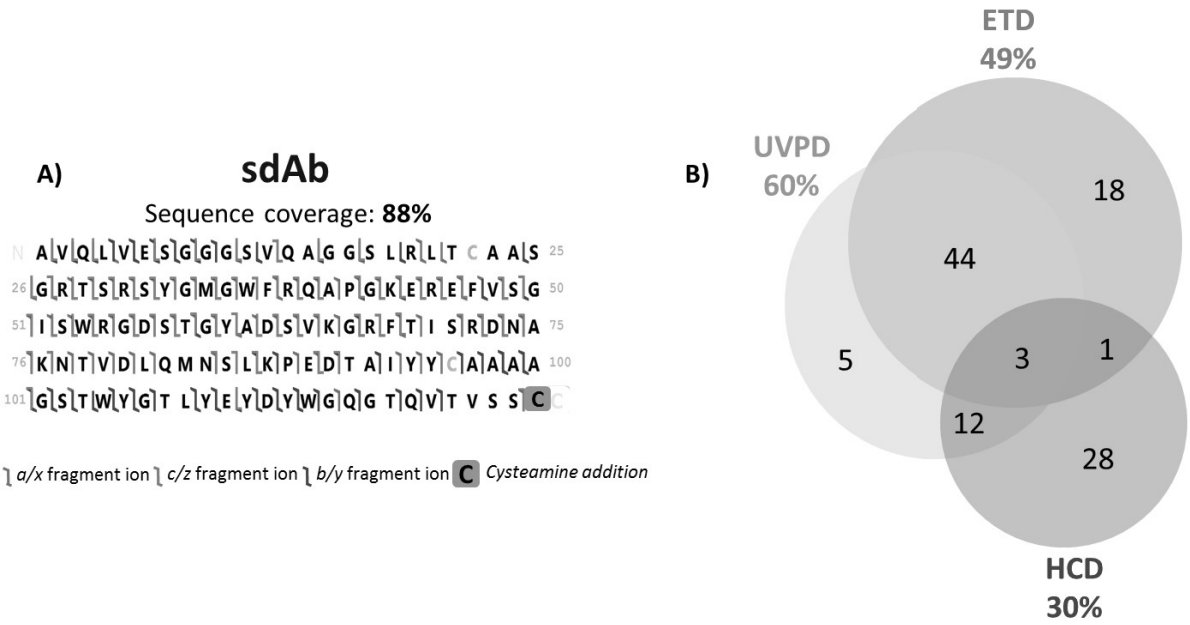


Figure 115: Total sequence coverage of sdAb upon combination of fragments obtained from three dissociation techniques (HCD, ETD and UVPD). (**A**) Fragmentation map showing the localization of the different fragment ion types and showing the localization of the cysteamine addition depicted in a blue square. (**B**) Venn diagram showing the sequence coverage (outside the circles in percentage %) obtained upon each individual technique, with number of shared fragments inside the circles.

In summary, combining the fragment ion data obtained from the three dissociation techniques is a promising strategy to provide a more comprehensive characterization of nanobodies. With this in mind, we followed the same strategy for the characterization of the AF488 sdADC using the previously optimized parameters.

5.3 Top-down MS for the characterization of the primary structure of the sdADC

In a similar way, we first reduced the disulfide bridge of the sdADC upon TCEP mild treatment. The reduced samples exhibited a mass of 14,147.6 ± 0.2 Da, consistent with the reduced nanobody bearing one AF488 payload (+ 698 Da), and in-line with our previous SEC-nMS results (**Figure 116**).

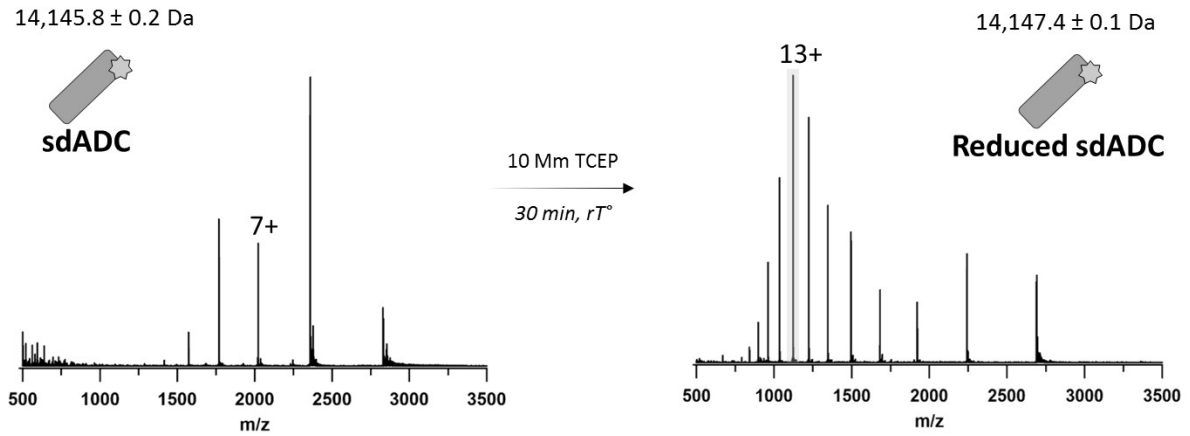


Figure 116: RPLC-MS analysis of anti-EGFR sdADC before and after 10 mM TCEP reduction to unfold the proteins and release additional charge states. The nanobody is depicted by a pink square and the addition of AF488 is depicted by an orange star corroborating the addition of 698 Da. The precursor ion selected for fragmentation is highlighted in orange (13+ charge state).

a) Sequence coverage and conjugation site assessment of the sdADC

The selected parameters upon sdAb TD-MS optimizations were applied for the fragmentation of the 13+ charged precursor ion of the sdADC (1089.27 *m/z*). The resulted MS/MS raw spectra upon the fragmentation of the sdADC using 20% NCE HCD, 4 ms ETD and 213 nm UVPD laser irradiation for 20 ms were dramatically different. Consequently, the different deconvoluted MS/MS spectra matched with the sdADC sequence allowed achieving different sequence coverage of; 35%, 52% and 55% upon HCD, ETD and UVPD fragmentation, respectively, along with assessment of the conjugation site on the C-terminus.

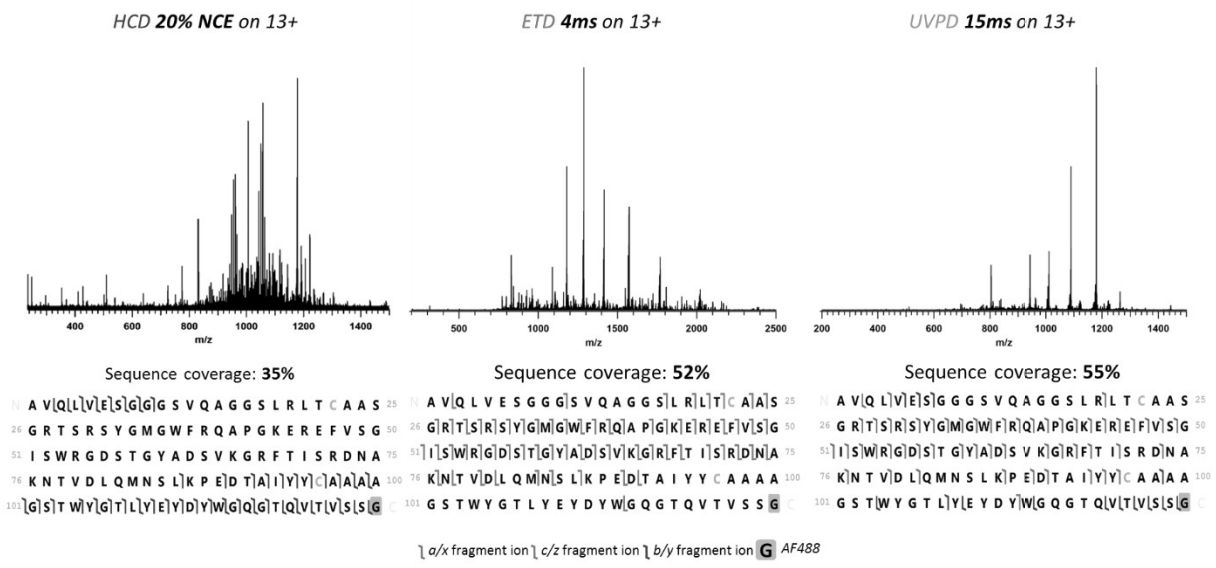


Figure 117: TD-MS experiments of the sdADC. The upper panel represents the obtained MS/MS spectra upon the sdADC fragmentation using 20% NCE HCD, 4 ms ETD and 20 ms UVPD. The lower panel shows the fragmentation map obtained from each single run using one fragmentation technique at a time; HCD, ETD or UVPD fragmentation. *b/y* fragment ions are depicted in blue, *c/z* in red and *a/x* in green. AF488 conjugation is outlined in an orange square. AF488 conjugation site is highlighted in an orange square.

After the addition of the AF488 fluorescent molecule on the C-terminal cysteamine, the results were comparable with sdAb bearing cysteamine only. Thus, in a similar manner and with the objective of confirming the conjugation on the added cysteamine, we first compared the MS/MS spectra obtained from both samples. Indeed, the MS/MS spectra of the sdADC was different from that obtained upon HCD fragmentation of the sdAb with fragment ions observed only upon AF488 conjugation. In the highest mass range window (800-1,400 *m/z*), many fragments were observed in the case of the sdADC, corresponding to the amino acid sequence region V120-S124). Indeed, in these regions many fragments were diagnostic to the AF488 modification, confirming the sdADC conjugation. In addition, in the low mass range (200-900 *m/z*), a fragment ion corresponding to the G125 and the Cysteamine-AF488 cleavage was observed. Those diagnostic fragments, provided confidence in the conjugation confirmation at the C-ter of the nanobody sequence. (**Figure 118**).

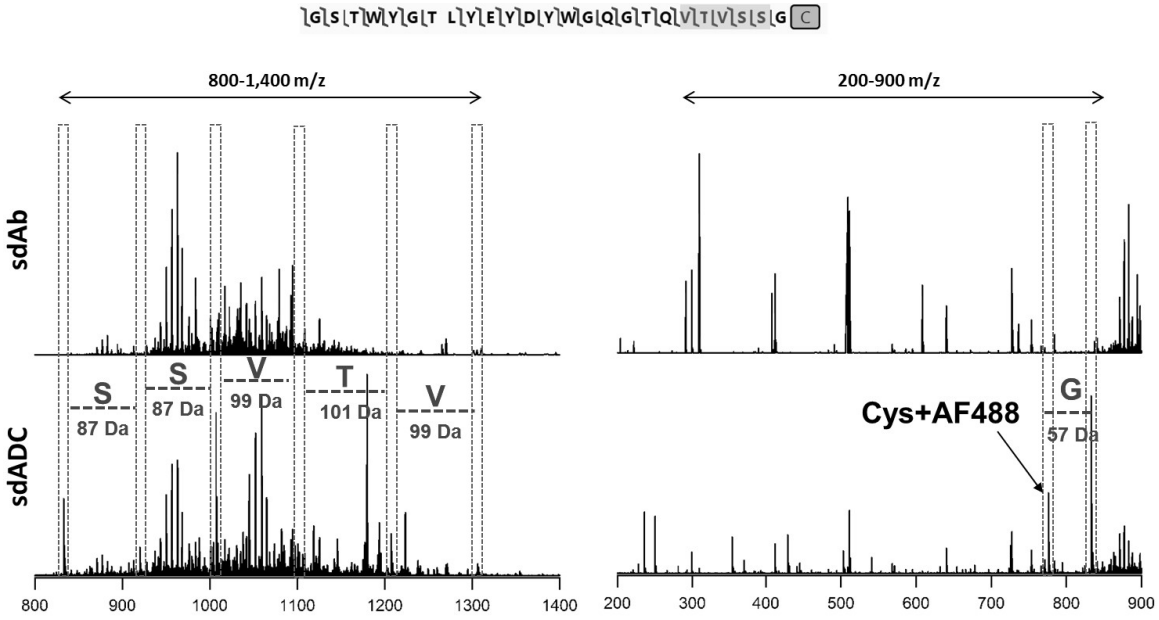


Figure 118: Comparison of the MS/MS spectra upon 20% NCE HCD fragmentation of sdAb (**top**) and sdADC (**bottom**). In the high mass range (800-1,400 *m/z*), five cleavages corresponded to the amino acid region bearing the AF488 conjugation (highlighted in orange in the fragmentation map). In the low mass range (200-900 *m/z*), a fragment ion corresponding to the cysteamine-AF488 cleavage was observed, confirming the conjugation of the nanobody at C-ter glycine.

Moreover, the position of the conjugation on the sdADC was modified to the two cysteine residues at position 22 and 96, respectively, which showed significant differences in the sequence coverage and AF488 specific fragments number (**Figure 119**). Indeed, when assuming the conjugation occurred on the other cysteines, the sequence coverage drastically decreases from 35% (in case of C-ter labelling) to 12% and 6% in case of C96 and C22 labelling, respectively. Furthermore, no diagnostic fragment ions

bearing the AF488 molecule in position 22, and/or 96 were assigned this time, which increases our confidence in locating the conjugation site at the C-ter cysteamine.

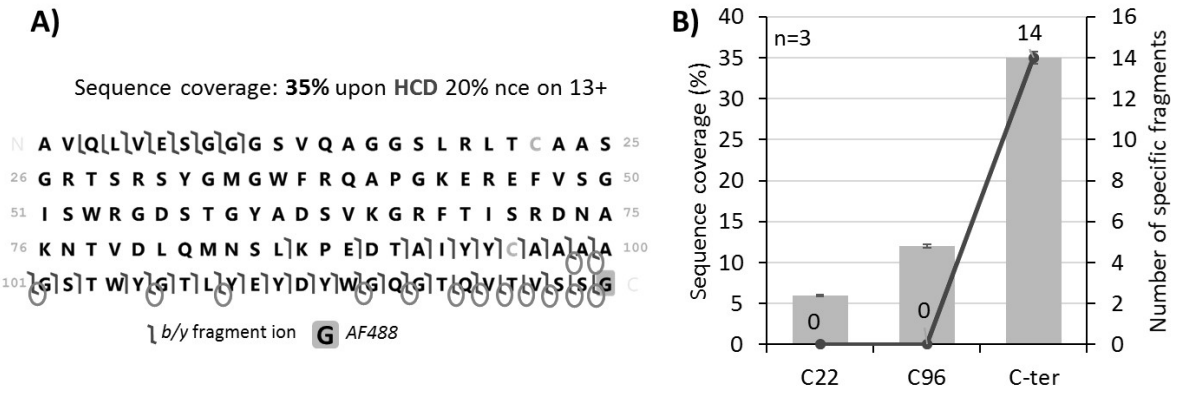


Figure 119: SdADC fragmentation upon 20% nce HCD. (**A**) Fragmentation map of matched *b/y* ions with the sdAb sequence. AF488 modification is outlined in an orange frame. Specific fragments are depicted with red circles. (**B**) Graph representing the sequence coverage with different potential conjugation sites. The numbers on the bars represent the specific fragment to the modification.

In addition, similarly to the sdAb fragmentation, each technique allowed to dissociate the amino acid sequence in specific regions. In more details, HCD mainly provided fragments on the N-terminus of the amino acid sequence allowing us to confirm the site of conjugation, while the fragment ions obtained *via* UVPD and ETD were located throughout different parts of the sequence including the interior region of the sdADC sequence. More importantly, yielded in a large number of fragment ions resulting from different dissociation pathways. (**Figure 120**). Overall, combining the fragment ion data obtained from the three dissociation techniques, resulted in 87% sdADC sequence coverage along with the localization of the AF488 modification at the C-terminus cysteamine. Furthermore, among the 238 fragments matched with the sequence, 16 fragment ions (7% of the sequence) containing the AF488 molecule allowed to confirm that the labelling occurred on the C-terminal cysteamine.

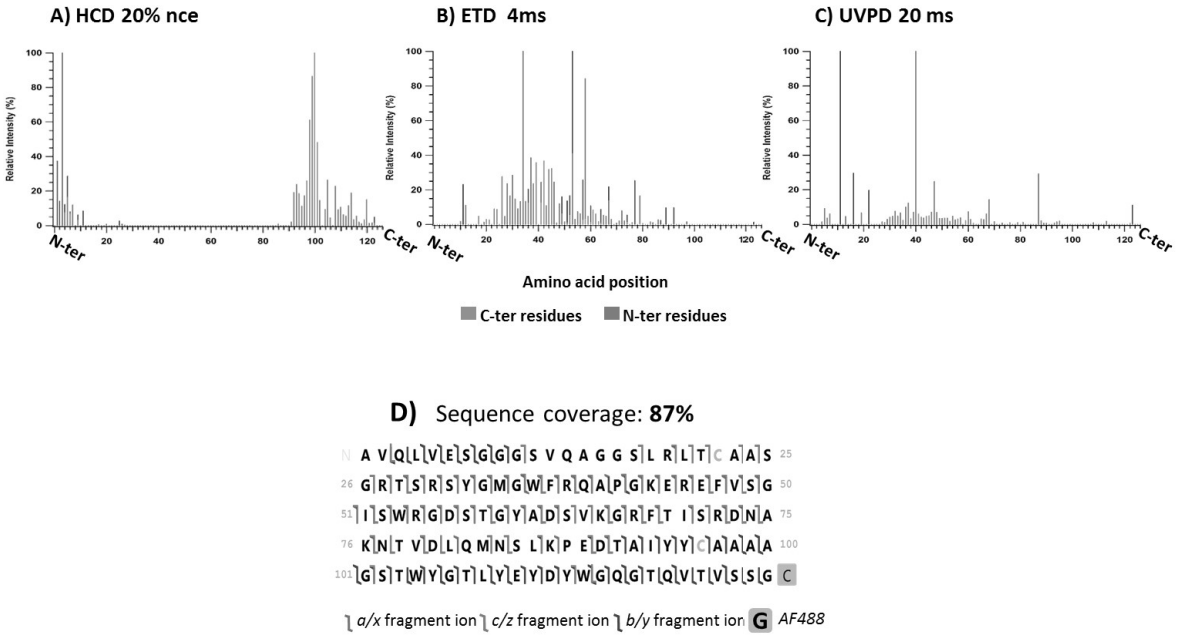


Figure 120: Location of C-terminal and N-terminal residues upon sdADC fragmentation using **(A)** 20% nce HCD, **(B)** 4 ms ETD and **(C)** 20 ms 213 nm UVPD. This schematic shows that HCD fragments are mainly located in the C-ter and N-ter regions of the sdADC. ETD fragments are present in the central region of the nanobody conjugate. Fragments obtained upon UVPD are located all along the sdADC backbone. C-ter residues are represented in red while N-ter are in blue. (**D**) Fragmentation maps of total sequence coverage of the sdADC obtained upon combination of the fragments from 20% nce HCD, 4 ms ETD and 20 ms UVPD. *b/y* fragment ions are depicted in blue, *c/z* in red and *a/x* in green. AF488 conjugation is outlined in an orange square.

Altogether, our TD-MS results upon fragmentation of sdADC illustrated a clear complementarity between HCD, ETD and UVPD for a more comprehensive characterization of the amino acid sequence along with the conjugation site identification.

b) Investigation of the disulfide bonds incorporated in the sdAb and sdADC proteins

Since AF488 covalent labelling is accomplished through a maleimide reaction with a thiol group, potentially, the cysteine residues contained within the sdAb disulfide bond could also be conjugated. Albeit the results afforded in the previous sections clearly show that the conjugation is solely performed on the C-terminal cysteamine, suggesting that the conjugation process does not disrupt the intramolecular disulfide bridge, we aimed at evaluating our TD-MS approach to see if it affords information about the presence of intra-molecular disulfide bridges. To investigate this, we aimed at determining the terminal fragments that contained specifically the cysteines of interest by using traditional collisional activation methods such as HCD, on the unreduced form of the sdADC.

The resulting MS/MS spectrum represented in **Figure 121** was searched for products arising from the regions of interest. Despite the significantly limited sequence coverage (4%), we successfully identified fragments that clearly exhibited the presence of both cysteines in their oxidized form, in addition to

the AF488 modification. The peak intensity and the isotopic distribution were key factors to validate the specific fragments to the modification, thus we selected *m/z* 2,308.88 (y122), *m/z* 2,290.03 (y121), *m/z* 2,273.52 (y120) and *m/z* 2,252.02 (y119) which were clear evidence that the C22 and C96 amino acids were involved in a disulfide bridge and that the C-terminal Glycine incorporated one AF488 payload.

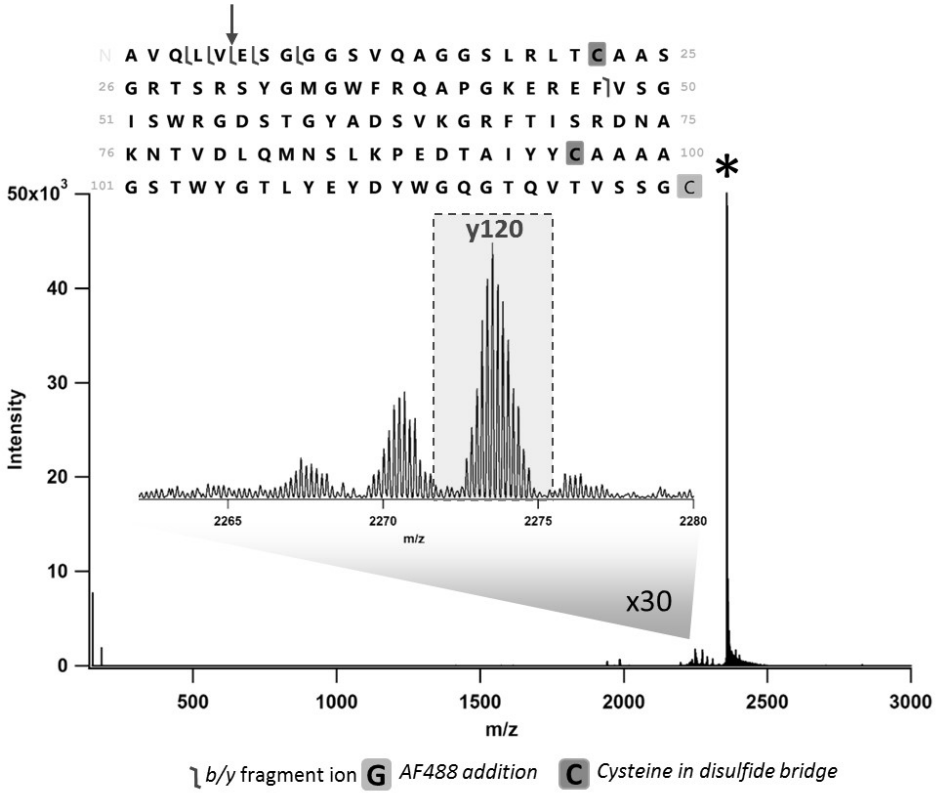


Figure 121: TD-MS experiments of unreduced sdADC showing MS/MS spectra upon 10% nce HCD fragmentation with fragmentation map on top and a zoom on the fragment ion y120. *b/y* fragments are depicted in blue. AF488 modification is outlined in orange frame and hydrogen loss modifications are in grey frames. The precursor ion is depicted with a star.

Interestingly, upon 213 nm UVPD many fragments could be matched to the sequence increasing the sequence coverage to 19%, confirming the multiple pathways of dissociation that UVPD involves. The observed fragment ions were containing either the unreduced C22 exclusively, such as the ion *m/z* 1,243.14 represented in **Figure 122,** or both cysteines C22 and C96 involved in a disulfide linkage: 2,289.88 *m/z.* This pointed out to the existence of the S-S bond.

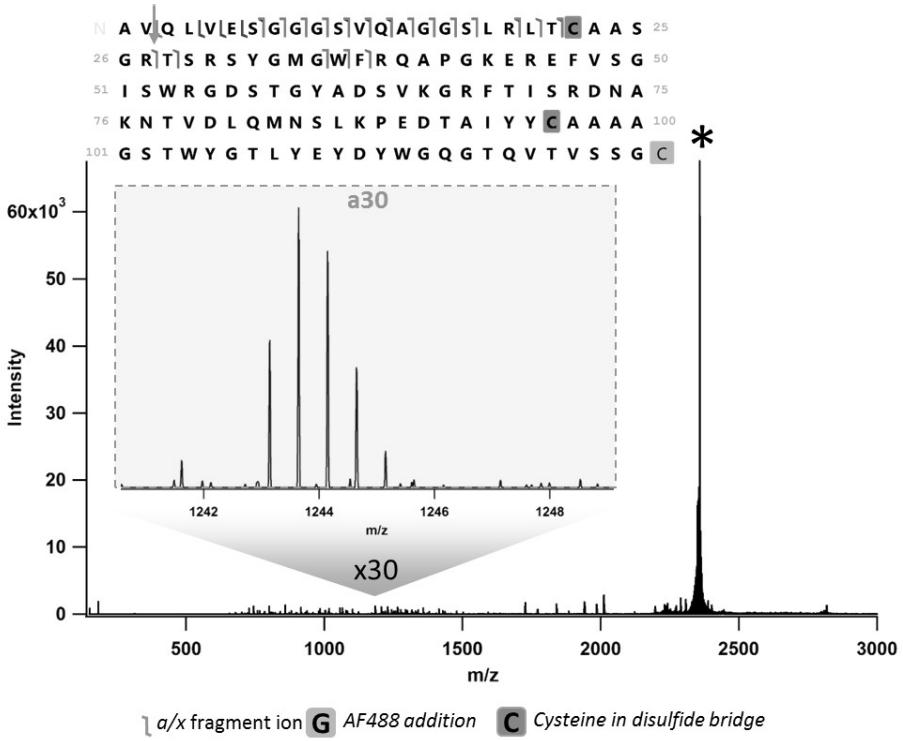


Figure 122: TD-MS experiments of unreduced sdADC showing MS/MS spectra upon 30 ms UVPD fragmentation with fragmentation map on top and a zoom on the fragment ion a30. *a/x* fragments are depicted in green. AF488 modification is outlined in orange frame and hydrogen loss modifications are in grey frames. The precursor ion is depicted with a star.

Assigning the internal fragments resulting from the different dissociation experiments could also be considered [129], to allow comprehensive sequence coverage of proteins containing either the intact or fragmented disulfide linkage. Indeed, after submitting our deconvoluted MS/MS spectra to ClipsMS [229] software, we could allocate many ion products in addition to terminal fragments that were prioritized. In the case of UVPD in particular, inclusion of internal fragments raised the sequence coverage to 35%, mostly due to the access of the interior regions between C22 and C96 that contained 8 fragment ions, including 6 internal fragments incorporating one or both cysteines as depicted in the fragment location map in **Figure 123**. We believe that this drastic increase is mostly due to the multiple possible pathways of internal fragmentation in UVPD that can take place after each reaction, which explains the generation of a plethora of fragment ion types due to the cleavage of many backbone bonds, namely *ax*, *ay*, *az*, *bx*, *by*, *bz*, *cx*, *cy* and *cz* ions. Moreover, the improvement of the sequence coverage led to characterize specific fragments to the disulfide bridge leading to achieving a more comprehensive characterization of our sdADC.

However, as we mentioned in the previous chapter (**Part III, chapter 1**), the risk of false positive when including internal fragments is significantly high. Therefore, these internal fragments should be considered with caution. In this context, the 8 internal fragments that contained the cysteines involved in the sdADC disulfide bridge, were searched in the raw MS/MS spectrum. Indeed, among the six

matched internal fragments, 2 were false positives (ay26 and cx46). Conversely, 4 internal fragments were confirmed in the MS/MS spectra (cy48, cx43, bz40 and az28) and allowed to confirm the linkage of the disulfide bond containing C22 and C96 residues. In fact, those fragments ions have relatively low abundance which explains why they are ignored from classical TD-MS softwares in the first place.

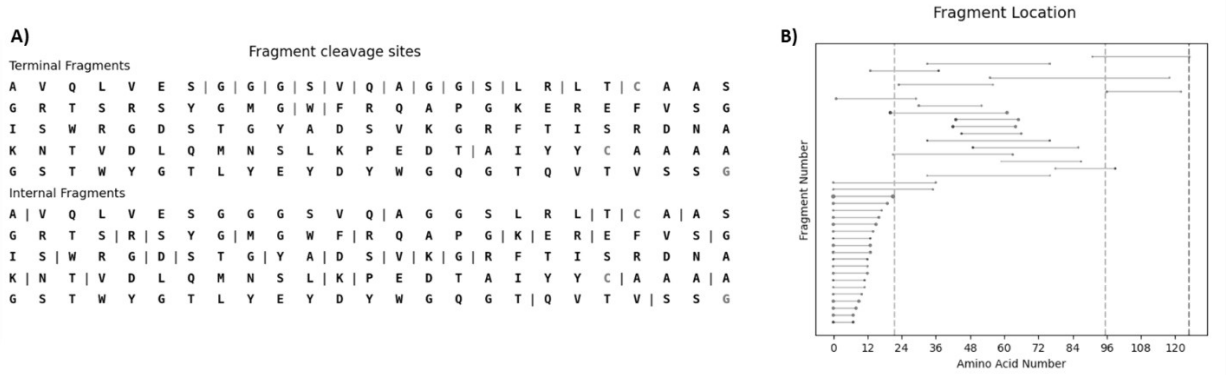


Figure 123: Added value of internal fragments to determine fragments in the S-S bond area. **(A)** Fragment cleavage sites showing terminal and internal fragments. **(B)** Fragment location map representing internal fragments between Cys22 and Cys96 depicted in orange dashed lines, and incorporating the AF488 modification depicted in blue lines.

In summary, the fragmentation of the sdADC in its intact form generates few terminal fragments, resulting in very limited sequence coverage. This is due to the difficulty to reach the core of the nanobody, which is normally maintained by a set of cysteines that form a disulfide bond. To demonstrate this hypothesis we first determined terminal fragments upon HCD fragmentation of the unreduced sdADC. Albeit very limited sequence coverage was obtained (4%), some few fragments were informative of the conservation of the disulfide bond. This information was further confirmed by UVPD that generated a larger number of terminal fragments and increased the confidence in assigning the disulfide bridge (19%). Taking into account the plethora of internal fragments that UVPD generates, we searched for these fragments using ClipsMS for further disulfide bond assessment. As expected, the inclusion of internal fragments increased the sequence coverage up to 35%, which helped confirming the preservation of the disulfide linkage between C22 and C96, along with assessment of the AF488 modification in the C-terminus. However, the inclusion of these internal fragments can increase the risk of false positive and should be considered with caution.

6. Conclusions

The main objective of this chapter was to test the validity of our TD-MS workflow by characterizing our in-lab produced anti-EGFR nanobody conjugate. First, we demonstrated the successful expression and subsequent purification of the anti-EGFR nanobody, along with its conjugation to an AF488 molecule. Thanks to a straightforward SEC-nMS analysis, we could confirm the successful AF488 labelling through precise mass measurement of the conjugated sdADC (avDAR of 1). Subsequently, using our TD-MS approach we determined the specific site of conjugation at the C-ter amino acid by providing a total

sequence coverage of 87%. This high sequence coverage was obtained by combining results from three fragmentation techniques (HCD, ETD and UVPD) which highlighted the complementarity of these activation techniques. Moreover, we confirmed the preservation of the disulfide bridge upon the sdADC by UVPD fragmentation that gives specific fragments including the S-S bond along with the conjugation on the C-ter. In addition, inclusion internal fragments confirmed the preservation of the S-S bond upon manual validation of those fragment ions.

7. Scientific communications

Benazza R., Papadakos G., Diemer H., Cotton G., Cianférani S. and Hernandez-Alba O. **Full characterization of single-domain antibody-drug-conjugates (sdADC) with assessment of modification and conjugation sites using top-down mass spectrometry (TD-MS)**.*Oral communication*. *Journées Club Jeunes de la FPS*. Avril **2023**, Strasbourg, France.

Benazza R., Papadakos G., Cotton G., Hernandez-Alba O. and Cianférani S. **TD-MS approaches for straightforward characterization of an anti-EGFR single-domain antibody-drug conjugates**. *Poster communication. JFSM*. September **2023**, Nantes, France.

Benazza R., Papadakos G., Cotton G., Hernandez-Alba O. and Cianférani S. **Full characterization of sdAb conjugates using top-down mass spectrometry (TD-MS)**. *Poster communication*. Analytics. June **2022**, Nantes, France.

Benazza R., Papadakos G., Diemer H., Cotton G., Cianférani S. and Hernandez-Alba O. **Characterization of anti-EGFR nanobodies using top-down mass spectrometry (TD-MS).** *Article in preparation*.

Conclusions of Part III

In the first chapter of this part, we first evaluated the latest generation Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific) upon TD-MS characterization of reference proteins (Myoglobin 17 kDa and Carbonic anhydrase 26 kDa). Upon optimization of several parameters using Myoglobin, we pinpointed to the importance of the precursor ion choice, which is a crucial step for an efficient fragmentation. Indeed, a highly charged and well intense precursor ion lead to better yield in residue cleavage and thus higher sequence coverage. Next, we demonstrated that each fragmentation technique used (HCD, ETD and UVPD), leads to completely different fragmentation patterns and different fragment ion types, which consequently provide coverage of different regions of the protein sequence. Therefore, the combination of fragment ions resulted from the three dissociation techniques (HCD, ETD and UVPD) allow to reach 99% of the Myoglobin sequence. In fact, yielding 100% of the sequence coverage is not challenging when it comes to low MW proteins. Therefore, we explored this hypothesis through the fragmentation of Carbonic anhydrase (29 kDa) which contains an acetylation at the N-ter residue. Indeed, the fragmentation of Carbonic anhydrase using similar Myoglobin TD-MS conditions, allowed reaching only 64% of the sequence coverage, highlighting the impact of the size on the fragmentation efficiency.

Then, we moved forward to the characterization of a more challenging protein (trastuzumab, 150 kDa), which required a mild digestion/reduction to downsize the intact mAb to ~25 kDa subunits. This required an additional chromatographic dimension, which allowed the separation of the resulted subunits: Fc/2, Lc and Fd subunits. Indeed, the successful outcome of characterizing the intact mAb sequence with localization of the N-glycosylation site was achieved thanks to the fine-tuning of different parameters, again. Our optimizations studies started first with evaluating different energies (for HCD) and reaction times (for ETD and UVPD), to find a compromise between maximum obtained and explained fragments. The results revealed that at lower reaction times or when using lower energies, the precursor ions depletion is not completely achieved resulting in low residue cleavage. Conversely, at higher dissociation times/energies the ions undergo additional fragmentation giving rise to low abundant and non-informative fragment ions accompanied with an increasing number of internal fragment ions. Second, the combination of the three activation techniques at their optimal conditions, allowed covering different regions of the subunits sequence, which enhanced the sequence coverage and increased the confidence in the glycosylation site assignment. Furthermore, we tested several combinations of precursor ions selection; four multiplexed ions, two multiplexed ions or one unique precursor ion selection either one of the most charged or the most intense ones. Through these experiments, we found that the fragmentation relies highly on the choice of the number of precursor ions to be fragmented. Multiplexing several ions lead to lower sequence coverage, due to mainly

acquiring less scans for each ion (5 scans per precursor ion *versus* 20 scans when the ion is selected alone). Moreover, multiplexing leads to the combination of many low-intensity fragment ions that reduces the overall quality of the spectra (lower S/N ratio), which directly affects the deconvolution. Therefore, focusing on one precursor ion can significantly increase the sequence coverage. This specific ion should be at a decent abundance (high intensity, high charge state) to undergo better fragmentation. In the same propensity, widening the isolation window induces a drop in the sequence coverage, as the averaging of the resulted ions reduces the relative intensity of some fragment ions of interest. Overall, our preliminary optimizations could achieve 75%, 78% and 71% sequence coverage obtained for Fc/2, Lc and Fd subunits. Finally, different key points could help improving the sequence coverage. Notably, PTCR reaction performed at the MS^3 level added a remarkable value of simplifying the fragmentation MS/MS spectra (mostly after ETD fragmentation) and thus allowing identifying more fragment ions. Assigning internal fragments in addition to terminal ones could also, help reach to an important information that is usually ignored by common TD-MS softwares, which increased the sequence coverage up to 98%, 100% and 100% for the Fc/2, the Lc and the Fd subunits, respectively. However, we demonstrated through the subunits sequence scrambling, that inclusion of internal fragments could lead to false-positive assignment especially in the case of UVPD. This is due to similarities in masses of different combinations of amino acids, which could not be differentiated due to lack of scoring parameters that reflect the confidence of sequence assignment. We believe that this limitation could be tackled by strengthening the deconvolution and fragment matching parameters, available on classical or newly developed TD-MS softwares. When possible, sometimes it is worth adding an additional fragmentation level (MS^3) to confirm these fragment ions.

Overall, we showed in this study that the achievement of high sequence coverage of a monoclonal antibody is now possible when taking into consideration the variety of tunable MS^n parameters. Even so, in this study we only focused on the use of two conventional fragmentation techniques (HCD and ETD) in comparison of the most cutting-edge dissociation techniques (UVPD). Achieving near complete sequence coverage can be even enhanced by taking advantage of other techniques available on the latest generation Orbitrap instruments, namely EthcD [243] and EtciD that combine the features of electron-driven and collisional dissociations to yield higher sequence coverage. Moreover, combining the capabilities of Orbitrap analyzers with an Omnitrap platform (FasmaTech, Greece) can also yield in efficient proteins fragmentation due to the combination of multi-stage MS/MS methods available on this platform (AI-ECD, AI-UVPD, IRPMD… etc.).

However, among the most challenging structural information to obtain is the conjugation sites on ADCs for instance, which could possibly be provided thanks to fine instrumental tuning or combination of several fragmentation techniques, as showed in a recent study [129]. This ability of TD- and MD-MS

approaches to provide information on the primary sequence of mAb-based products, could be valuable for the comprehensive characterization of multiple therapeutics formats, such as nanobodies.

Therefore, the main objective of the second chapter was to test the validity of our TD-MS workflow by characterizing our in-lab produced anti-EGFR nanobody conjugate. First, we demonstrated the successful expression and subsequent purification of the anti-EGFR nanobody. This nanobody was strategically engineered to include a free thiol group at its C-terminus, which facilitated further AF488 conjugation through a maleimide reaction. Thanks to a straightforward SEC -nMS analysis, we could confirm the successful AF488 labelling through precise mass measurement of the conjugated sdADC. The drug load distribution showed a homogeneous distribution of D1 species that could allow us to calculate an avDAR value of 1. Subsequently, using our TD-MS approach we aimed at determining the specific site of this conjugation, along with assessment of the disulfide bond, through direct fragmentation using mainly three fragmentation techniques (HCD, ETD and UVPD). The first experiments performed on a reduced form of the sdADC (and sdAb as a reference), aimed at obtaining a maximum sequence coverage to localize the AF488 position, after denaturation of the compact form of the nanobody. Indeed, this step allowed releasing a broader range of charge states that we evaluated using HCD fragmentation. Our results revealed that the choice of the precursor ion is extremely important to enhance the fragmentation efficiency. The latter should be highly intense and well charged to energetically collide with the neutral gas molecules (in this case, using HCD) and thus dissociate the backbone of the nanobody. Bearing in mind that ETD fragmentation depends also on the high charge state of the precursor ion and that UVPD has no preference for the charge state or intensity of the latter, the 13+ seemed to be a good compromise for our further optimizations. Moreover, the fragmentations performed using HCD provided valuable information on the site of conjugation, where we obtained several diagnostic fragments of the cysteamine and AF488 sites, despite the limited sequence coverage that this technique offers. To enhance the nanobodies sequence coverage, we took advantage of ETD and UVPD techniques available on the Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific). Both techniques generated different and various fragments ions that were located in diverse regions of the nanobody sequence. For instance, ETD favored the interior region fragmentation in contrary to HCD that fragmented the nanobody in its C- and N-ter regions. Whereas UVPD provided fragments ions that were located all along the peptide backbone. This study highlighted the complementarity of these activation techniques, which allowed to achieve 87% of the sdADC (and 88% for the reference sdAb) sequence coverage with more confidence in the AF488 localization site assessment.

In fact, to confirm that our conjugation was at the C-terminus with the disulfide bond (between C22 and C96) was preserved, we sought to characterize our sdADC at its unreduced form. Despite the

limited sequence coverage that HCD provided upon fragmentation of the 7+, some fragments were overall good intensity and good isotopic distribution, could confirm the involvement of C22 and C96 in a disulfide linkage. This hypothesis was supported by UVPD data, which showed higher sequence coverage with more terminal fragment bearing the disulfide bridge along with the AF488 conjugation. However, a large number of the obtained fragments *via* UVPD were mainly internal ones that were explained further using ClipsMS algorithm. Inclusion of these internal fragments was a key step to confirm the conservation of the disulfide bridge in the core of the nanobody, which additionally contained a labelled AF488 in its C-terminus. However, we must bear in mind that the automatic assignment of these small internal fragments comes with a big risk of false positive even when using low mass tolerance (< 2 ppm), therefore a manual validation is required. The obtained fragments could be either directly checked on the MS/MS spectra, which is far more complicated when characterizing larger proteins (mAbs, ADCs…), or upon matching those fragments with a scrambled sequence [324]. Overall, whilst our study seems to be one of the first TD-MS experiments performed on nanobodies, it paves the way to several optimizations and improvements. Among the main challenges for any top down workflow, disulfide bonds seem to challenge the fragmentation efficiency by being inaccessible with the classical dissociation techniques. This limitation could be tackled by using additional activation of the protein of interest along with its fragmentation, as for instance infrared photoactivation during ETD reaction (AI-ETD) [26, 240-242]. This technique has outperformed conventional ETD and HCD fragmentation and improved the sequence coverage of ~29 kDa proteins. In addition, hybrid techniques that combine electron-based and photon-based activations (ETuvPD [252], ECuvPD [224]) could probably boost the sequence coverage of disulfide bond-containing proteins, by accessing the regions that are constrained by this linkage. Therefore, we believe that our promising results give a valuable insight and lead the way to more optimizations for a comprehensive characterization of nanobody therapeutics.

General conclusions and perspectives

General conclusions and perspectives

The aim of my PhD work was to develop tailored and straightforward MS-based techniques for the characterization of a specific class of therapeutic proteins: ADCs/PDCs. My research was conducted on state-of-the-art mass spectrometers, with the aim of evaluating these platforms and developing dedicated methodologies for biotherapeutics assessment. The results obtained in the frame of my PhD work contributed to discuss the following questions:

Are native MS methods mature enough to enter biopharmaceutical companies for routine analysis?

One of the central aspects described in my thesis is the evaluation of a new benchtop mass spectrometer (The BioAccord, Waters) to develop tailored LC-MS approaches for biotherapeutics characterization. From an instrumental standpoint, this platform is indeed user-friendly and allows going from routine quality control analyses to thorough method development. For instance, I demonstrated the possibility to perform SEC-nMS [124] and CEX-nMS [174] experiments through reference mAbs characterization, which allowed highlighting the robustness, sensitivity and reproducibility of the instrument under native conditions. In particular, I showed that the automated SEC-nMS method is important for monitoring the development of new ADCs/PDCs upon a site-specific reaction (Ugi multicomponent reaction [32]). From a data processing point of view, this platform offers detailed reports comprising accurate mass measurements, and straightforward assessment of the different CQAs (size variants, D0, avDAR and DLD).

However, the bottlenecks of this platform rely mainly on the limited *m/z* range (until 7000 *m/z*) which I highlighted through the study of different bsAb[299] formats. Compared to other ToF analyzers where the *m/z* range is extended up to 10,000 m/z and more, using the BioAccord platform large MW could not be identified (such as adeno associated viruses, AAVs [344] and antibody-nanoparticle conjugates [345], ANCs), and HMWS that can be formed upon therapeutic mAbs aggregation [165]. This limitation of the BioAccord will be certainly overcome in the near future to widen the application scope of this benchtop LC-MS instrument. Nevertheless, I showed that MP [133], is a good alternative to nMS as it provides extremely fast measurements (< 1min) in a wide mass range (from 30kDa to 5MDa) using less concentrated samples (100 pm – 100 nM). Thus, this technique could be more involved in biotherapeutics characterization projects. Nevertheless, this technique lacks of detailed structural and l information, a limitation that I anticipate to be tackled in the near future. Moreover, it is still limited in term of resolution compared to nMS. Overall, even if I showed that the BioAccord is a suitable platform for straightforward characterization of newly developed bioconjugates, as well as for method

development of more tailored LC-MS methods for their characterization, I still anticipate that its capabilities need to be improved.

Perspectives

Upon the ability of characterizing mAb complexes, the biopharma field is more and more interested in the characterization of gene therapy products such as AAVs [344, 346], and in other immunoconjugates such as ANCs [345], in native conditions. These biomolecules reach higher mass ranges (> 200 kDa) and thus their characterization using conventional nMS is challenging. MP would be a valuable tool for their straightforward assessment and thus could allow to assess the heterogeneity and the macromolecular assemblies of AAVs and ANCs, which will make it soon the go-to approach for straightforward biotherapeutics characterization upon slight instrumental developments. Another track for characterizing those complexes, could be taking benefit of the charge detection MS (CDMS) approach, which allows determination of single ions on various mass analyzers [347]. Although this technique is major breakthrough in the analytical field, it is still in its infancy and could be improved in order to be applied on complex molecules. One example of the evolution of this technique is the recent publication of Bones and coworkers where SEC separation was coupled to CDMS [348]. Beyond improving the resolution to detect those single ions, I anticipate that in the near future it would be taking benefits from coupling with other separation techniques, and from improvement of straightforward data processing softwares [349].

At this stage, and to answer the former question, I believe that nMS is certainly ready to be imbedded in biopharmaceutical laboratories for the characterization of biotherapeutics. Obviously, biopharma companies are already adopting nMS workflows for mAb characterization for instance, but the need for expertise and extensive method development is still required for more challenging products.

Are TD/MD-MS methods valuable alternatives to peptide mapping for mAb-based product primary sequence assessment?

The expanding need to characterize new bioconjugates formats in a prompt manner, and to replace more time-consuming approaches such as peptide mapping, was supported by the development of cutting-edge mass spectrometers. I had the opportunity to evaluate one of these instruments: the Orbitrap Eclipse™ Tribrid™ MS (Thermo Fisher Scientific), for the characterization of the primary structure of mAb-based formats. The evaluation of the instrument through TD-MS analysis of reference proteins (Myoglobin ~17 kDa, carbonic anhydrase ~26 kDa) and a reference mAb (trastuzumab ~150 kDa) allowed to pinpoint the importance of different parameters optimization, in particular the choice of the precursor ion and the choice of the energy and/or the time of dissociation. While for proteins with molecular weights lower than 30 kDa TD-MS is relatively easy, for mAbs it is

still challenging due to their size, the presence of multiple intra-, and inter-chain disulfide bridges and the presence of multiple chains. These structural characteristics prompted me to opt for a MD-MS approach instead. Upon the optimization of the MS/MS parameters, I showed the benefits of PTCR [254, 256] for spectral decongestion, which allowed improving the subunits (Fc/2, Lc and Fd) sequence coverage and increasing the confidence in the identification of the glycosylation site. The overall optimizations allowed to increase the sequence coverage up to ~70% for the subunits, in less time (~15 min) compared to peptide mapping that are usually time-consuming due to the extensive digestion/alkylation process.

This high sequence coverage along with the PTM localization, was obtained upon the combination of fragments from different fragmentation techniques (HCD, ETD and UVPD in this case), which allows to highlight the complementarity of these techniques.

In terms of data processing, TD/MD-MS is more challenging as it does not offer many tailored softwares for MS/MS spectral interpretation and fragments/sequence matching. In addition, there are not general agreement on the used data processing parameters. In this context, I evaluated a new TD-MS software that takes into consideration the internal fragments (ClipsMS) [229]. The inclusion of internal fragments allowed increasing dramatically the sequence coverage (~100% for all subunits). However, I could confirm that a large number of internal fragments were false-positives due the similarities in masses of different combinations of amino acids (isobaric fragments) and due to the absence of scoring systems.

This overall optimized TD-MS workflow was useful for the assessment of a conjugated nanobody that I produced myself in a secondment at Almac Discovery, in Edinburg. On the one hand, it allowed me to confirm the conjugation site of the sdADC, through nearly complete sequence coverage (87%) and specific fragments localization near the conjugation site region. On the other hand, the data from TD-MS could be correlated to those obtained from SEC-nMS confirming the conjugation one molecule (avDAR of 1), a correlation that cannot not be obtained through peptide mapping. Moreover, TD-MS was a valuable approach for assessing the disulfide bond among the sdADC, through consideration of terminal and internal fragments (validated manually). However, in TD-MS the fragmentation spectra are complex especially for large intact proteins, which requires more appropriate handling and more softwares that are powerful.

Perspectives

While I demonstrated the clear benefits of state-of-the-art mass spectrometers that comprise several fragmentation techniques, I also highlighted the fact that a substantial portion of the fragmentation spectra remains undisclosed, which limits yielding complete sequence coverage. Indeed, a major bottleneck of the TD-MS field is the lack of tailored softwares for processing the data from complex

and large MW biomolecules. In fact, there are multiple softwares that assign terminal fragments [257], but still produce inaccuracies upon miss-deconvolution and miss-interpretation of overlapped species rising from the dissociation of more complex biomolecules. To date, there is only one software that deals with internal fragments assignments (ClipsMS) [229], but lacks robust scoring systems which require time-consuming manual validation of these fragments.

I expect that this point will be the main focus of the TD-MS field in the near future, as the interest increase of the multinational Consortium for Top-Down Proteomics (CTDP) [131] is concentrated in the nTDMS [259] characterization of more complex and challenging biomolecules. In the context of nTDMS, which is by far the most attractive technique to extensively characterize bimolecular complexes in their intact form, novel technologies could be useful. The recently developed novel ion trap (Omnitrap, Fasmatech, Greece) [350] that allow multidimensional MS^n fragmentation thanks to the incorporation of multiple fragmentation techniques within one MS platform, has demonstrated fascinating capabilities of complete cleavage of unfolded proteins, with successful detection either on an orbitrap or ToF analyzer [351-352]. Thus, I anticipate that the use of this platform will be further strengthened for the characterization of biotherapeutics. A last point that I would like to briefly discuss is the potential use of the new high sensitivity, high resolution and high mass accuracy mass analyzers for nTDMS, such as the recently developed SELECT SERIES MRT (Waters) [353] and The Orbitrap Astral MS (Thermo Fisher) [354]. The high sensitivity and high resolution thanks to the nearly lossless ion transfer could result in better identification of fragment ions and consequently yield in high sequence coverage.

Experimental section

Experimental section

This part describes the used instrumentations, protein sequences, sample preparation, data acquisition and data processing methods.

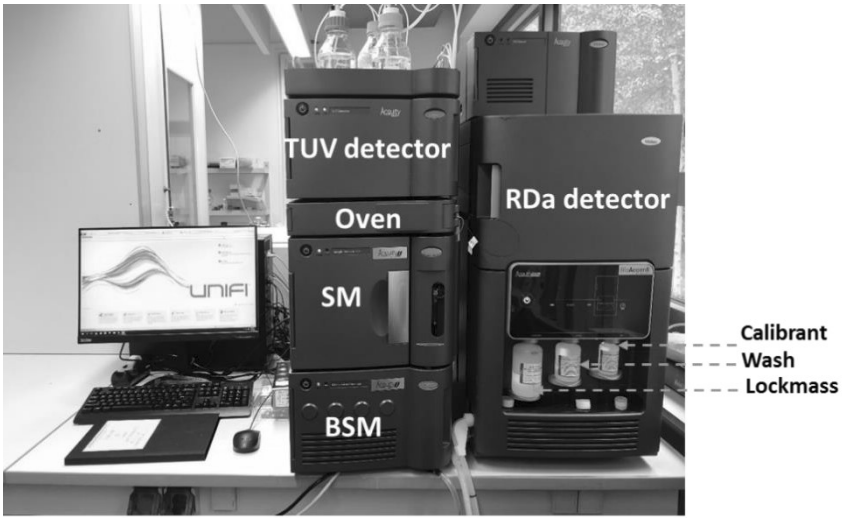
1. Instrumentation

1.1 The BioAccord LC-MS system (Waters, UK)

The BioAccord LC-MS system is a continuous coupling of an Acquity UPLC M-Class liquid chromatography system that includes, a binary solvent manager (BSM), a sample manager (SM), a column oven and a tunable UV detector (TUV detector), continuously coupled to an RDa™ time-of-flight (ToF) mass spectrometer with ESI source. Of note, the sample loop can inject a maximum volume of 10 µL. The whole system is fully controlled by UNIFI software v3.1.0.16 from data acquisition to data processing (Waters, UK).

The mass spectrometer is calibrated prior to every native or denaturing MS analysis, in the 400-7000 m/z range in the positive mode using a solution containing, 50 ng/µL of sodium iodide in isopropanol/water (80/20 v/v) and 0.5 ng/µL of rubidium iodide in isopropanol/water (80/20 v/v). A LockMass solution containing 3.75 ng/µL of leucine encephalin, 12.5 ng/µL of caffeine and 2.5 ng/µL of 1-pentanesulfonic acid in ACN/water (80/20 v/v) is injected automatically prior to each analysis.

This platform was used for the SEC-nMS and CEX-nMS analyses of **Part II chapter 1**, the SEC-nMS analyses of **Part II chapter 2**, the SEC-nMS analyses of **Part II chapter 3**, the dSEC-MS analyses of **Part II chapter 4** and for the SEC-nMS analysis of **Part III chapter 2.**

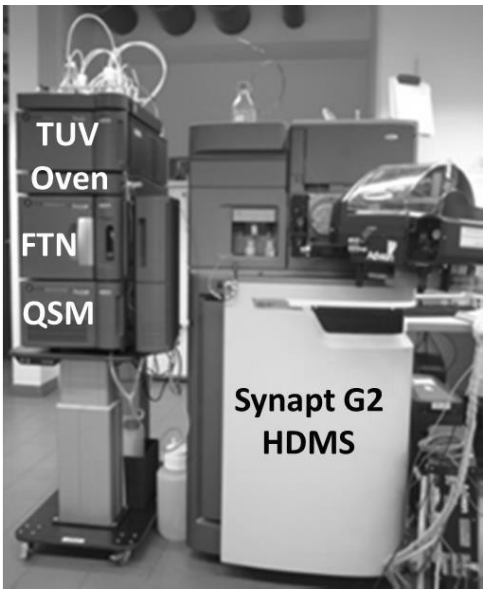


Picture of the BioAccord LC-MS system (Waters, UK)

1.2 The Synapt G2 HDMS (Waters, UK)

The Acquity H-Class composed of a quaternary pump (QSM), a sample manager (FTN sample manager), a column oven and a TUV detector was coupled to a Synapt G2 HDMS is a time-of-flight mass spectrometer. It was used for the analysis of reference mAbs and ADCs in **Part II chapter 1** and for the analysis of bispecifics in **Part II chapter 3**. Of note, the sample loop can inject a maximum volume of 20 µL.

The instrument was calibrated and tuned using a 2 g/L solution of cesium iodide in 2-propanol/water (50/50 v/v), and then was operated in sensitivity mode using a + 1.5 kV capillary voltage. The acquisitions were recorded in the 1000-5,000 m/z range with a 4 s scan time. MS data processing was performed using Mass Lynx V4.1 (Waters, Manchester, UK).



Picture of the Acquity H-Class system coupled to the Synapt G2 HDMS (Waters, UK)

1.3 The Orbitrap™ Eclipse™ Tribrid™ MS (Thermo Fisher Scientific, USA)

The Orbitrap™ Eclipse™ Tribrid™ MS (Thermo Fisher Scientific, USA) is an MS equipped with three analyzers; ion trap, quadrupole and orbitrap. It's also equipped with 4 fragmentation techniques: CID, HCD, ETD and UVPD along with PTCR technique. The ETD and PTCR are injected in the source and the UVPD laser is located at the back of the spectrometer. The MS was used directly upon ESI injection of samples analyzed in **Part III chapter 1**. An Ultimate 3000 LC including a pump module, a column oven and an autosampler was coupled to the MS for the analyses in **Part III chapter 2.** The sample loop can inject until 20 µL. The system was calibrated using a Pierce™ Flexmix™ calibration solution at a flowrate of 5 µL/min, using an HESI source and a low flow needle.

Picture of the the Ultimate™ 3000 LC coupled to The Orbitrap™ Eclipse™ Tribrid™ MS (Thermo Fisher Scientific, USA)

1.4 TWO MP (Refeyn Ltd, UK)

The mass photometer used for analysis of bsAbs in **Part II chapter 3** was a Two MP from Refeyn Ltd, UK.

Picture of the two MP (Refeyn Ltd, UK)

2. Protein sequences

Sample	Amino acid sequence	Theoritical mass (Da)
Anticalin D11vs	QDSTSDLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDKDPGKMFATI YELKEDKSYNVTYVWSGQKKCMYSIVTFVPGSQPGEFTLGNIKSAPGRTSWLVRVVS TNYNQHAMVFFKSVTQNREGFAITLYGRTKELTSELKENFIRFSKSLGLPENHIVFPVP IDQCIDGSAWSHPQFEK	21,302
Anticalin (K46R)-D11vs	QDSTSDLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDRDPGKMFATI YELKEDKSYNVTYVWSGQKKCMYSIVTFVPGSQPGEFTLGNIKSAPGRTSWLVRVVS TNYNQHAMVFFKSVTQNREGFAITLYGRTKELTSELKENFIRFSKSLGLPENHIVFPVP IDQCIDGSAWSHPQFEK	21,330
Anticalin (Q1-D6)-D11vs	NLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDKDPGKMFATIYELKE DKSYNVTYVWSGQKKCMYSIVTFVPGSQPGEFTLGNIKSAPGRTSWLVRVVSTNYN QHAMVFFKSVTQNREGFAITLYGRTKELTSELKENFIRFSKSLGLPENHIVFPVPIDQCI DGSAWSHPQFEK	20,782
Anticalin (Q1-D6)-(K46R)-D11vs	NLIPAPPLSKVPLQQNFQDNQFHGKWYVVGRAGNTGLREDRDPGKMFATIYELKE DKSYNVTYVWSGQKKCMYSIVTFVPGSQPGEFTLGNIKSAPGRTSWLVRVVSTNYN QHAMVFFKSVTQNREGFAITLYGRTKELTSELKENFIRFSKSLGLPENHIVFPVPIDQCI DGSAWSHPQFEK	20,810
AntiEGFR sdAb	MAVQLVESGGGSVQAGGSLRLTCAASGRTSRSYGMGWFRQAPGKEREFVSGISW RGDSTGYADSVKGRFTISRDNAKNTVDLQMNSLKPEDTAIYYCAAAAGSTWYGTLY EYDYWGQGTQVTVSSG	13,521
Atezolizumab	**LC**:DIQMTQSPSSLSASVGDRVTITCRASQDVSTAVAWYQQKPGKAPKLLIYSASFLY SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYLYHPATFGQGTKVEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	144,356 (G0F)2
	HC:EVQLVESGGGLVQPGGSLRLSCAASGFTFSDSWIHWVRQAPGKGLEWVAWIS PYGGSTYYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARRHWPGGFDY WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYASTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	
Bevacizumab	**LC**:DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSLHS GVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYSTVPWTFGQGTKVEIKRTVAAPS VFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	149,199 (0GF)2

	HC :EVQLVESGGGLVQPGGSLRLSCAASGYTFTNYGMNWVRQAPGKGLEWVGWI NTYTGEPTYAADFKRRFTFSLDTSKSTAYLQMNSLRAEDTAVYYCAKYPHYYGSSHW YFDVWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEP KSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF NWYVDGVEVHNAKTKPREEQYJSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA PIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE NNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLS PG	
Carbonic anhydrase	SHHWGYGKHNGPEHWHKDFPIANGERQSPVDIDTKAVVQDPALKPLALVYGEATS RRMVNNGHSFNVEYDDSQDKAVLKDGPLTGTYRLVQFHFHWGSSDDQGSEHTVD RKKYAAELHLVHWNTKYGDFGTAAQQPDGLAVVGVFLKVGDANPALQKVLDALDS IKTKGKSTDFPNFDPGSLLPNVLNYWTYPGSLTTPPLLESVTWIVLKEPISVSSQQMLK FRTLNFNAEGEPELLMLANWRPAQPLKNRQVRGFPK	28,981
Durvalumab	**LC**:EIVLTQSPGTLSLSPGERATLSCRASQRVSSSYLAWYQQKPGQAPRLLIYDASSRA TGIP DRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGSLPWTFGQGTKVEIKRTVAAPSVFI FPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTY SLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	148,973 (G0F)2
	HC:EVQLVESGGGLVQPGGSLRLSCAASGFTFSRYWMSWVRQAPGKGLEWVANIK QDGSEKYYVDSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCAREGGWFGELAF DYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPK SCDKTHTCPPCPAPEFEGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPASI EKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK	
Ipilimumab	**LC**:EIVLTQSPGTLSLSPGERATLSCRASQSVGSSYLAWYQQKPGQAPRLLIYGAFSRA TGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCQQYGSSPWTFGQGTKVEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	147,992 (G0F)2
	HC:QVQLVESGGGVVQPGRSLRLSCAASGFTFSSYTMHWVRQAPGKGLEWVTFISY DGNNKYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAIYYCARTGWLGPFDYW GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALT SGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDK THTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV DGVEVHNAKTKPREEQYBSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTI SKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT TPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG	

Myoglobin	GLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDKFKHLKTEAEM KASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPIKYLEFISDAIIHVL HSKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFQG	16,940
Nivolumab	**LC**:EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDASNRAT GIPA RFSGSGSGTDFTLTISSLEPEDFAVYYCQQSSNWPRTFGQGTKVEIKRTVAAPSVFIFP PDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLS STLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	146,240 (G0F)2
	HC:QVQLVESGGGVVQPGRSLRLDCKASGITFSNSGMHWVRQAPGKGLEWVAVI WYDGSKRYYADSVKGRFTISRDNSKNTLFLQMNSLRAEDTAVYYCATNDDYWGQG TLVTVSSASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVH TFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRVESKYGPPCPPC PAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFBSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQP REPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD SDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLG	
OKT3	**LC**:QIVLTQSPAIMSASPGEKVTMTCSASSSVSYMNWYQQKSGTSPKRWIYDTSKLA SGVPAHFRGSGSGTSYSLTISGMEAEDAATYYCQQWSSNPFTFGSGTKLEINRADTA PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSK DSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC	146,189
	HC:QVQLQQSGAELARPGASVKMSCKASGYTFTRYTMHWVKQRPGQGLEWIGYI NPSRGYTNYNQKFKDKATLTTDKSSSTAYMQLSSLTSEDSAVYYCARYYDDHYCLDY WGQGTTLTVSSAKTTAPSVYPLAPVCGGTTGSSVTLGCLVKGYFPEPVTLTWNSGSL SSGVHTFPAVLQSDLYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVDKKIEPRPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	
Panitumumab	**LC**:DIQMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKAPKLLIYDASNLE TGVPSRFSGSGSGTDFTFTISSLQPEDIATYFCQHFDHLPLAFGGGTKVEIKRTVAAPS VFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	147,094 (G0F)
	HC:VQLQESGPGLVKPSETLSLTCTVSGGSVSSGDYYWTWIRQSPGKGLEWIGHIYY SGNTNYNPSLKSRLTISIDTSKTQFSLKLSSVTAADTAIYYCVRDRVTGAFDIWGQGT MVTVSSASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVH TFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTCNVDHKPSNTKVDKTVERKCCVECPPC PAPPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVHN AKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQP REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLD SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG	

Pertuzumab	**LC**:DIQMTQSPSSLSASVGDRVTITCKASQDVSIGVAWYQQKPGKAPKLLIYSASYRY TGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQYYIYPYTFGQGTKVEIKRTVAAPS VFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	148,102 (G0F)2
	HC:EVQLVESGGGLVQPGGSLRLSCAASGFTFTDYTMDWVRQAPGKGLEWVADV NPNSGGSIYNQRFKGRFTLSVDRSKNTLYLQMNSLRAEDTAVYYCARNLGPSFYFDY WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG	
Ramucirumab	**LC**:DIQMTQSPSSVSASIGDRVTITCRASQGIDNWLGWYQQKPGKAPKLLIYDASNL DTGVPSRFSGSGSGTYFTLTISSLQAEDFAVYFCQQAKAFPPTFGGGTKVDIKGTVAA PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	144,264 (deglycosylated)
	HC:EVQLVQSGGGLVKPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSSISS SSSYIYYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARVTDAFDIWGQGT MVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVH TFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE VHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	
Rituximab	**LC**:QIVLSQSPAILSASPGEKVTMTCRASSSVSYIHWFQQKPGSSPKPWIYATSNLAS GVPVRFSGSGSGTSYSLTISRVEAEDAATYYCQQWTSNPPTFGGGTKLEIKRTVAAPS VFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	147,075 (G0F)2
	HC:QVQLQQPGAELVKPGASVKMSCKASGYTFTSYNMHWVKQTPGRGLEWIGAIY PGNGDTSYNQKFKGKATLTADKSSSTAYMQLSSLTSEDSAVYYCARSTYYGGDWYF NVWGAGTTVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPK SCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAP IEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK	

Scrambled trastuzumab Fc/2	YVCEQQGQFKYSIPMREHVKDRSAPVTSKLPTPAGVWVPKNIETVKDGTPKSTENG LVKPKLTPHHVNQVYGKPLNPTEASNFDTHTQTLEKKEFNTVCSSKMWLRESNKQA FKNVTPYTSDLVAKSSEPGYVRPQPEAQYFNEIYVLVPSGLQSILEKKLPRPFLVCWSV DVVDSTVNDLLPNVEFSHHLKRSCVSVYMESDGESYGDW	23,790
Scrambled trastuzumab Fd	GFNNTYYADGSTSVYLTVLAPGYDVFVTKYGVLGPNGAIAWYHADQATSGDDMHS SSSTHPYQNGKLRMWGNRYTDTCEPIVSGSKALRSLKGTNVSTGPLGDPVASAEEV DKVTSGTPDAKPVGSTTLLGFPSSPTPPFYWGKQQGACSCKGVIVRTGSPSQAGQE VCERVTPPSSLLFKTYLVNLLAAKNSVWKLHGSLVCSFCYWVEISKERRGSTVIVSLQY CATSTSALKSGAP	25,383
Scrambled trastuzumab LC	GKYEVTGPTYQGVVYHRFDQSTLPYCLQDDSGASVKSFYDFATTNQSLTRHTRSFSP VTKCTNIVFGVENSVSSHSLIGSLAAMISQDRCKDTASTRQNVNQASQSKFGFSEPY ASDLRISGAVLPEYVLPTKTITWSECNKPKSKKTSETAPEDEPLKEQPRFQDVVCFGSA AVYDTSGIQKTLAQKPQSGVPSLASTYSAQLWSLSTSVLQY	23,443
Trastuzumab	**LC**:DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLY SGVPSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPS VFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKD STYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC	148,056 (G0F)2
	HC:EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPT NGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDY WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	
	Fd:EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPT NGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDY WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLG	
	Fc/2:GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE PQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD GSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGYVCEQQGQFKY SIPMREHVKDRSAPVTSKLPTPAGVWVPKNIETVKDGTPKSTENGLVKPKLTPHHVN QVYGKPLNPTEASNFDTHTQTLEKKEFNTVCSSKMWLRESNKQAFKNVTPYTSDLV AKSSEP	

3. Sample preparation and acquisition methods

3.1 SEC-nMS experiments

Intact level SEC-nMS analysis of intact mAbs was performed after injecting 10 µg of the analyzed mAbs through a Maxpeak BEH 250 Å 2,5 µm, 4.6 x 150 mm column (Waters) unless it's mentioned otherwise. Mobile phase was 150 mM $AcONH_4$ pH 6.9. The flowrate was 250 µl/min for 6 min. Voltage cone was 120V and desolvation temperature was 300 °C. For Anticalin conjugates a cone voltage of 80V was used with a desolvation temperature of 100 °C to avoid denaturation of the protein.

For the middle level SEC-nMS analysis of trastuzumab, the mAb was digested using IdeS (1 unit/µg) and incubated for 30 min at 37 °C. Similar conditions than for intact trastuzumab were used, except for voltage cone that was reduced to 80V.

3.2 CEX-nMS experiments

For CEX-nMS analysis, 5-10 µg of intact or digested mAb were injected through a BioResolve SCX mAb column (3 µm, 2.1 mm x 100 mm) using a flowrate of 100 µL/min at room temperature. Mobile phase A was 50 mM $AcONH_4$ at pH 5.0 and mobile phase B was 160 mM $AcONH_4$ at pH 8.6. Similar voltage and desolvation temperatures were used than for SEC-nMS.

3.3 Peptide mapping studies of antibodies and anticalines

a) Sample preparation

20 µg of sample were solubilized in 150 mM NH_4HCO_3, 0.1% RapiGest (Waters) at pH 7.8, to obtain a final volume of 24 µL. Disulfide reduction was performed by incubating the solution with 5 mM DTT for 30 min at 57°C. Alkylation of cysteins was performed in 10 mM IAM in the dark at room temperature for 40 min. The enzyme was prepared by suspending 20 µg of trypsin in 100 µL of H_2O. Digestion was performed by adding 1 µL of trypsin (Promega, V5111), which corresponds to a 1:100 enzyme:substrate ratio. Samples were incubated overnight at 37°C. The reaction was stopped by 1% of TFA. RapiGest was eliminated by incubation at 37°C for 30 min and centrifugation at 10,000 g for 5 min.

b) nanoLC-MS/MS

The analysis were performed using a nanoACQUITY Ultra-Performance-LC (Waters) coupled to a Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (ThermoFisher). A volume equivalent to 140 ng of digest were trapped on a Symmetry C18 pre-column (180 µm x 20 mm, 5 µm particle size, Waters) and the peptides were separated on an ACQUITY UPLC® BEH130 C18 separation column (75 µm × 250 mm, 1.7 µm particle size, Waters). The solvent system consisted of 0.1% FA in water (solvent

A) and 0.1% FA in acetonitrile (solvent B). Peptide trapping was performed during 3 min at a flow rate of 5 µL/min with 99% A and 1% B and elution was performed at 60 °C at a flow rate of 350 nL/min from 6% to 40% of B in 43 minutes. MS and MS/MS acquisition were performed in positive mode, with the following settings: spray voltage 1800 V and capillary temperature 250°C. The MS scan had a resolution of 70000, the AGC target was $3x10^6$ and the maximum IT was 50 ms on m/z [300-1800] range. The MS/MS scans were acquired at a resolution of 17500, the AGC target was $1x10^5$ and the maximum IT was 100 ms with fixed first mass of 100 m/z and Isolation window of 2 m/z. Top 10 HCD was selected with intensity threshold of $5x10^4$ and dynamic exclusion of 3 s. The normalized collision energy (NCE) was fixed at 27 V. The complete system was fully controlled by Thermo Scientific™ Xcalibur™ software. Raw data collected were processed and converted with MSConvert into .mgf peak list format.

c) Peptide identification for Trastuzumab samples

Identification of peptides was performed by using the search engine MASCOT 2.6.2 algorithm (Matrix Science) and Byos® 5.0 software (Protein Metrics). The search was performed against the amino acid sequence of trastuzumab. Spectra were searched with a mass tolerance of 10 ppm for MS and 0.05 Da for MS/MS data. The search was made without enzyme specified for MASCOT search, in order to allow the identification of any non-specific peptide cleavage. For Byos® search, trypsin was specified as enzyme with a maximum of three missed cleavages. Variable modifications were specified: carbamidomethylation of cysteine residues, oxidation of methionine residues and adduct of Ugi (769.431 Da for ik-025-02 and 773.389 for ik-025-05) and Passerini (787.442 Da for ik-025-02 and 791.400 for ik-025-05) payload on lysine, aspartate and glutamate residues. Peptide identifications were validated with a minimal ion score of 25 for Mascot and 300 for Byos.

Peptides containing Ugi or Passerini payload were validated with the following criteria: be a tryptic peptide (without unspecific cleavage), the retention time higher than that of the unmodified peptide, identification of signature fragment ions at m/z 286.172 and 637.343, (characteristic of payload fragmentation in the spectrum), and identification with both search engines (Byos® and MASCOT).

d) Peptide identification for Anticalin samples

Identification of peptides was performed by using the search engine MASCOT 2.6.2 algorithm (Matrix Science) and Byos® 5.0 software (Protein Metrics). The searches were performed against the amino acid sequence of proteins. Spectra were searched with a mass tolerance of 10 ppm for MS and 0.05 Da for MS/MS data. The search was made without enzyme specified for MASCOT search, in order to allow the identification of any non-specific peptide cleavage. For Byos® search, trypsin was specified as enzyme with a maximum of three missed cleavages. Variable modifications were specified:

carbamidomethylation of cysteine residues, oxidation of methionine residues and adduct of Ugi (308.185 Da) and Passerini (326.195 Da) payload on lysine, aspartate and glutamate residues. Peptide identifications were validated with a minimal ion score of 25 for Mascot and 300 for Byos.

Peptides containing Ugi or Passerini payload were validated with the following criteria: be a tryptic peptide (without unspecific cleavage), the retention time higher than that of the unmodified peptide, and identification with both search engines (Byos® and MASCOT).

3.4 Peptide mapping studies of nanobodies

a) Sample preparation

Ten µg of the sdADC was solubilized in 150 Mm NH4HCO3 0.1% RapiGest (Waters) at pH 7.8, to obtain a final volume of 24 µL. Disulfide reduction was performed by incubation of the solution with 5 mM DTT for 30 min at 57°C. 10 mM of IAM for 40 min in the dark at room temperature, was added to alkylate free thiol groups on cysteine residues and prevent reformation of disulfide bridges.

Digestion was performed by adding 5 µL of trypsin solution; i.e. 20 µg of trypsin (Promega) suspended in 100 µl of H2O which corresponds to 1:50 enzyme:substrate ratio, at 37°C for 5h. The reaction was stopped by adding 1% of TFA. RapiGest was eliminated by incubation during 30 min at 37°C and centrifugation at 13,000 g for 10 min.

b) LC-MS/MS analysis

nanoLC-MS/MS analysis was performed using a Dionex Ultimate 3000 LC system. 100 ng of sdADC digest was trapped on a Symmetry C18 pre-column (180 µm x 20 mm, 5 µm particle size, Waters) and the peptides were separated on an ACQUITY UPLC® BEH130 C18 separation column (75 µm x 250 mm, 1.7 µm particle size, Waters). The solvent system consisted of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). Peptide trapping was performed during 3 min at a flow rate of 300 nL/min with 97.5% A and 2.5% B and elution was performed at 40 °C at a flow rate of 300 nL/min from 7.5% to 50% of B in 37 minutes. For optimal nanoLC-MS/MS, the mass spectrometer was operated in positive mode, with the following settings: spray voltage 2000 V and capillary temperature 275°C. The MS scan had a resolution of 120 000, the AGC target was 106 and the maximum IT was 50 ms on m/z [300-1800] range. The MS/MS scans had a resolution of 15 000, the AGC target was 105 and the maximum IT was 22 ms with an Isolation window of 2 m/z. Top 10 HCD was selected with intensity threshold of 104 and dynamic exclusion of 5 s. The normalized collision energy (NCE) was fixed at 30%. The complete system was fully controlled by Thermo Scientific™ Xcalibur™ software. Raw data collected were processed and converted with MSConvert into .mgf peak list format.

c) Data processing

search engines MASCOT 2.6.2 algorithm (Matrix Science) was used. The search was performed against the sequence of the light and heavy chains of the ADC. Spectra were searched with a mass tolerance of 10 ppm for MS and 0.05 Da for MS/MS data. The search was made without enzyme specified, in order to allow the identification of any non-specific cleavage peptide. Variable modifications were specified: oxidation of methionine residues, pyro-glutamylation of the N-termini, deamidation of asparagine, isomerization of aspartic acid residues and drug-linker conjugation (C52 F H56 N9 O13 and C26 O8 N6 H34) on cysteine residues. Peptide identifications were validated with a minimal Mascot ion score of 30.

3.5 Mass photometry

The bsAb samples were diluted to 50 µM in PBS and then 2 µL were mixed with 18 µL of PBS in the wells on the gasket.

3.6 RPLC-MS of rebriged Fabs and intact trastuzumab

Less than 1 µg was injected on a Bioresolve RP mAb polyphenyl (450Å, 2,7 µm 2.1x50 mm column, Waters) at a flowrate of 300 µL/min at 80 °C. Mobile phases consisted of 0.1% FA in H2O (mobile phase A) and 0.1% FA in ACN (mobile phase B). The separation was carried out using a gradient from 5 to 95% of mobile phase B in 25 min. The column was washed with 95% mobile phase B for 1 minute and then was equilibrate with 5% mobile phase B for 3 minutes. The BioAccord was operated in the positive mode with a capillary voltage of 1.5 kV. Desolvation temperature was set to 330 °C, the cone voltage to 60V and the source pressure was fixed at 2 mbar. Acquisitions were performed on the m/z range 400-7000 with a 1 s scan time.

3.7 dSEC-MS of rebriged Fabs and intact trastuzumab

SEC analysis was performed using the AdvanceBioSEC3 (300 Å 2,5 µm, 4,6 x 150 mm, Agilent) or the Maxpeak Protein SEC (4.6 mm x 150 mm, 2.5 µm, 250 Å, Waters) column kept at room temperature. The separation was carried out using an isocratic gradient of mobile phase (20% ACN+0.1% FA+0.1%TFA) at a flowrate of 0.1 mL/min for 15 minutes. For dSEC-MS, 1 to 3 µg of protein samples were injected. The BioAccord was operated in the same conditions than for rpLC-MS. Full scan acquisition was performed on the high mass range (400 – 7000 m/z) with a 1 s can time.

3.8 TD-MS experiments of reference proteins

Myoglobin from equine heart (Sigma-Aldrich) and carbonic anhydrase from bovine erythrocytes (Sigma-Aldrich) were diluted to 2 µM in 49.5/49.5/1 H2O/ACN/FA. Both proteins were infused with an

H-ESI source (Thermo Fisher Scientific) at a syringe flowrate of 5 µL/min. The source temperature was 275°C.

The Orbitrap Eclipse Tribrid MS (Thermo Fisher Scientific) was in Intact Protein mode. Spray voltage, sheath gas and auxiliary gas were set to 3,800 V, 4 arbitrary units (a. u.) and 2 a. u., respectively. The MS/MS fragmentation was performed at an Orbitrap resolution of 120,000 (at 200 m/z) using HCD, ETD and UVPD using different optimization parameters discussed in the **Part III chapter 1**. The AGC target was set to 7x10e5 and the maximum injection time to 200 ms. The MS/MS spectra were collected after averaging 10 µscans for 1 min.

For data analysis, spectra averaged over 1 min were deconvoluted using Xtract™ algorithm included in the FreeStyle software (Thermo Fisher Scientific) using the following parameters: S/N of 7, fit factor of 80%, remainder threshold of 25 and maximum charge state to the charge state of the precursor ion and minimum number of charges to 1.

For the determination of terminal fragments, deconvoluted [MH+] fragments were matched to the corresponding sequences with ProSight Lite using a mass tolerance of 10 ppm. For each fragmentation, the type of fragment ions was prior selected.

3.9 MD-MS experiments of trastuzumab subunits

For subunits generation, intact trastuzumab was digested using IdeS (1 unit/µg) and incubated for 30 min at 37 °C and then was reduced using 100 mM DTT and incubated for 1h at 57°C. Around 1 µg of sample were loaded on the column.

RPLC separation of the subunits was carried out on the MabPac RP 4 µm 1.0 x 100 mm column (Thermo Fisher Scientific) for 12 min. Mobile phase A was 0.1% FA in water and mobile phase B was 0.1% FA in ACN. Antibody The flow rate was set to 100 µL/min following a linear gradient of 12 min: from 20% to 50% B in 5.5 min, 90% B for 2 min and equilibration at 20% B for 4 min.

MS and MS/MS analyses were performed under similar conditions than TD-MS workflow, except for source parameters that were set as follows: ion transfer tube to 320°C, vaporizer to 300°C, spray voltage to 3,500 V, sheath gas to 25 a. u., auxiliary gas to 10 a. u. and in-source fragmentation of 10 V.

Terminal fragments were assigned similarly for TD-MS experiments for reference proteins. For internal fragments analysis, ClipsMS was used with a mass tolerance of 2 ppm and a smallest internal fragment size of 5 amino acids. For each fragmentation techniques, terminal and internal corresponding fragment types were searched.

3.10 TD-MS experiments of nanobodies

A Dionex Ultimate 3000 LC system was used to inject one to 3 µg of samples through an Agilent Zorbax 300 SB-C8 (2,1 x 50 mm, 1,8 microns) at 60°C. The solvent system consisted of 0.1% TFA in water (solvent A) and 0.1% TFA in acetonitrile (solvent B). Elution was performed at a flow rate of 200 µL/min from 20% to 60% of B for 7 minutes. All experiments were performed on an Orbitrap Tribrid Eclipse MS (Thermo Scientific). The LC was hyphenated to an Orbitrap Tribrid Eclipse MS (Thermo Scientific) equipped with ETD, HCD and 213 nm UVPD options. For all experiments, the spray voltage was set to 3.4 kV, and the ion transfer tube temperature at 320 °C.

The MS scans were acquired at a resolution of 15 000, the AGC target was 106 and the maximum IT at 200 ms on a range of [200-3000] m/z. MS/MS scans had a resolution of 120 000, the AGC target was 106 and the maximum IT was 200 ms on m/z [180-2000] with an IW of 2 m/z. The MS scan had a resolution of 120 000, the normalized AGC target was 250% and the maximum IT was 50 ms on m/z [300-1800] range. The MS/MS scans had a resolution of 15 000, the nAGC target was 200% and the maximum IT was 22 ms with an Isolation window of 2 m/z. For HCD fragmentation, the ions were accelerated under a constant N2 pressure of 10-9 mbar with 12 eV. For performing ETD, anionic fluoranthene radicals were generated in the source region of the instrument. For UVPD, ions were activated with 213 nm laser delivering a total energy of 150 µJ (2 µJ /pulse).

References

References

1. Kaplon, H.; Crescioli, S.; Chenoweth, A.; Visweswaraiah, J.; Reichert, J. M., Antibodies to watch in 2023. *MAbs* **2023,** *15* (1), 2153410.
2. Singh, S.; Kumar, N. K.; Dwiwedi, P.; Charan, J.; Kaur, R.; Sidhu, P.; Chugh, V. K., Monoclonal Antibodies: A Review. *Curr Clin Pharmacol* **2018,** *13* (2), 85-99.
3. Yang, E. Y.; Shah, K., Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Front Oncol* **2020,** *10*, 1182.
4. Sun, Y.; Yu, X.; Wang, X.; Yuan, K.; Wang, G.; Hu, L.; Zhang, G.; Pei, W.; Wang, L.; Sun, C.; Yang, P., Bispecific antibodies in cancer therapy: Target selection and regulatory requirements. *Acta Pharmaceutica Sinica B* **2023**.
5. Carter, P. J.; Lazar, G. A., Next generation antibody drugs: pursuit of the 'high-hanging fruit'. *Nat Rev Drug Discov* **2018,** *17* (3), 197-223.
6. Fu, Z.; Li, S.; Han, S.; Shi, C.; Zhang, Y., Antibody drug conjugate: the "biological missile" for targeted cancer therapy. *Signal Transduct Target Ther* **2022,** *7* (1), 93.
7. Kostova, V.; Desos, P.; Starck, J. B.; Kotschy, A., The Chemistry Behind ADCs. *Pharmaceuticals (Basel)* **2021,** *14* (5), 442-449.
8. Xu, Y.; Wang, D.; Mason, B.; Rossomando, T.; Li, N.; Liu, D.; Cheung, J. K.; Xu, W.; Raghava, S.; Katiyar, A.; Nowak, C.; Xiang, T.; Dong, D. D.; Sun, J.; Beck, A.; Liu, H., Structure, heterogeneity and developability assessment of therapeutic antibodies. *MAbs* **2019,** *11* (2), 239-264.
9. Wakankar, A.; Chen, Y.; Gokarn, Y.; Jacobson, F. S., Analytical methods for physicochemical characterization of antibody drug conjugates. *MAbs* **2011,** *3* (2), 161-72.
10. Beck, A.; D'Atri, V.; Ehkirch, A.; Fekete, S.; Hernandez-Alba, O.; Gahoual, R.; Leize-Wagner, E.; Francois, Y.; Guillarme, D.; Cianferani, S., Cutting-edge multi-level analytical and structural characterization of antibody-drug conjugates: present and future. *Expert Rev Proteomics* **2019,** *16* (4), 337-362.
11. Beck, A.; Sanglier-Cianferani, S.; Van Dorsselaer, A., Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. *Anal Chem* **2012,** *84* (11), 4637-46.
12. Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S., Characterization of therapeutic antibodies and related products. *Anal Chem* **2013,** *85* (2), 715-36.
13. Zhu, X.; Huo, S.; Xue, C.; An, B.; Qu, J., Current LC-MS-based strategies for characterization and quantification of antibody-drug conjugates. *J Pharm Anal* **2020,** *10* (3), 209-220.
14. Leney, A. C.; Heck, A. J., Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom* **2017,** *28* (1), 5-13.
15. Goyon, A.; D'Atri, V.; Colas, O.; Fekete, S.; Beck, A.; Guillarme, D., Characterization of 30 therapeutic antibodies and related products by size exclusion chromatography: Feasibility assessment for future mass spectrometry hyphenation. *J Chromatogr B Analyt Technol Biomed Life Sci* **2017,** *1065-1066*, 35-43.
16. Murisier, A.; Andrie, M.; Fekete, S.; Lauber, M.; D'Atri, V.; Iwan, K.; Guillarme, D., Direct coupling of size exclusion chromatography and mass spectrometry for the characterization of complex monoclonal antibody products. *J Sep Sci* **2022,** *45* (12), 1997-2007.
17. Ehkirch, A.; Hernandez-Alba, O.; Colas, O.; Beck, A.; Guillarme, D.; Cianferani, S., Hyphenation of size exclusion chromatography to native ion mobility mass spectrometry for the analytical characterization of therapeutic antibodies and related products. *J Chromatogr B Analyt Technol Biomed Life Sci* **2018,** *1086*, 176-183.
18. Fekete, S.; Beck, A.; Veuthey, J. L.; Guillarme, D., Ion-exchange chromatography for the characterization of biopharmaceuticals. *J Pharm Biomed Anal* **2015,** *113*, 43-55.
19. Murisier, A.; Duivelshof, B. L.; Fekete, S.; Bourquin, J.; Schmudlach, A.; Lauber, M. A.; Nguyen, J. M.; Beck, A.; Guillarme, D.; D'Atri, V., Towards a simple on-line coupling of ion exchange

chromatography and native mass spectrometry for the detailed characterization of monoclonal antibodies. *J Chromatogr A* **2021,** *1655*, 462499.
20. Neil L. Kelleher; Hong Y. Lin; Gary A. Valaskovic; David J. Aaserud; Einar K. Fridriksson; McLafferty, F. W., Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J Am Chem Soc* **1999,** *121*, 806-812.
21. Catherman, A. D.; Skinner, O. S.; Kelleher, N. L., Top Down proteomics: facts and perspectives. *Biochem Biophys Res Commun* **2014,** *445* (4), 683-93.
22. Fornelli, L.; Toby, T. K.; Schachner, L. F.; Doubleday, P. F.; Srzentic, K.; DeHart, C. J.; Kelleher, N. L., Top-down proteomics: Where we are, where we are going? *J Proteomics* **2018,** *175*, 3-4.
23. Fornelli, L.; Parra, J.; Hartmer, R.; Stoermer, C.; Lubeck, M.; Tsybin, Y. O., Top-down analysis of 30–80 kDa proteins by electron transfer dissociation time-of-flight mass spectrometry. *Anal Bioanal Chem* **2013,** *405*, 8505–8514.
24. Riley, N. M.; Westphall, M. S.; Coon, J. J., Sequencing Larger Intact Proteins (30-70 kDa) with Activated Ion Electron Transfer Dissociation. *J Am Soc Mass Spectrom* **2018,** *29* (1), 140-149.
25. Srzentic, K.; Fornelli, L.; Tsybin, Y. O.; Loo, J. A.; Seckler, H.; Agar, J. N.; Anderson, L. C.; Bai, D. L.; Beck, A.; Brodbelt, J. S.; van der Burgt, Y. E. M.; Chamot-Rooke, J.; Chatterjee, S.; Chen, Y.; Clarke, D. J.; Danis, P. O.; Diedrich, J. K.; D'Ippolito, R. A.; Dupre, M.; Gasilova, N.; Ge, Y.; Goo, Y. A.; Goodlett, D. R.; Greer, S.; Haselmann, K. F.; He, L.; Hendrickson, C. L.; Hinkle, J. D.; Holt, M. V.; Hughes, S.; Hunt, D. F.; Kelleher, N. L.; Kozhinov, A. N.; Lin, Z.; Malosse, C.; Marshall, A. G.; Menin, L.; Millikin, R. J.; Nagornov, K. O.; Nicolardi, S.; Pasa-Tolic, L.; Pengelley, S.; Quebbemann, N. R.; Resemann, A.; Sandoval, W.; Sarin, R.; Schmitt, N. D.; Shabanowitz, J.; Shaw, J. B.; Shortreed, M. R.; Smith, L. M.; Sobott, F.; Suckau, D.; Toby, T.; Weisbrod, C. R.; Wildburger, N. C.; Yates, J. R., 3rd; Yoon, S. H.; Young, N. L.; Zhou, M., Interlaboratory Study for Characterizing Monoclonal Antibodies by Top-Down and Middle-Down Mass Spectrometry. *J Am Soc Mass Spectrom* **2020,** *31* (9), 1783-1802.
26. Lodge, J. M.; Schauer, K. L.; Brademan, D. R.; Riley, N. M.; Shishkova, E.; Westphall, M. S.; Coon, J. J., Top-Down Characterization of an Intact Monoclonal Antibody Using Activated Ion Electron Transfer Dissociation. *Anal Chem* **2020,** *92* (15), 10246-10251.
27. Shaw, J. B.; Cooper-Shepherd, D. A.; Hewitt, D.; Wildgoose, J. L.; Beckman, J. S.; Langridge, J. I.; Voinov, V. G., Enhanced Top-Down Protein Characterization with Electron Capture Dissociation and Cyclic Ion Mobility Spectrometry. *Anal Chem* **2022,** *94* (9), 3888-3896.
28. Hernandez-Alba, O.; Houel, S.; Hessmann, S.; Erb, S.; Rabuka, D.; Huguet, R.; Josephs, J.; Beck, A.; Drake, P. M.; Cianferani, S., A Case Study to Identify the Drug Conjugation Site of a Site-Specific Antibody-Drug-Conjugate Using Middle-Down Mass Spectrometry. *J Am Soc Mass Spectrom* **2019,** *30* (11), 2419-2429.
29. Watts, E.; Williams, J. D.; Miesbauer, L. J.; Bruncko, M.; Brodbelt, J. S., Comprehensive Middle-Down Mass Spectrometry Characterization of an Antibody-Drug Conjugate by Combined Ion Activation Methods. *Anal Chem* **2020,** *92* (14), 9790-9798.
30. Dhenin, J.; Dupre, M.; Druart, K.; Krick, A.; Mauriac, C.; Chamot-Rooke, J., A multiparameter optimization in middle-down analysis of monoclonal antibodies by LC-MS/MS. *J Mass Spectrom* **2023,** *58* (3), e4909.
31. Schroeder, H. W., Jr.; Cavacini, L., Structure and function of immunoglobulins. *J Allergy Clin Immunol* **2010,** *125* (2 Suppl 2), S41-52.
32. Sornay, C.; Hessmann, S.; Erb, S.; Dovgan, I.; Ehkirch, A.; Botzanowski, T.; Cianferani, S.; Wagner, A.; Chaubet, G., Investigating Ugi/Passerini Multicomponent Reactions for the Site-Selective Conjugation of Native Trastuzumab*. *Chemistry* **2020,** *26* (61), 13797-13805.
33. Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M., Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* **2007,** *4* (9), 709-12.
34. Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* **2004,** *101* (26), 9528-33.
35. Brodbelt, J. S.; Morrison, L. J.; Santos, I., Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chem Rev* **2020,** *120* (7), 3328-3380.

36. Muhammed, Y., The Best IgG Subclass for the Development of Therapeutic Monoclonal Antibody Drugs and their Commercial Production: A Review. *Immunome Research* **2020,** *16* (1).
37. Sun, Y.; Huang, T.; Hammarstrom, L.; Zhao, Y., The Immunoglobulins: New Insights, Implications, and Applications. *Annu Rev Anim Biosci* **2020,** *8*, 145-169.
38. Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A., The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* **2007,** *25*, 21-50.
39. Sela-Culang, I.; Kunik, V.; Ofran, Y., The structural basis of antibody-antigen recognition. *Front Immunol* **2013,** *4*, 302.
40. Vidarsson, G.; Dekkers, G.; Rispens, T., IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* **2014,** *5*, 520.
41. Loureiro, L. R.; Carrascal, M. A.; Barbas, A.; Ramalho, J. S.; Novo, C.; Delannoy, P.; Videira, P. A., Challenges in Antibody Development against Tn and Sialyl-Tn Antigens. *Biomolecules* **2015,** *5* (3), 1783-809.
42. Tang, Y.; Cain, P.; Anguiano, V.; Shih, J. J.; Chai, Q.; Feng, Y., Impact of IgG subclass on molecular properties of monoclonal antibodies. *MAbs* **2021,** *13* (1), 1993768.
43. Aalberse, R. C.; Stapel, S. O.; Schuurman, J.; Rispens, T., Immunoglobulin G4: an odd antibody. *Clin Exp Allergy* **2009,** *39* (4), 469-77.
44. Shade, K.-T.; Anthony, R., Antibody Glycosylation and Inflammation. *Antibodies* **2013,** *2* (4), 392-414.
45. Wang, L. X.; Tong, X.; Li, C.; Giddens, J. P.; Li, T., Glycoengineering of Antibodies for Modulating Functions. *Annu Rev Biochem* **2019,** *88*, 433-459.
46. Ferrara, C.; Grau, S.; Jager, C.; Sondermann, P.; Brunker, P.; Waldhauer, I.; Hennig, M.; Ruf, A.; Rufer, A. C.; Stihle, M.; Umana, P.; Benz, J., Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. *Proc Natl Acad Sci U S A* **2011,** *108* (31), 12669-74.
47. Shi, W.; Li, W.; Zhang, J.; Li, T.; Song, Y.; Zeng, Y.; Dong, Q.; Lin, Z.; Gong, L.; Fan, S.; Tang, F.; Huang, W., One-step synthesis of site-specific antibody-drug conjugates by reprograming IgG glycoengineering with LacNAc-based substrates. *Acta Pharm Sin B* **2022,** *12* (5), 2417-2428.
48. Ma, B.; Guan, X.; Li, Y.; Shang, S.; Li, J.; Tan, Z., Protein Glycoengineering: An Approach for Improving Protein Properties. *Front Chem* **2020,** *8*, 622.
49. Wang, L. X.; Lomino, J. V., Emerging technologies for making glycan-defined glycoproteins. *ACS Chem Biol* **2012,** *7* (1), 110-22.
50. Umaña, P.; Jean-Mairet, J.; Moudry, R.; Amstutz, H.; Bailey, J. E., Engineered glycoforms of an antineuro-blastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *NATURE BIOTECHNOLOGY* **1999,** *17*.
51. Xu, J.; Shao, Z.; Wang, Z.; Huang, Y.; Zou, X.; Shen, Y., Developing a medium combination to attain similar glycosylation profile to originator by DoE and cluster analysis method. *Sci Rep* **2021,** *11* (1), 7103.
52. Liu, H.; May, K., Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function. *MAbs* **2012,** *4* (1), 17-23.
53. Lefranc, M. P.; Pommie, C.; Kaas, Q.; Duprat, E.; Bosc, N.; Guiraudou, D.; Jean, C.; Ruiz, M.; Da Piedade, I.; Rouard, M.; Foulquier, E.; Thouvenin, V.; Lefranc, G., IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. *Dev Comp Immunol* **2005,** *29* (3), 185-203.
54. Gevondyan, N. M.; Volynskaia, A. M.; Gevondyan, V. S., Four free cysteine residues found in human IgG1 of healthy donors. *Biochemistry (Mosc)* **2006,** *71* (3), 279-84.
55. Cohen, S. L.; Price, C.; Vlasak, J., Elimination and Peptide Bond Hydrolysis: Two Distinct Mechanisms of Human IgG1 Hinge Fragmentation upon Storage. *JACS Communications* **2007,** *129*, 6976-6977.
56. Forthal, D. N., Function of antibodies. *Microbiol Spectr.* **2014,** *2* (4), 1-17.

57. Gómez Román, V. R.; Murray, J. C.; Weiner, L. M., Antibody-Dependent Cellular Cytotoxicity (ADCC). **2014**, 1-27.
58. Cao, X.; Chen, J.; Li, B.; Dang, J.; Zhang, W.; Zhong, X.; Wang, C.; Raoof, M.; Sun, Z.; Yu, J.; Fakih, M. G.; Feng, M., Promoting antibody-dependent cellular phagocytosis for effective macrophage-based cancer immunotherapy. *SCIENCE ADVANCES* **2022,** *8*.
59. Meyer, S.; Leusen, J. H.; Boross, P., Regulation of complement and modulation of its activity in monoclonal antibody therapy of cancer. *MAbs* **2014,** *6* (5), 1133-44.
60. Rodriguez-Nava, C.; Ortuno-Pineda, C.; Illades-Aguiar, B.; Flores-Alfaro, E.; Leyva-Vazquez, M. A.; Parra-Rojas, I.; Del Moral-Hernandez, O.; Vences-Velazquez, A.; Cortes-Sarabia, K.; Alarcon-Romero, L. D. C., Mechanisms of Action and Limitations of Monoclonal Antibodies and Single Chain Fragment Variable (scFv) in the Treatment of Cancer. *Biomedicines* **2023,** *11* (6).
61. Lipman, N. S.; Jackson, L. R.; Trudel, L. J.; Weis-Garcia, F., Monoclonal Versus Polyclonal Antibodies: Distinguishing Characteristics, Applications, and Information Resources. *ILAR Journal* **2023**.
62. Hwang, W. Y.; Foote, J., Immunogenicity of engineered antibodies. *Methods* **2005,** *36* (1), 3-10.
63. Waldmann, H., Human Monoclonal Antibodies: The Benefits of Humanization. *Methods Mol Biol* **2019,** *1904*, 1-10.
64. Torres, M.; Fernandez-Fuentes, N.; Fiser, A.; Casadevall, A., Exchanging murine and human immunoglobulin constant chains affects the kinetics and thermodynamics of antigen binding and chimeric antibody autoreactivity. *PLoS One* **2007,** *2* (12), e1310.
65. Harding, i. A.; Stickler, M. M.; Razo, J.; DuBridge, R. B., The immunogenicity of humanized and fully human antibodies. *mAbs* **2010,** *2* (3).
66. Tan, H. T.; Sugita, K.; Akdis, C. A., Novel Biologicals for the Treatment of Allergic Diseases and Asthma. *Curr Allergy Asthma Rep* **2016,** *16* (10), 70.
67. Guimaraes Koch, S. S.; Thorpe, R.; Kawasaki, N.; Lefranc, M. P.; Malan, S.; Martin, A. C. R.; Mignot, G.; Pluckthun, A.; Rizzi, M.; Shubat, S.; Weisser, K.; Balocco, R., International nonproprietary names for monoclonal antibodies: an evolving nomenclature system. *MAbs* **2022,** *14* (1), 2075078.
68. Sifniotis, V.; Cruz, E.; Eroglu, B.; Kayser, V., Current Advancements in Addressing Key Challenges of Therapeutic Antibody Design, Manufacture, and Formulation. *Antibodies (Basel)* **2019,** *8* (2).
69. Beck, A., Biosimilar, biobetter and next generation therapeutic antibodies. *MAbs* **2011,** *3* (2), 107-10.
70. Baah, S.; Laws, M.; Rahman, K. M., Antibody-Drug Conjugates-A Tutorial Review. *Molecules* **2021,** *26* (10).
71. Beck, A.; Goetsch, L.; Dumontet, C.; Corvaia, N., Strategies and challenges for the next generation of antibody-drug conjugates. *Nat Rev Drug Discov* **2017,** *16* (5), 315-337.
72. Kaur, R.; Kaur, G.; Gill, R. K.; Soni, R.; Bariwal, J., Recent developments in tubulin polymerization inhibitors: An overview. *Eur J Med Chem* **2014,** *87*, 89-124.
73. Walczak, C. E., Microtubule dynamics and tubulin interacting proteins. *Current Opinion in Cell Biology* **2000,** *12*.
74. Cheung-Ong, K.; Giaever, G.; Nislow, C., DNA-damaging agents in cancer chemotherapy: serendipity and chemical biology. *Chem Biol* **2013,** *20* (5), 648-59.
75. Katz, J.; Janik, J. E.; Younes, A., Brentuximab Vedotin (SGN-35). *Clin Cancer Res* **2011,** *17* (20), 6428-36.
76. Girish, S.; Gupta, M.; Wang, B.; Lu, D.; Krop, I. E.; Vogel, C. L.; Burris Iii, H. A.; LoRusso, P. M.; Yi, J. H.; Saad, O.; Tong, B.; Chu, Y. W.; Holden, S.; Joshi, A., Clinical pharmacology of trastuzumab emtansine (T-DM1): an antibody-drug conjugate in development for the treatment of HER2-positive cancer. *Cancer Chemother Pharmacol* **2012,** *69* (5), 1229-40.
77. Hoffmann, R. M.; Coumbe, B. G. T.; Josephs, D. H.; Mele, S.; Ilieva, K. M.; Cheung, A.; Tutt, A. N.; Spicer, J. F.; Thurston, D. E.; Crescioli, S.; Karagiannis, S. N., Antibody structure and engineering considerations for the design and function of Antibody Drug Conjugates (ADCs). *Oncoimmunology* **2018,** *7* (3), e1395127.

78. Lyon, R. P.; Bovee, T. D.; Doronina, S. O.; Burke, P. J.; Hunter, J. H.; Neff-LaFord, H. D.; Jonas, M.; Anderson, M. E.; Setter, J. R.; Senter, P. D., Reducing hydrophobicity of homogeneous antibody-drug conjugates improves pharmacokinetics and therapeutic index. *Nat Biotechnol* **2015,** *33* (7), 733-5.
79. Akkapeddi, P.; Azizi, S. A.; Freedy, A. M.; Cal, P.; Gois, P. M. P.; Bernardes, G. J. L., Construction of homogeneous antibody-drug conjugates using site-selective protein chemistry. *Chem Sci* **2016,** *7* (5), 2954-2963.
80. King, T. A.; Walsh, S. J.; Kapun, M.; Wharton, T.; Krajcovicova, S.; Glossop, M. S.; Spring, D. R., Disulfide re-bridging reagents for single-payload antibody-drug conjugates. *Chem Commun (Camb)* **2023,** *59* (65), 9868-9871.
81. Dannheim, F. M.; Walsh, S. J.; Orozco, C. T.; Hansen, A. H.; Bargh, J. D.; Jackson, S. E.; Bond, N. J.; Parker, J. S.; Carroll, J. S.; Spring, D. R., All-in-one disulfide bridging enables the generation of antibody conjugates with modular cargo loading. *Chemical Science* **2022,** *13* (30), 8781-8790.
82. Marquard, A. N.; Carlson, J. C. T.; Weissleder, R., Expanding the Scope of Antibody Rebridging with New Pyridazinedione-TCO Constructs. *Bioconjug Chem* **2020,** *31* (6), 1616-1623.
83. Xu, L.; Raabe, M.; Zegota, M. M.; Nogueira, J. C. F.; Chudasama, V.; Kuan, S. L.; Weil, T., Site-selective protein modification via disulfide rebridging for fast tetrazine/trans-cyclooctene bioconjugation. *Org Biomol Chem* **2020,** *18* (6), 1140-1147.
84. Thoreau, F.; Rochet, L. N. C.; Baker, J. R.; Chudasama, V., Enabling the formation of native mAb, Fab' and Fc-conjugates using a bis-disulfide bridging reagent to achieve tunable payload-to-antibody ratios (PARs). *Chem Sci* **2023,** *14* (14), 3752-3762.
85. Bahou, C.; Chudasama, V., The use of bromopyridazinedione derivatives in chemical biology. *Org Biomol Chem* **2022,** *20* (30), 5879-5890.
86. Vhora, I.; Patil, S.; Bhatt, P.; Misra, A., Protein- and Peptide-drug conjugates: an emerging drug delivery technology. *Adv Protein Chem Struct Biol* **2015,** *98*, 1-55.
87. Labrijn, A. F.; Janmaat, M. L.; Reichert, J. M.; Parren, P., Bispecific antibodies: a mechanistic review of the pipeline. *Nat Rev Drug Discov* **2019,** *18* (8), 585-608.
88. Dahlén, E.; Veitonmäki, N.; Norlén, P., Bispecific antibodies in cancer immunotherapy. *Therapeutic Advances in Vaccines and Immunotherapy* **2018,** *6* (1), 3 –17.
89. Bates, A.; Power, C. A., David vs. Goliath: The Structure, Function, and Clinical Prospects of Antibody Fragments. *Antibodies (Basel)* **2019,** *8* (2).
90. Wu, L.; Seung, E.; Xu, L.; Rao, E.; Lord, D. M.; Wei, R. R.; Cortez-Retamozo, V.; Ospina, B.; Posternak, V.; Ulinski, G.; Piepenhagen, P.; Francesconi, E.; El-Murr, N.; Beil, C.; Kirby, P.; Li, A.; Fretland, J.; Vicente, R.; Deng, G.; Dabdoubi, T.; Cameron, B.; Bertrand, T.; Ferrari, P.; Pouzieux, S.; Lemoine, C.; Prades, C.; Park, A.; Qiu, H.; Song, Z.; Zhang, B.; Sun, F.; Chiron, M.; Rao, S.; Radosevic, K.; Yang, Z. Y.; Nabel, G. J., Trispecific antibodies enhance the therapeutic efficacy of tumor-directed T cells through T cell receptor co-stimulation. *Nat Cancer* **2020,** *1* (1), 86-98.
91. Kang, J.; Sun, T.; Zhang, Y., Immunotherapeutic progress and application of bispecific antibody in cancer. *Front Immunol* **2022,** *13*, 1020003.
92. Moghimi, S. M.; Rahbarizadeh, F.; Ahmadvand, D.; Parhamifar, L., Heavy Chain Only Antibodies: A New Paradigm in Personalized HER2+ Breast Cancer Therapy. *Bioimpacts* **2013,** *3* (1), 1-4.
93. Tang, H.; Gao, Y.; Han, J., Application Progress of the Single Domain Antibody in Medicine. *Int J Mol Sci* **2023,** *24* (4).
94. Pillay, T. S.; Muyldermans, S., Application of Single-Domain Antibodies ("Nanobodies") to Laboratory Diagnosis. *Annals of Laboratory Medicine* **2021,** *41* (6), 549-558.
95. Muyldermans, S., Applications of Nanobodies. *Annu Rev Anim Biosci* **2021,** *9*, 401-421.
96. Kang, W.; Ding, C.; Zheng, D.; Ma, X.; Yi, L.; Tong, X.; Wu, C.; Xue, C.; Yu, Y.; Zhou, Q., Nanobody Conjugates for Targeted Cancer Therapy and Imaging. *Technol Cancer Res Treat* **2021,** *20*, 15330338211010117.
97. Verhaar, E. R.; Woodham, A. W.; Ploegh, H. L., Nanobodies in cancer. *Semin Immunol* **2021,** *52*, 101425.

98. Jin, B. K.; Odongo, S.; Radwanska, M.; Magez, S., Nanobodies: A Review of Generation, Diagnostics and Therapeutics. *Int J Mol Sci* **2023,** *24* (6).
99. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013,** *82*, 775-97.
100. Hassanzadeh-Ghassabeh, G.; Devoogdt, N.; De Pauw, P.; Vincke, C.; Muyldermans, S., Nanobodies and their potential applications. *Perspective* **2013,** *8* (6), 1013–1026.
101. Goli, V. A. R.; Butreddy, A., Biosimilar monoclonal antibodies: Challenges and approaches towards formulation. *Chem Biol Interact* **2022,** *366*, 110116.
102. Castel, J.; Delaux, S.; Hernandez-Alba, O.; Cianferani, S., Recent advances in structural mass spectrometry methods in the context of biosimilarity assessment: from sequence heterogeneities to higher order structures. *J Pharm Biomed Anal* **2023,** *236*, 115696.
103. Duivelshof, B. L.; Jiskoot, W.; Beck, A.; Veuthey, J. L.; Guillarme, D.; D'Atri, V., Glycosylation of biosimilars: Recent advances in analytical characterization and clinical implications. *Anal Chim Acta* **2019,** *1089*, 1-18.
104. Ho, E. C. H.; Qiu, R.; Miller, E.; Bilotta, M. T.; FitzGerald, D.; Antignani, A., Antibody drug conjugates, targeting cancer-expressed EGFR, exhibit potent and specific antitumor activity. *Biomed Pharmacother* **2023,** *157*, 114047.
105. Porebska, N.; Ciura, K.; Chorazewska, A.; Zakrzewska, M.; Otlewski, J.; Opalinski, L., Multivalent protein-drug conjugates - An emerging strategy for the upgraded precision and efficiency of drug delivery to cancer cells. *Biotechnol Adv* **2023,** *67*, 108213.
106. Kaplon, H.; Chenoweth, A.; Crescioli, S.; Reichert, J. M., Antibodies to watch in 2022. *MAbs* **2022,** *14* (1), 2014296.
107. Dahlén, E.; Veitonmäki, N.; Norlén, P., Bispecific antibodies in cancer immunotherapy. *Therapeutic Advances in Vaccines and Immunotherapy* **2017,** *6* (1), 3 –17.
108. Kim, E. G.; Kim, K. M., Strategies and Advancement in Antibody-Drug Conjugate Optimization for Targeted Cancer Therapeutics. *Biomol Ther (Seoul)* **2015,** *23* (6), 493-509.
109. Norsworthy, K. J.; Ko, C. W.; Lee, J. E.; Liu, J.; John, C. S.; Przepiorka, D.; Farrell, A. T.; Pazdur, R., FDA Approval Summary: Mylotarg for Treatment of Patients with Relapsed or Refractory CD33-Positive Acute Myeloid Leukemia. *Oncologist* **2018,** *23* (9), 1103-1108.
110. Tong, J. T. W.; Harris, P. W. R.; Brimble, M. A.; Kavianinia, I., An Insight into FDA Approved Antibody-Drug Conjugates for Cancer Therapy. *Molecules* **2021,** *26* (19).
111. Federal Register of the ICH. **1999,** *64* (159), 44817–45148.
112. Kretsinger, J.; Frantz, N.; Hart, S. A.; Kelley, W. P.; Kitchen, B.; Novick, S.; Rellahan, B.; Stranges, D.; Stroop, C. J. M.; Yin, P.; Gastens, M. H., Expectations for Phase-Appropriate Drug Substance and Drug Product Specifications for Early-Stage Protein Therapeutics. *J Pharm Sci* **2019,** *108* (4), 1442-1452.
113. Alt, N.; Zhang, T. Y.; Motchnik, P.; Taticek, R.; Quarmby, V.; Schlothauer, T.; Beck, H.; Emrich, T.; Harris, R. J., Determination of critical quality attributes for monoclonal antibodies using quality by design principles. *Biologicals* **2016,** *44* (5), 291-305.
114. Jefferis, R., Posttranslational Modifications and the Immunogenicity of Biotherapeutics. *J Immunol Res* **2016,** *2016*, 5358272.
115. Ambrogelly, A.; Gozo, S.; Katiyar, A.; Dellatore, S.; Kune, Y.; Bhat, R.; Sun, J.; Li, N.; Wang, D.; Nowak, C.; Neill, A.; Ponniah, G.; King, C.; Mason, B.; Beck, A.; Liu, H., Analytical comparability study of recombinant monoclonal antibody therapeutics. *MAbs* **2018,** *10* (4), 513-538.
116. Butre, C. I.; D'Atri, V.; Diemer, H.; Colas, O.; Wagner, E.; Beck, A.; Cianferani, S.; Guillarme, D.; Delobel, A., Interlaboratory Evaluation of a User-Friendly Benchtop Mass Spectrometer for Multiple-Attribute Monitoring Studies of a Monoclonal Antibody. *Molecules* **2023,** *28* (6).
117. Camperi, J.; Guillarme, D.; Stella, C., Targeted Bottom-up Characterization of Recombinant Monoclonal Antibodies by Multidimensional LC/MS. *Anal Chem* **2020,** *92* (19), 13420-13426.
118. Wagner-Rousset, E.; Janin-Bussat, M. C.; Colas, O.; Excoffier, M.; Ayoub, D.; Haeuw, J. F.; Rilatt, I.; Perez, M.; Corvaia, N.; Beck, A., Antibody-drug conjugate model fast characterization by LC-MS following IdeS proteolytic digestion. *MAbs* **2014,** *6* (1), 273-85.

119. Zhu, W.; Li, M.; Zhang, J., Integrating Intact Mass Analysis and Middle-Down Mass Spectrometry Approaches to Effectively Characterize Trastuzumab and Adalimumab Structural Heterogeneity. *J Proteome Res* **2021,** *20* (1), 270-278.
120. Marcoux, J.; Champion, T.; Colas, O.; Wagner-Rousset, E.; Corvaia, N.; Van Dorsselaer, A.; Beck, A.; Cianferani, S., Native mass spectrometry and ion mobility characterization of trastuzumab emtansine, a lysine-linked antibody drug conjugate. *Protein Sci* **2015,** *24* (8), 1210-23.
121. Desligniere, E.; Diemer, H.; Erb, S.; Coliat, P.; Pivot, X.; Detappe, A.; Hernandez-Alba, O.; Cianferani, S., A Combination of Native LC-MS Approaches for the Comprehensive Characterization of the Antibody-Drug Conjugate Trastuzumab Deruxtecan. *Front Biosci (Landmark Ed)* **2022,** *27* (10), 290.
122. Debaene, F.; Boeuf, A.; Wagner-Rousset, E.; Colas, O.; Ayoub, D.; Corvaia, N.; Van Dorsselaer, A.; Beck, A.; Cianferani, S., Innovative native MS methodologies for antibody drug conjugate characterization: High resolution native MS and IM-MS for average DAR and DAR distribution assessment. *Anal Chem* **2014,** *86* (21), 10674-83.
123. Fussl, F.; Strasser, L.; Carillo, S.; Bones, J., Native LC-MS for capturing quality attributes of biopharmaceuticals on the intact protein level. *Curr Opin Biotechnol* **2021,** *71*, 32-40.
124. Jones, J.; Pack, L.; Hunter, J. H.; Valliere-Douglass, J. F., Native size-exclusion chromatography-mass spectrometry: suitability for antibody-drug conjugate drug-to-antibody ratio quantitation across a range of chemotypes and drug-loading levels. *MAbs* **2020,** *12* (1), 1682895.
125. Duivelshof, B. L.; Beck, A.; Guillarme, D.; D'Atri, V., Bispecific antibody characterization by a combination of intact and site-specific/chain-specific LC/MS techniques. *Talanta* **2022,** *236*, 122836.
126. Desligniere, E.; Ehkirch, A.; Botzanowski, T.; Beck, A.; Hernandez-Alba, O.; Cianferani, S., Toward Automation of Collision-Induced Unfolding Experiments through Online Size Exclusion Chromatography Coupled to Native Mass Spectrometry. *Anal Chem* **2020,** *92* (19), 12900-12908.
127. Desligniere, E.; Ehkirch, A.; Duivelshof, B. L.; Toftevall, H.; Sjogren, J.; Guillarme, D.; D'Atri, V.; Beck, A.; Hernandez-Alba, O.; Cianferani, S., State-of-the-Art Native Mass Spectrometry and Ion Mobility Methods to Monitor Homogeneous Site-Specific Antibody-Drug Conjugates Synthesis. *Pharmaceuticals (Basel)* **2021,** *14* (6).
128. Larson, E. J.; Roberts, D. S.; Melby, J. A.; Buck, K. M.; Zhu, Y.; Zhou, S.; Han, L.; Zhang, Q.; Ge, Y., High-Throughput Multi-attribute Analysis of Antibody-Drug Conjugates Enabled by Trapped Ion Mobility Spectrometry and Top-Down Mass Spectrometry. *Anal Chem* **2021,** *93* (29), 10013-10021.
129. Wei, B.; Lantz, C.; Liu, W.; Viner, R.; Ogorzalek Loo, R. R.; Campuzano, I. D. G.; Loo, J. A., Added Value of Internal Fragments for Top-Down Mass Spectrometry of Intact Monoclonal Antibodies and Antibody-Drug Conjugates. *Anal Chem* **2023**.
130. Masson, G. R.; Jenkins, M. L.; Burke, J. E., An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opin Drug Discov* **2017,** *12* (10), 981-994.
131. Zhou, M.; Lantz, C.; Brown, K. A.; Ge, Y.; Pasa-Tolic, L.; Loo, J. A.; Lermyte, F., Higher-order structural characterisation of native proteins and complexes by top-down mass spectrometry. *Chem Sci* **2020,** *11* (48), 12918-12936.
132. Wu, D.; Piszczek, G., Measuring the affinity of protein-protein interactions on a single-molecule level by mass photometry. *Anal Biochem* **2020,** *592*, 113575.
133. Wu, D.; Piszczek, G., Rapid Determination of Antibody-Antigen Affinity by Mass Photometry. *J Vis Exp* **2021,** (168).
134. den Boer, M. A.; Lai, S. H.; Xue, X.; van Kampen, M. D.; Bleijlevens, B.; Heck, A. J. R., Comparative Analysis of Antibodies and Heavily Glycosylated Macromolecular Immune Complexes by Size-Exclusion Chromatography Multi-Angle Light Scattering, Native Charge Detection Mass Spectrometry, and Mass Photometry. *Anal Chem* **2022,** *94* (2), 892-900.
135. FENN, J. B.; MANN, M.; MENG, C. K.; WONG, S. F.; WHITEHOUSE, C. M., Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989,** *246*, 64-71.
136. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T., Protein and Polymer Analyses up to mlz 100 000 by Laser Ionization Time-of-flight Mass Spectrometry. *Rapid Commun Mass Spectrom* **1988,** *2* (8), 151-153.

137. Katta, V.; Chait, B. T., Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J. Am. Chem. Soc.* **1991,** *113*, 8534–8535.
138. Ganem, B.; Li, Y. T.; Henion, J. D., Detection of noncovalent receptor-ligand complexes by mass spectrometry. *J. Am. Chem. Soc.* **1991,** *113* (16), 6294–6296.
139. Smith, R. D.; Light-Wahl, K. J.; Winger, B. E.; Loo, J. A., Preservation of non-covalent associations in electrospray ionization mass spectrometry: Multiply charged polypeptide and protein dimers. *Journal of mass spectrometry* **1992,** *27*.
140. Winger, B. E.; Light-Wahl, K. J.; Loo, R. R. O.; Udseth, H. R.; Smith, R. D., Observation and Implications of High Mass-to-Charge Ratio Ions from Electrospray Ionization Mass Spectrometry *ASMS* **1993,** *4*, 536-545.
141. Collings, B. A.; Douglas, D. J., An extended mass range quadrupole for electrospray mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **1997**, 121-127.
142. Dodonov, A. F. C., I. V.; Laiko, V. V. , Atmospheric pressure ionization time-of-flight mass spectrometer. *International Mass Spectrometry Conference; Amsterdam.* **1991**.
143. Fitzgerald, M. C.; Chernushevich, I.; Standing, K. G.; Whitman, C. P.; Kent, S. B., Probing the oligomeric structure of an enzyme by electrospray ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci U S A.* **1996,** *14*.
144. Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H., High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Commun Mass Spectrom.* **1996,** *10* (8).
145. Schmidt, A.; Bahr, U.; Karas, M., Influence of Pressure in the First Pumping Stage on Analyte Desolvation and Fragmentation in Nano-ESI MS. *Anal. Chem.* **2001,** *73*.
146. Sobott, F.; Hernández, H.; McCammon, M. G.; Tito, M. A.; Robinson, C. V., A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal Chem* **2002,** *74* (6).
147. Fort, K. L.; van de Waterbeemd, M.; Boll, D.; Reinhardt-Szyba, M.; Belov, M. E.; Sasaki, E.; Zschoche, R.; Hilvert, D.; Makarov, A. A.; Heck, A. J. R., Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. *Analyst* **2017,** *143* (1), 100-105.
148. Hecht, E. S.; Scigelova, M.; Eliuk, S.; Makarov, A., Fundamentals and Advances of Orbitrap Mass Spectrometry. **2019**, 1-40.
149. Desligniere, E.; Rolland, A.; Ebberink, E.; Yin, V.; Heck, A. J. R., Orbitrap-Based Mass and Charge Analysis of Single Molecules. *Acc Chem Res* **2023,** *56* (12), 1458-1468.
150. Heck, A. J.; Van Den Heuvel, R. H., Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom Rev* **2004,** *23* (5), 368-89.
151. Martin Samonig; Swart, R., SEC-MS with Volatile Buffers for Characterization of Bipharmaceuticals. *Thermo Fisher app note* **2016**.
152. Le-Minh, V.; Halgand, F.; Van der Rest, G.; Taverna, M.; Smadja, C., Conformation assessment of therapeutic monoclonal antibodies by SEC-MS: Unravelling analytical biases for application to quality control. *J Pharm Biomed Anal* **2020,** *185*, 113252.
153. Goyon, A.; Beck, A.; Colas, O.; Sandra, K.; Guillarme, D.; Fekete, S., Evaluation of size exclusion chromatography columns packed with sub-3mum particles for the analysis of biopharmaceutical proteins. *J Chromatogr A* **2017,** *1498*, 80-89.
154. Chen, B.; Lin, Z.; Alpert, A. J.; Fu, C.; Zhang, Q.; Pritts, W. A.; Ge, Y., Online Hydrophobic Interaction Chromatography-Mass Spectrometry for the Analysis of Intact Monoclonal Antibodies. *Anal Chem* **2018,** *90* (12), 7135-7138.
155. Yan, Y.; Liu, A. P.; Wang, S.; Daly, T. J.; Li, N., Ultrasensitive Characterization of Charge Heterogeneity of Therapeutic Monoclonal Antibodies Using Strong Cation Exchange Chromatography Coupled to Native Mass Spectrometry. *Anal Chem* **2018,** *90* (21), 13013-13020.

156. Fussl, F.; Trappe, A.; Carillo, S.; Jakes, C.; Bones, J., Comparative Elucidation of Cetuximab Heterogeneity on the Intact Protein Level by Cation Exchange Chromatography and Capillary Electrophoresis Coupled to Mass Spectrometry. *Anal Chem* **2020,** *92* (7), 5431-5438.
157. Koza, S. M.; Yu, Y. Q., Rapid Size Variant Analysis of Monoclonal Antibodies Using UPLC™ and HPLC Compatible MaxPeak™ Premier Protein SEC Columns. *Waters* **2022**.
158. Vimer, S.; Ben-Nissan, G.; Marty, M.; Fleishman, S. J.; Sharon, M., Direct-MS analysis of antibody-antigen complexes. *Proteomics* **2021,** *21* (21-22), e2000300.
159. Valliere-Douglass, J. F.; McFee, W. A.; Salas-Solano, O., Native intact mass determination of antibodies conjugated with monomethyl Auristatin E and F at interchain cysteine residues. *Anal Chem* **2012,** *84* (6), 2843-9.
160. Larson, E. J.; Zhu, Y.; Wu, Z.; Chen, B.; Zhang, Z.; Zhou, S.; Han, L.; Zhang, Q.; Ge, Y., Rapid Analysis of Reduced Antibody Drug Conjugate by Online LC-MS/MS with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal Chem* **2020,** *92* (22), 15096-15103.
161. Corradini, D., Buffer additives other than the surfactant sodium dodecyl sulfate for protein separations by capillary electrophoresis. **1997**.
162. Jooss, K.; McGee, J. P.; Melani, R. D.; Kelleher, N. L., Standard procedures for native CZE-MS of proteins and protein complexes up to 800 kDa. *Electrophoresis* **2021,** *42* (9-10), 1050-1059.
163. Belov, A. M.; Viner, R.; Santos, M. R.; Horn, D. M.; Bern, M.; Karger, B. L.; Ivanov, A. R., Analysis of Proteins, Protein Complexes, and Organellar Proteomes Using Sheathless Capillary Zone Electrophoresis - Native Mass Spectrometry. *J Am Soc Mass Spectrom* **2017,** *28* (12), 2614-2634.
164. Astefanei, A.; Dapic, I.; Camenzuli, M., Different Stationary Phase Selectivities and Morphologies for Intact Protein Separations. *Chromatographia* **2017,** *80* (5), 665-687.
165. Ehkirch, A.; Goyon, A.; Hernandez-Alba, O.; Rouviere, F.; D'Atri, V.; Dreyfus, C.; Haeuw, J. F.; Diemer, H.; Beck, A.; Heinisch, S.; Guillarme, D.; Cianferani, S., A Novel Online Four-Dimensional SECxSEC-IMxMS Methodology for Characterization of Monoclonal Antibody Size Variants. *Anal Chem* **2018,** *90* (23), 13929-13937.
166. Fekete, S.; Beck, A.; Veuthey, J. L.; Guillarme, D., Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *J Pharm Biomed Anal* **2014,** *101*, 161-73.
167. Bailey, A. O.; Han, G.; Phung, W.; Gazis, P.; Sutton, J.; Josephs, J. L.; Sandoval, W., Charge variant native mass spectrometry benefits mass precision and dynamic range of monoclonal antibody intact mass analysis. *MAbs* **2018,** *10* (8), 1214-1225.
168. Leblanc, Y.; Bihoreau, N.; Chevreux, G., Characterization of Human Serum Albumin isoforms by ion exchange chromatography coupled on-line to native mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* **2018,** *1095*, 87-93.
169. Samantha Ippoliti, A. S., Matthew A. Lauber, Ying Qing Yu, Online IEX-MS of mAb Charge Variants
Using a BioResolve SCX mAb Column,
IonHance CX-MS pH Concentrates, and
BioAccord System. **2019**.
170. Matsuda, Y.; Kliman, M.; Mendelsohn, B. A., Application of Native Ion Exchange Mass Spectrometry to Intact and Subunit Analysis of Site-Specific Antibody-Drug Conjugates Produced by AJICAP First Generation Technology. *J Am Soc Mass Spectrom* **2020**.
171. Cummins, P. M.; Rochfort, K. D.; O'Connor, B. F., Ion-Exchange Chromatography: Basic Principles and Application. *Methods Mol Biol* **2017,** *1485*, 209-223.
172. Millan-Martin, S.; Carillo, S.; Fussl, F.; Sutton, J.; Gazis, P.; Cook, K.; Scheffler, K.; Bones, J., Optimisation of the use of sliding window deconvolution for comprehensive characterisation of trastuzumab and adalimumab charge variants by native high resolution mass spectrometry. *Eur J Pharm Biopharm* **2021,** *158*, 83-95.
173. Fussl, F.; Cook, K.; Scheffler, K.; Farrell, A.; Mittermayr, S.; Bones, J., Charge Variant Analysis of Monoclonal Antibodies Using Direct Coupled pH Gradient Cation Exchange Chromatography to High-Resolution Native Mass Spectrometry. *Anal Chem* **2018,** *90* (7), 4669-4676.

174. Goyon, A.; McDonald, D.; Fekete, S.; Guillarme, D.; Stella, C., Development of an innovative salt-mediated pH gradient cation exchange chromatography method for the characterization of therapeutic antibodies. *J Chromatogr B Analyt Technol Biomed Life Sci* **2020,** *1160*, 122379.
175. Leblanc, Y.; Faid, V.; Lauber, M. A.; Wang, Q.; Bihoreau, N.; Chevreux, G., A generic method for intact and subunit level characterization of mAb charge variants by native mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* **2019,** *1133*, 121814.
176. Bobaly, B.; Fleury-Souverain, S.; Beck, A.; Veuthey, J. L.; Guillarme, D.; Fekete, S., Current possibilities of liquid chromatography for the characterization of antibody-drug conjugates. *J Pharm Biomed Anal* **2018,** *147*, 493-505.
177. D'Atri, V.; Fekete, S.; Beck, A.; Lauber, M.; Guillarme, D., Hydrophilic Interaction Chromatography Hyphenated with Mass Spectrometry: A Powerful Analytical Tool for the Comparison of Originator and Biosimilar Therapeutic Monoclonal Antibodies at the Middle-up Level of Analysis. *Anal Chem* **2017,** *89* (3), 2086-2092.
178. Ehkirch, A.; D'Atri, V.; Rouviere, F.; Hernandez-Alba, O.; Goyon, A.; Colas, O.; Sarrut, M.; Beck, A.; Guillarme, D.; Heinisch, S.; Cianferani, S., An Online Four-Dimensional HICxSEC-IMxMS Methodology for Proof-of-Concept Characterization of Antibody Drug Conjugates. *Anal Chem* **2018,** *90* (3), 1578-1586.
179. Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J., Jr.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T., Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom Rev* **2019,** *38* (3), 291-320.
180. Dodds, J. N.; Baker, E. S., Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *Journal of the American Society for Mass Spectrometry* **2019,** *30* (11), 2185-2195.
181. May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L.; Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A., Conformational ordering of biomolecules in the gas phase: nitrogen collision cross sections measured on a prototype high resolution drift tube ion mobility-mass spectrometer. *Anal Chem* **2014,** *86* (4), 2107-16.
182. Hernandez-Alba, O.; Wagner-Rousset, E.; Beck, A.; Cianferani, S., Native Mass Spectrometry, Ion Mobility, and Collision-Induced Unfolding for Conformational Characterization of IgG4 Monoclonal Antibodies. *Anal Chem* **2018,** *90* (15), 8865-8872.
183. Desligniere, E.; Ollivier, S.; Beck, A.; Ropartz, D.; Rogniaux, H.; Cianferani, S., Benefits and Limitations of High-Resolution Cyclic IM-MS for Conformational Characterization of Native Therapeutic Monoclonal Antibodies. *Anal Chem* **2023,** *95* (8), 4162-4171.
184. Giles, K.; Williams, J. P.; Campuzano, I., Enhancements in travelling wave ion mobility resolution. *Rapid Commun Mass Spectrom* **2011,** *25* (11), 1559-66.
185. Shvartsburg, A. A.; Smith, R. D., Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **2008,** *80*.
186. Hernandez, D. R.; Debord, J. D.; Ridgeway, M. E.; Kaplan, D. A.; Park, M. A.; Fernandez-Lima, F., Ion dynamics in a trapped ion mobility spectrometer. *Analyst* **2014,** *139* (8), 1913-21.
187. Michelmann, K.; Silveira, J. A.; Ridgeway, M. E.; Park, M. A., Fundamentals of trapped ion mobility spectrometry. *J Am Soc Mass Spectrom* **2015,** *26* (1), 14-24.
188. Kolakowski, B. M.; Mester, Z., Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). *Analyst* **2007,** *132* (9), 842-64.
189. Cooper, H. J., To What Extent is FAIMS Beneficial in the Analysis of Proteins? *J Am Soc Mass Spectrom* **2016,** *27* (4), 566-77.

190. Gerbasi, V. R.; Melani, R. D.; Abbatiello, S. E.; Belford, M. W.; Huguet, R.; McGee, J. P.; Dayhoff, D.; Thomas, P. M.; Kelleher, N. L., Deeper Protein Identification Using Field Asymmetric Ion Mobility Spectrometry in Top-Down Proteomics. *Anal Chem* **2021,** *93* (16), 6323-6328.
191. Pease, L. F., 3rd; Elliott, J. T.; Tsai, D. H.; Zachariah, M. R.; Tarlov, M. J., Determination of protein aggregation with differential mobility analysis: application to IgG antibody. *Biotechnol Bioeng* **2008,** *101* (6), 1214-22.
192. Rabuck, J. N.; Hyung, S. J.; Ko, K. S.; Fox, C. C.; Soellner, M. B.; Ruotolo, B. T., Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry. *Anal Chem* **2013,** *85* (15), 6995-7002.
193. Tian, Y.; Lippens, J. L.; Netirojjanakul, C.; Campuzano, I. D. G.; Ruotolo, B. T., Quantitative collision-induced unfolding differentiates model antibody-drug conjugates. *Protein Sci* **2019,** *28* (3), 598-608.
194. Campuzano, I. D. G.; Robinson, J. H.; Hui, J. O.; Shi, S. D.; Netirojjanakul, C.; Nshanian, M.; Egea, P. F.; Lippens, J. L.; Bagal, D.; Loo, J. A.; Bern, M., Native and Denaturing MS Protein Deconvolution for Biopharma: Monoclonal Antibodies and Antibody-Drug Conjugates to Polydisperse Membrane Proteins and Beyond. *Anal Chem* **2019,** *91* (15), 9472-9480.
195. Paul, S. S.; Lyons, A.; Kirchner, R.; Woodside, M. T., Quantifying Oligomer Populations in Real Time during Protein Aggregation Using Single-Molecule Mass Photometry. *ACS Nano* **2022,** *16* (10), 16462-16470.
196. Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E. G.; Collier, M. P.; Chandler, S. A.; Tkachenko, O.; Allen, J.; Crispin, M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Eichmann, C.; Selenko, P.; Frey, L.; Riek, R.; Galpin, M. R.; Struwe, W. B.; Benesch, J. L. P.; Kukura, P., Quantitative mass imaging of single
biological macromolecules. *Biophysics* **2018,** *360*, 123-427.
197. Soltermann, F.; Foley, E. D. B.; Pagnoni, V.; Galpin, M.; Benesch, J. L. P.; Kukura, P.; Struwe, W. B., Quantifying Protein-Protein Interactions by Molecular Counting with Mass Photometry. *Angew Chem Int Ed Engl* **2020,** *59* (27), 10774-10779.
198. Sonn-Segev, A.; Belacic, K.; Bodrug, T.; Young, G.; VanderLinden, R. T.; Schulman, B. A.; Schimpf, J.; Friedrich, T.; Dip, P. V.; Schwartz, T. U.; Bauer, B.; Peters, J. M.; Struwe, W. B.; Benesch, J. L. P.; Brown, N. G.; Haselbach, D.; Kukura, P., Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat Commun* **2020,** *11* (1), 1772.
199. Wu, D.; Hwang, P.; Li, T.; Piszczek, G., Rapid characterization of adeno-associated virus (AAV) gene therapy vectors by mass photometry. *Gene Ther* **2022,** *29* (12), 691-697.
200. Lermyte, F.; Tsybin, Y. O.; O'Connor, P. B.; Loo, J. A., Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *J Am Soc Mass Spectrom* **2019,** *30* (7), 1149-1157.
201. Cui, W.; Rohrs, H. W.; Gross, M. L., Top-down mass spectrometry: recent developments, applications and perspectives. *Analyst* **2011,** *136* (19), 3854-64.
202. Toby, T. K.; Fornelli, L.; Kelleher, N. L., Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu Rev Anal Chem (Palo Alto Calif)* **2016,** *9* (1), 499-519.
203. Kellie, J. F.; Tran, J. C.; Lee, J. E.; Ahlf, D. R.; Thomas, H. M.; Ntai, I.; Catherman, A. D.; Durbin, K. R.; Zamdborg, L.; Vellaichamy, A.; Thomas, P. M.; Kelleher, N. L., The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. *Mol Biosyst* **2010,** *6* (9), 1532-9.
204. Resemann, A.; Wunderlich, D.; Rothbauer, U.; Warscheid, B.; Leonhardt, H.; Fuchser, J.; Kuhlmann, K.; Suckau, D., Top-Down de Novo Protein Sequencing of a 13.6 kDa Camelid Single Heavy Chain Antibody by Matrix-Assisted Laser Desorption Ionization-Time-of-Flight/Time-of-Flight Mass Spectrometry. *J Proteome Res* **2010,** *9* (4), 1854-63.
205. Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R., Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal Chem* **2011,** *83* (23), 8919-27.

206. Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O., Analysis of intact monoclonal antibody IgG1 by electron transfer dissociation Orbitrap FTMS. *Mol Cell Proteomics* **2012,** *11* (12), 1758-67.
207. Fornelli, L.; Ayoub, D.; Aizikov, K.; Liu, X.; Damoc, E.; Pevzner, P. A.; Makarov, A.; Beck, A.; Tsybin, Y. O., Top-down analysis of immunoglobulin G isotypes 1 and 2 with electron transfer dissociation on a high-field Orbitrap mass spectrometer. *J Proteomics* **2017,** *159*, 67-76.
208. Chen, B.; Lin, Z.; Zhu, Y.; Jin, Y.; Larson, E.; Xu, Q.; Fu, C.; Zhang, Z.; Zhang, Q.; Pritts, W. A.; Ge, Y., Middle-Down Multi-Attribute Analysis of Antibody-Drug Conjugates with Electron Transfer Dissociation. *Anal Chem* **2019,** *91* (18), 11661-11669.
209. Fornelli, L.; Ayoub, D.; Srzentic, K.; Nagornov, K.; Kozhinov, A.; Gasilova, N.; Menin, L.; Beck, A.; Tsybin, Y., Structural Analysis of Monoclonal Antibodies with Top-down and Middle-down Electron Transfer Dissociation Mass Spectrometry: The First Decade. *Chimia* **2022,** *76* (1-2), 114.
210. Ho, C.; Lam, C.; Chan, M.; Cheung, R.; Law, L.; Lit, L.; Ng, K.; Suen, M.; Tai, H., Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *Clin Biochem Rev* **2003**.
211. Mclafferty, F., High-Resolution Tandem FT Mass Spectrometry above 10 kDa. *Acc. Chem. Res.* **1994,** *27* (379-386).
212. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W., Electron Capture Dissociation of Multiply ChargedProtein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998,** *120*.
213. Meng, F.; Cargile, B. J.; Miller, L. M.; Forbes, A. J.; Johnson, J. R.; L.Kelleher, N., Informatics and multiplexing of intact protein identification in bacteria and the archaea. *Nature publishing group* **2001**.
214. LeDuc, R. D.; Taylor, G. K.; Kim, Y. B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L., ProSight PTM: an integrated environment for protein identification and characterization by top-down mass spectrometry. *Nucleic Acids Res* **2004,** *32* (Web Server issue), W340-5.
215. Zamdborg, L.; LeDuc, R. D.; Glowacz, K. J.; Kim, Y. B.; Viswanathan, V.; Spaulding, I. T.; Early, B. P.; Bluhm, E. J.; Babai, S.; Kelleher, N. L., ProSight PTM 2.0: improved protein identification and characterization for top down mass spectrometry. *Nucleic Acids Res* **2007,** *35* (Web Server issue), W701-6.
216. Fellers, R. T.; Greer, J. B.; Early, B. P.; Yu, X.; LeDuc, R. D.; Kelleher, N. L.; Thomas, P. M., ProSight Lite: graphical software to analyze top-down mass spectrometry data. *Proteomics* **2015,** *15* (7), 1235-8.
217. Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.-H.; Begley, T. P.; McLafferty, F. W., Top Down Characterization of Larger Proteins (45 kDa) by Electron Capture Dissociation Mass Spectrometry. *JACS articles* **2002,** *124* (4).
218. Mao, Y.; Valeja, S. G.; Rouse, J. C.; Hendrickson, C. L.; Marshall, A. G., Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-Fourier transform ion cyclotron resonance-mass spectrometry. *Anal Chem* **2013,** *85* (9), 4239-46.
219. Fornelli, L.; Ayoub, D.; Aizikov, K.; Beck, A.; Tsybin, Y. O., Middle-down analysis of monoclonal antibodies with electron transfer dissociation orbitrap fourier transform mass spectrometry. *Anal Chem* **2014,** *86* (6), 3005-12.
220. Macek, B.; Waanders, L. F.; Olsen, J. V.; Mann, M., Top-down Protein Sequencing and MS3 on a Hybrid Linear Quadrupole Ion Trap-Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* **2006,** *5* (5), 949-958.
221. Brunner, A. M.; Lossl, P.; Liu, F.; Huguet, R.; Mullen, C.; Yamashita, M.; Zabrouskov, V.; Makarov, A.; Altelaar, A. F.; Heck, A. J., Benchmarking multiple fragmentation methods on an orbitrap fusion for top-down phospho-proteoform characterization. *Anal Chem* **2015,** *87* (8), 4152-8.
222. Fornelli, L.; Srzentic, K.; Huguet, R.; Mullen, C.; Sharma, S.; Zabrouskov, V.; Fellers, R. T.; Durbin, K. R.; Compton, P. D.; Kelleher, N. L., Accurate Sequence Analysis of a Monoclonal Antibody by Top-Down and Middle-Down Orbitrap Mass Spectrometry Applying Multiple Ion Activation Techniques. *Anal Chem* **2018,** *90* (14), 8421-8429.
223. Fort, K. L.; Cramer, C. N.; Voinov, V. G.; Vasil'ev, Y. V.; Lopez, N. I.; Beckman, J. S.; Heck, A. J. R., Exploring ECD on a Benchtop Q Exactive Orbitrap Mass Spectrometer. *J Proteome Res* **2018,** *17* (2), 926-933.

224. Shaw, J. B.; Malhan, N.; Vasil'ev, Y. V.; Lopez, N. I.; Makarov, A.; Beckman, J. S.; Voinov, V. G., Sequencing Grade Tandem Mass Spectrometry for Top-Down Proteomics Using Hybrid Electron Capture Dissociation Methods in a Benchtop Orbitrap Mass Spectrometer. *Anal Chem* **2018,** *90* (18), 10819-10827.
225. Kline, J. T.; Mullen, C.; Durbin, K. R.; Oates, R. N.; Huguet, R.; Syka, J. E. P.; Fornelli, L., Sequential Ion-Ion Reactions for Enhanced Gas-Phase Sequencing of Large Intact Proteins in a Tribrid Orbitrap Mass Spectrometer. *J Am Soc Mass Spectrom* **2021,** *32* (9), 2334-2345.
226. Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S., Ultrafast Ultraviolet Photodissociation at 193 nm and its Applicability to Proteomic Workflow. *Journal of proteom research* **2010,** *9*, 4205–4214.
227. Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S., Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation. *J Am Chem Soc* **2013,** *135* (34), 12646-51.
228. Cotham, V. C.; Brodbelt, J. S., Characterization of Therapeutic Monoclonal Antibodies at the Subunit-Level using Middle-Down 193 nm Ultraviolet Photodissociation. *Anal Chem* **2016,** *88* (7), 4004-13.
229. Lantz, C.; Zenaidee, M. A.; Wei, B.; Hemminger, Z.; Ogorzalek Loo, R. R.; Loo, J. A., ClipsMS: An Algorithm for Analyzing Internal Fragments Resulting from Top-Down Mass Spectrometry. *J Proteome Res* **2021,** *20* (4), 1928-1935.
230. MCLAFFERTY, F. W.; BRYCE, T. A., Metastable-ion Characteristics: Characterization of Isomeric Molecules. **1967**.
231. Dongre, s. R.; Jones, J. L.; Somogyii, A. d.; Wysocki, V. H., Influence of Peptide Composition, Gas-Phase Basicity, andChemical Modification on Fragmentation Efficiency: Evidencefor the Mobile Proton Model. *J. Am. Chem. Soc.* **1996,** *118*.
232. Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A., Mobile and localized protons: a framework for understanding peptide dissociation. *Journal of mass spectrometry* **2000**.
233. McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K., Electron Capture Dissociation of Gaseous Multiply Charged Ions by Fourier-Transform Ion Cyclotron Resonance. *ASMS* **2000**.
234. Jones, A. W.; Cooper, H. J., Probing the mechanisms of electron capture dissociation mass spectrometry with nitrated peptides. *Phys Chem Chem Phys* **2010,** *12* (41), 13394-9.
235. Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E., Electron transfer dissociation of peptide anions. *J Am Soc Mass Spectrom* **2005,** *16* (6), 880-2.
236. Riley, N. M.; Mullen, C.; R.Weisbrod, C.; Sharma, S.; Senko, M. W.; Zabrouskov, V.; S.Westphall, M.; Syka, J. E. P.; Coon, J. J., Enhanced Dissociation of Intact Proteins with High Capacity Electron Transfer Dissociation. *JASMS* **2015,** *27* (3), 520–531.
237. Liu, J.; McLuckey, S. A., Electron Transfer Dissociation: Effects of Cation Charge State on Product Partitioning in Ion/Ion Electron Transfer to Multiply Protonated Polypeptides. *Int J Mass Spectrom* **2012,** *330-332*, 174-181.
238. Lermyte, F.; Łącki, M. K.; Valkenborg, D.; Baggerman, G.; Gambin, A.; Sobott, F., Understanding reaction pathways in top-down ETD by dissecting isotope distributions: A mammoth task. *International Journal of Mass Spectrometry* **2015,** *390*, 146-154.
239. Lermyte, F.; Lacki, M. K.; Valkenborg, D.; Gambin, A.; Sobott, F., Conformational Space and Stability of ETD Charge Reduction Products of Ubiquitin. *J Am Soc Mass Spectrom* **2017,** *28* (1), 69-76.
240. Riley, N. M.; Westphall, M. S.; Coon, J. J., Activated Ion Electron Transfer Dissociation for Improved Fragmentation of Intact Proteins. *Anal Chem* **2015,** *87* (14), 7109-16.
241. Riley, N. M.; Westphall, M. S.; Coon, J. J., Activated Ion-Electron Transfer Dissociation Enables Comprehensive Top-Down Protein Fragmentation. *J Proteome Res* **2017,** *16* (7), 2653-2659.
242. Riley, N. M.; Sikora, J. W.; Seckler, H. S.; Greer, J. B.; Fellers, R. T.; LeDuc, R. D.; Westphall, M. S.; Thomas, P. M.; Kelleher, N. L.; Coon, J. J., The Value of Activated Ion Electron Transfer Dissociation for High-Throughput Top-Down Characterization of Intact Proteins. *Anal Chem* **2018,** *90* (14), 8553-8560.

243. Yu, Q.; Wang, B.; Chen, Z.; Urabe, G.; Glover, M. S.; Shi, X.; Guo, L. W.; Kent, K. C.; Li, L., Electron-Transfer/Higher-Energy Collision Dissociation (EThcD)-Enabled Intact Glycopeptide/Glycoproteome Characterization. *J Am Soc Mass Spectrom* **2017,** *28* (9), 1751-1764.
244. Brodbelt, J. S., Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem Soc Rev* **2014,** *43* (8), 2757-83.
245. Madsen, J. A.; Brodbelt, J. S., Comparison of infrared multiphoton dissociation and collision-induced dissociation of supercharged peptides in ion traps. *J Am Soc Mass Spectrom* **2009,** *20* (3), 349-58.
246. Raspopov, S. A.; El-Faramawy, A.; Thomson, B. A.; Siu, K. W. M., Infrared Multiphoton Dissociation in Quadrupole Time-of-Flight Mass Spectrometry: Top-Down Characterization of Proteins. *Anal. Chem.* **2006,** *78*.
247. R, R. J., The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. *J Am Soc Mass Spectrom* **2017,** *28* (9), 1823-1826.
248. Macias, L. A.; Wang, X.; Davies, B. W.; Brodbelt, J. S., Mapping paratopes of nanobodies using native mass spectrometry and ultraviolet photodissociation. *Chem Sci* **2022,** *13* (22), 6610-6618.
249. Dunham, S. D.; Sanders, J. D.; Holden, D. D.; Brodbelt, J. S., Improving the Center Section Sequence Coverage of Large Proteins Using Stepped-Fragment Ion Protection Ultraviolet Photodissociation. *J Am Soc Mass Spectrom* **2022,** *33* (3), 446-456.
250. Macias, L. A.; Brodbelt, J. S., Investigation of Product Ions Generated by 193 nm Ultraviolet Photodissociation of Peptides and Proteins Containing Disulfide Bonds. *J Am Soc Mass Spectrom* **2022,** *33* (7), 1315-1324.
251. Lanzillotti, M.; Brodbelt, J. S., Comparison of Top-Down Protein Fragmentation Induced by 213 and 193 nm UVPD. *J Am Soc Mass Spectrom* **2023**.
252. Cannon, J. R.; Holden, D. D.; Brodbelt, J. S., Hybridizing ultraviolet photodissociation with electron transfer dissociation for intact protein characterization. *Anal Chem* **2014,** *86* (21), 10970-7.
253. Hale, O. J.; Cooper, H. J., In situ mass spectrometry analysis of intact proteins and protein complexes from biological substrates. *Biochem Soc Trans* **2020,** *48* (1), 317-326.
254. Huguet, R.; Mullen, C.; Srzentic, K.; Greer, J. B.; Fellers, R. T.; Zabrouskov, V.; Syka, J. E. P.; Kelleher, N. L.; Fornelli, L., Proton Transfer Charge Reduction Enables High-Throughput Top-Down Analysis of Large Proteoforms. *Anal Chem* **2019,** *91* (24), 15732-15739.
255. James L. Stephenson, J.; McLuckey, S. A., Simplification of Product Ion Spectra Derived from Multiply Charged Parent Ions via Ion/Ion Chemistry. *Anal. Chem.* **1998**.
256. Bailey, A. O.; Huguet, R.; Mullen, C.; Syka, J. E. P.; Russell, W. K., Ion-Ion Charge Reduction Addresses Multiple Challenges Common to Denaturing Intact Mass Analysis. *Anal Chem* **2022,** *94* (9), 3930-3938.
257. Zhong, J.; Sun, Y.; Xie, M.; Peng, W.; Zhang, C.; Wu, F. X.; Wang, J., Proteoform characterization based on top-down mass spectrometry. *Brief Bioinform* **2021,** *22* (2), 1729-1750.
258. Haverland, N. A.; Skinner, O. S.; Fellers, R. T.; Tariq, A. A.; Early, B. P.; LeDuc, R. D.; Fornelli, L.; Compton, P. D.; Kelleher, N. L., Defining Gas-Phase Fragmentation Propensities of Intact Proteins During Native Top-Down Mass Spectrometry. *J Am Soc Mass Spectrom* **2017,** *28* (6), 1203-1215.
259. Lantz, C.; Wei, B.; Zhao, B.; Jung, W.; Goring, A. K.; Le, J.; Miller, J.; Loo, R. R. O.; Loo, J. A., Native Top-Down Mass Spectrometry with Collisionally Activated Dissociation Yields Higher-Order Structure Information for Protein Complexes. *J Am Chem Soc* **2022,** *144* (48), 21826-21830.
260. Lermyte, F.; Habeck, T.; Brown, K.; Des Soye, B.; Lantz, C.; Zhou, M.; Alam, N.; Hossain, M. A.; Jung, W.; Keener, J.; Volny, M.; Wilson, J.; Ying, Y.; Agar, J.; Danis, P.; Ge, Y.; Kelleher, N.; Li, H.; Loo, J.; Marty, M.; Pasa-Tolic, L.; Sandoval, W., Top-down mass spectrometry of native proteoforms and their complexes: A community study. *Res Sq* **2023**.
261. Greisch, J. F.; den Boer, M. A.; Lai, S. H.; Gallagher, K.; Bondt, A.; Commandeur, J.; Heck, A. J. R., Extending Native Top-Down Electron Capture Dissociation to MDa Immunoglobulin Complexes Provides Useful Sequence Tags Covering Their Critical Variable Complementarity-Determining Regions. *Anal Chem* **2021,** *93* (48), 16068-16075.

262. Zhang, J.; Reza Malmirchegini, G.; Clubb, R. T.; Loo, J. A., Native top-down mass spectrometry for the structural characterization of human hemoglobin. *Eur J Mass Spectrom (Chichester)* **2015,** *21* (3), 221-31.
263. Ting, Y. S.; Egertson, J. D.; Payne, S. H.; Sangtae Kim§; MacLean, B.; Ka, L.; Aebersold, R.; Smith, R. D.; Noble, W. S.; MacCoss, M. J., Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data. *The American Society for Biochemistry and Molecular Biology* **2015**.
264. Brown, K. A.; Wilson, D. J., Bottom-up hydrogen deuterium exchange mass spectrometry: data analysis and interpretation. *Analyst* **2017,** *142* (16), 2874-2886.
265. Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jorgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D., Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods* **2019,** *16* (7), 595-602.
266. Uhrik, L.; Hernychova, L.; Muller, P.; Kalathiya, U.; Lisowska, M. M.; Kocikowski, M.; Parys, M.; Faktor, J.; Nekulova, M.; Nortcliffe, C.; Zatloukalova, P.; Ruetgen, B.; Fahraeus, R.; Ball, K. L.; Argyle, D. J.; Vojtesek, B.; Hupp, T. R., Hydrogen deuterium exchange mass spectrometry identifies the dominant paratope in CD20 antigen binding to the NCD1.2 monoclonal antibody. *Biochem J* **2021,** *478* (1), 99-120.
267. Zhang, M. M.; Huang, R. Y.; Beno, B. R.; Deyanova, E. G.; Li, J.; Chen, G.; Gross, M. L., Epitope and Paratope Mapping of PD-1/Nivolumab by Mass Spectrometry-Based Hydrogen-Deuterium Exchange, Cross-linking, and Molecular Docking. *Anal Chem* **2020,** *92* (13), 9086-9094.
268. Gotze, M.; Iacobucci, C.; Ihling, C. H.; Sinz, A., A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-wide Protein Interactions. *Anal Chem* **2019,** *91* (15), 10236-10244.
269. Yu, C.; Huang, L., Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal Chem* **2018,** *90* (1), 144-165.
270. Li, K. S.; Shi, L.; Gross, M. L., Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc Chem Res* **2018,** *51* (3), 736-744.
271. Liu, X. R.; Rempel, D. L.; Gross, M. L., Protein higher-order-structure determination by fast photochemical oxidation of proteins and mass spectrometry analysis. *Nat Protoc* **2020,** *15* (12), 3942-3970.
272. Shuken, S. R.; Rutledge, J.; Iram, T.; Losada, P. M.; Wilson, E. N.; Andreasson, K. I.; Leib, R. D.; Wyss-Coray, T., Limited Proteolysis-Mass Spectrometry Reveals Aging-Associated Changes in Cerebrospinal Fluid Protein Abundances and Structures. *Nat Aging* **2022,** *2* (5), 379-388.
273. Holfeld, A.; Schuster, D.; Sesterhenn, F.; Stalder, P.; Haenseler, W.; Barrio-Hernandez, I.; Ghosh, D.; Vowles, J.; Cowley, S. A.; Nagel, L.; Khanppnavar, B.; Beltrao, P.; Korkhov, V. M.; Riek, R.; Souza, N. d.; Picotti, P., Systematic identification of structure-specific protein–protein interactions. **2023**.
274. Waitt, G. M.; Xu, R.; Wisely, G. B.; Williams, J. D., Automated in-line gel filtration for native state mass spectrometry. *J Am Soc Mass Spectrom* **2008,** *19* (2), 239-45.
275. Cavanagh, J.; Benson, L. M.; Thompson, R.; Naylor, S., In-Line desalting Mass Spectrometry for the Study of Noncovalent biological complexes. *Anal. Chem.* **2003,** *75* ( ), 3281-3286.
276. Ma, F.; Raoufi, F.; Bailly, M. A.; Fayadat-Dilman, L.; Tomazela, D., Hyphenation of strong cation exchange chromatography to native mass spectrometry for high throughput online characterization of charge heterogeneity of therapeutic monoclonal antibodies. *MAbs* **2020,** *12* (1), 1763762.
277. Shi, R. L.; Xiao, G.; Dillon, T. M.; Ricci, M. S.; Bondarenko, P. V., Characterization of therapeutic proteins by cation exchange chromatography-mass spectrometry and top-down analysis. *MAbs* **2020,** *12* (1), 1739825.

278. Boeuf, A.; Debaene, F.; Ayoub, D.; Diemer, H.; Ehkirch, A.; Wagner-Rousset, E.; Dorsselaer, A. V.; Cianférani, S.; Beck, A., Mass Spectrometry-Based Strategies for Therapeutic Antibodies Extensive Characterization and Optimization (OptimAbs). *Structural Biology in Drug Discovery: Methods, Techniques, and Practices,* **2020,** *1*.
279. VanAernum, Z. L.; Busch, F.; Jones, B. J.; Jia, M.; Chen, Z.; Boyken, S. E.; Sahasrabuddhe, A.; Baker, D.; Wysocki, V. H., Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. *Nature Protocols* **2020,** *15* (3), 1132-1157.
280. Deslignière, E.; Ley, M.; Bourguet, M.; Ehkirch, A.; Botzanowski, T.; Erb, S.; Hernandez-Alba, O.; Cianférani, S., Pushing the limits of native MS: Online SEC-native MS for structural biology applications. *International Journal of Mass Spectrometry* **2021,** *461*, 116502.
281. Leblanc, Y.; Ramon, C.; Bihoreau, N.; Chevreux, G., Charge variants characterization of a monoclonal antibody by ion exchange chromatography coupled on-line to native mass spectrometry: Case study after a long-term storage at +5 degrees C. *J Chromatogr B Analyt Technol Biomed Life Sci* **2017,** *1048*, 130-139.
282. Trappe, A.; Fussl, F.; Carillo, S.; Zaborowska, I.; Meleady, P.; Bones, J., Rapid charge variant analysis of monoclonal antibodies to support lead candidate biopharmaceutical development. *J Chromatogr B Analyt Technol Biomed Life Sci* **2018,** *1095*, 166-176.
283. Tahallah, N.; Pinkse, M.; Maler, C. S.; Heck, A. J. R., The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument. *Rapid. Commun. Mass Spectrom.* **2001,** *15*, 596-601.
284. Grix, R.; Kutscher, R.; Li, G.; Grüner, U.; Wollnik, H., A Time-of-flight Mass Analyzer with High Resolving Power. *Rapid Commun Mass Spectrom* **1988,** *2* (5), 83-85.
285. Kebarle, P.; Verkerk, U. H., Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev* **2009,** *28* (6), 898-917.
286. Konermann, L., Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry. *J Am Soc Mass Spectrom* **2017,** *28* (9), 1827-1835.
287. Ventouri, I. K.; Malheiro, D. B. A.; Voeten, R. L. C.; Kok, S.; Honing, M.; Somsen, G. W.; Haselberg, R., Probing Protein Denaturation during Size-Exclusion Chromatography Using Native Mass Spectrometry. *Anal Chem* **2020,** *92* (6), 4292-4300.
288. Fussl, F.; Trappe, A.; Cook, K.; Scheffler, K.; Fitzgerald, O.; Bones, J., Comprehensive characterisation of the heterogeneity of adalimumab via charge variant analysis hyphenated on-line to native high resolution Orbitrap mass spectrometry. *MAbs* **2019,** *11* (1), 116-128.
289. Sanglier, S.; Atmanene, C.; Chevreux, G.; Dorsselaer, A. V., Nondenaturing Mass Spectrometry to Study Noncovalent Protein/Protein and Protein/Ligand Complexes: Technical Aspects and Application to the Determination of Binding Stoichiometries. *Methods Mol Biol* **2008,** *484*, 217-243

290. Ventouri, I. K.; Veelders, S.; Passamonti, M.; Endres, P.; Roemling, R.; Schoenmakers, P. J.; Somsen, G. W.; Haselberg, R.; Gargano, A. F. G., Micro-flow size-exclusion chromatography for enhanced native mass spectrometry of proteins and protein complexes. *Anal Chim Acta* **2023,** *1266*, 341324.
291. Duivelshof, B. L.; Denorme, S.; Sandra, K.; Liu, X.; Beck, A.; Lauber, M. A.; Guillarme, D.; D'Atri, V., Quantitative N-Glycan Profiling of Therapeutic Monoclonal Antibodies Performed by Middle-Up Level HILIC-HRMS Analysis. *Pharmaceutics* **2021,** *13* (11).
292. Chau, C. H.; Steeg, P. S.; Figg, W. D., Antibody-drug conjugates for cancer. *Lancet* **2019,** *394* (10200), 793-804.
293. Rocha, R. O.; Rodrigues, M. O.; Neto, B. A. D., Review on the Ugi Multicomponent Reaction Mechanism and the Use of Fluorescent Derivatives as Functional Chromophores. *ACS Omega* **2020,** *5* (2), 972-979.
294. Medeiros, G. A.; da Silva, W. A.; Bataglion, G. A.; Ferreira, D. A.; de Oliveira, H. C.; Eberlin, M. N.; Neto, B. A., Probing the mechanism of the Ugi four-component reaction with charge-tagged reagents by ESI-MS(/MS). *Chem Commun (Camb)* **2014,** *50* (3), 338-40.

295. Rothe, C.; Skerra, A., Anticalin((R)) Proteins as Therapeutic Agents in Human Diseases. *BioDrugs* **2018,** *32* (3), 233-243.
296. Gebauer, M.; Skerra, A., Anticalins small engineered binding proteins based on the lipocalin scaffold. *Methods Enzymol* **2012,** *503*, 157-88.
297. Deuschle, F. C.; Schiefner, A.; Brandt, C.; Skerra, A., Design of a surrogate Anticalin protein directed against CD98hc for preclinical studies in mice. *Protein Sci* **2020,** *29* (8), 1774-1783.
298. Deuschle, F. C.; Morath, V.; Schiefner, A.; Brandt, C.; Ballke, S.; Reder, S.; Steiger, K.; Schwaiger, M.; Weber, W.; Skerra, A., Development of a high affinity Anticalin((R)) directed against human CD98hc for theranostic applications. *Theranostics* **2020,** *10* (5), 2172-2187.
299. Shim, H., Bispecific Antibodies and Antibody-Drug Conjugates for Cancer Therapy: Technological Considerations. *Biomolecules* **2020,** *10* (3).
300. Wu, D.; Piszczek, G., Standard protocol for mass photometry experiments. *Eur Biophys J* **2021,** *50* (3-4), 403-409.
301. Nunes, J. P.; Morais, M.; Vassileva, V.; Robinson, E.; Rajkumar, V. S.; Smith, M. E.; Pedley, R. B.; Caddick, S.; Baker, J. R.; Chudasama, V., Functional native disulfide bridging enables delivery of a potent, stable and targeted antibody-drug conjugate (ADC). *Chem Commun (Camb)* **2015,** *51* (53), 10624-7.
302. Le Du, E.; Duhail, T.; Wodrich, M. D.; Scopelliti, R.; Fadaei-Tirani, F.; Anselmi, E.; Magnier, E.; Waser, J., Structure and Reactivity of N-Heterocyclic Alkynyl Hypervalent Iodine Reagents. *Chemistry* **2021,** *27* (42), 10979-10986.
303. Lazar, A. C.; Wang, L.; Blattler, W. A.; Amphlett, G.; Lambert, J. M.; Zhang, W., Analysis of the composition of immunoconjugates using size-exclusion chromatography coupled to mass spectrometry. *Rapid Commun Mass Spectrom* **2005,** *19* (13), 1806-14.
304. Brady, L. J.; Valliere-Douglass, J.; Martinez, T.; Balland, A., Molecular mass analysis of antibodies by on-line SEC-MS. *J Am Soc Mass Spectrom* **2008,** *19* (4), 502-9.
305. Liu, H.; Gaza-Bulseco, G.; Chumsae, C., Analysis of reduced monoclonal antibodies using size exclusion chromatography coupled with mass spectrometry. *J Am Soc Mass Spectrom* **2009,** *20* (12), 2258-64.
306. Grunert, I.; Heinrich, K.; Hingar, M.; Ernst, J.; Winter, M.; Bomans, K.; Wagner, K.; Fevre, A.; Reusch, D.; Wuhrer, M.; Bulau, P., Comprehensive Multidimensional Liquid Chromatography–Mass Spectrometry for the Characterization of Charge Variants of a Bispecific Antibody. *Journal of the American Society for Mass Spectrometry* **2022,** *33* (12), 2319-2327.
307. Janin-Bussat, M. C.; Dillenbourg, M.; Corvaia, N.; Beck, A.; Klinguer-Hamour, C., Characterization of antibody drug conjugate positional isomers at cysteine residues by peptide mapping LC-MS analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* **2015,** *981-982*, 9-13.
308. Armirotti, A.; Benatti, U.; Damonte, G., Top-down proteomics with a quadrupole time-of-flight mass spectrometer and collision-induced dissociation. *Rapid Commun Mass Spectrom* **2009,** *23* (5), 661-6.
309. Diedrich, J. K.; Pinto, A. F.; Yates, J. R., 3rd, Energy dependence of HCD on peptide fragmentation: stepped collisional energy finds the sweet spot. *J Am Soc Mass Spectrom* **2013,** *24* (11), 1690-9.
310. Jesper V Olsen, B. M., Oliver Lange, Alexander Makarov, Stevan Horning & Matthias Mann, Higher-energy C-trap dissociation (HCD) for precise peptide
modification analysis. **2007**.
311. Ganisl, B.; Breuker, K., Does Electron Capture Dissociation Cleave Protein Disulfide Bonds? *ChemistryOpen* **2012,** *1* (6), 260-8.
312. Fornelli, L.; Srzentic, K.; Toby, T. K.; Doubleday, P. F.; Huguet, R.; Mullen, C.; Melani, R. D.; Dos Santos Seckler, H.; DeHart, C. J.; Weisbrod, C. R.; Durbin, K. R.; Greer, J. B.; Early, B. P.; Fellers, R. T.; Zabrouskov, V.; Thomas, P. M.; Compton, P. D.; Kelleher, N. L., Thorough Performance Evaluation of 213 nm Ultraviolet Photodissociation for Top-down Proteomics. *Mol Cell Proteomics* **2020,** *19* (2), 405-420.

313. Bayat, P.; Lesage, D.; Cole, R. B., Tutorial: Ion Activation in Tandem Mass Spectrometry Using Ultra-High Resolution Instrumentation. *Mass Spectrom Rev* **2020,** *39* (5-6), 680-702.
314. Helms, A.; Escobar, E. E.; Vainauskas, S.; Taron, C. H.; Brodbelt, J. S., Ultraviolet Photodissociation Permits Comprehensive Characterization of O-Glycopeptides Cleaved with O-Glycoprotease IMPa. *Anal Chem* **2023,** *95* (24), 9280-9287.
315. Blevins, M. S.; Juetten, K. J.; James, V. K.; Butalewicz, J. P.; Escobar, E. E.; Lanzillotti, M. B.; Sanders, J. D.; Fort, K. L.; Brodbelt, J. S., Nanohydrophobic Interaction Chromatography Coupled to Ultraviolet Photodissociation Mass Spectrometry for the Analysis of Intact Proteins in Low Charge States. *J Proteome Res* **2022,** *21* (10), 2493-2503.
316. Moelleken, J.; Endesfelder, M.; Gassner, C.; Lingke, S.; Tomaschek, S.; Tyshchuk, O.; Lorenz, S.; Reiff, U.; Molhoj, M., GingisKHAN protease cleavage allows a high-throughput antibody to Fab conversion enabling direct functional assessment during lead identification of human monoclonal and bispecific IgG1 antibodies. *MAbs* **2017,** *9* (7), 1076-1087.
317. von Pawel-Rammingen, U.; Johansson, B. P.; Bjorck, L., IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *EMBO J* **2002,** *21* (7), 1607-15.
318. Johansson, B. P.; Shannon, O.; Bjorck, L., IdeS: A Bacterial Proteolytic Enzyme with Therapeutic Potential. *PLoS ONE* **2008,** *3* (2).
319. Jin, Y.; Lin, Z.; Xu, Q.; Fu, C.; Zhang, Z.; Zhang, Q.; Pritts, W. A.; Ge, Y., Comprehensive characterization of monoclonal antibody by Fourier transform ion cyclotron resonance mass spectrometry. *MAbs* **2019,** *11* (1), 106-115.
320. Bayat, P.; Lesage, D.; Cole, R. B., Low-energy collision-induced dissociation (low-energy CID), collision-induced dissociation (CID), and higher energy collision dissociation (HCD) mass spectrometry for structural elucidation of saccharides and clarification of their dissolution mechanism in DMAc/LiCl. *J Mass Spectrom* **2018,** *53* (8), 705-716.
321. Shrivastava, A.; Joshi, S.; Guttman, A.; Rathore, A. S., N-Glycosylation of monoclonal antibody therapeutics: A comprehensive review on significance and characterization. *Anal Chim Acta* **2022,** *1209*, 339828.
322. Jhingree, J. R.; Beveridge, R.; Dickinson, E. R.; Williams, J. P.; Brown, J. M.; Bellina, B.; Barran, P. E., Electron transfer with no dissociation ion mobility–mass spectrometry (ETnoD IM-MS). The effect of charge reduction on protein conformation. *International Journal of Mass Spectrometry* **2017,** *413*, 43-51.
323. Chen, Q.; Dai, R.; Yao, X.; Chaihu, L.; Tong, W.; Huang, Y.; Wang, G., Improving Accuracy in Mass Spectrometry-Based Mass Determination of Intact Heterogeneous Protein Utilizing the Universal Benefits of Charge Reduction and Alternative Gas-Phase Reactions. *Anal Chem* **2022,** *94* (40), 13869-13878.
324. Dunham, S. D.; Wei, B.; Lantz, C.; Loo, J. A.; Brodbelt, J. S., Impact of Internal Fragments on Top-Down Analysis of Intact Proteins by 193 nm UVPD. *J Proteome Res* **2023,** *22* (1), 170-181.
325. Zenaidee, M. A.; Lantz, C.; Perkins, T.; Jung, W.; Loo, R. R. O.; Loo, J. A., Internal Fragments Generated by Electron Ionization Dissociation Enhance Protein Top-Down Mass Spectrometry. *J Am Soc Mass Spectrom* **2020,** *31* (9), 1896-1902.
326. Pronk, S. D.; Schooten, E.; Heinen, J.; Helfrich, E.; Oliveira, S.; van Bergen En Henegouwen, P. M. P., Single Domain Antibodies as Carriers for Intracellular Drug Delivery: A Proof of Principle Study. *Biomolecules* **2021,** *11* (7), 927.
327. de Beer, M. A.; Giepmans, B. N. G., Nanobody-Based Probes for Subcellular Protein Identification and Visualization. *Front Cell Neurosci* **2020,** *14*, 573278.
328. Debie, P.; Devoogdt, N.; Hernot, S., Targeted Nanobody-Based Molecular Tracers for Nuclear Imaging and Image-Guided Surgery. *Antibodies (Basel)* **2019,** *8* (1).
329. Teodori, L.; Omer, M.; Marcher, A.; Skaanning, M. K.; Andersen, V. L.; Nielsen, J. S.; Oldenburg, E.; Lin, Y.; Gothelf, K. V.; Kjems, J., Site-specific nanobody-oligonucleotide conjugation for super-resolution imaging. *J Biol Methods* **2022,** *9* (1), e159.
330. Sharifi, J.; Khirehgesh, M. R.; Safari, F.; Akbari, B., EGFR and anti-EGFR nanobodies: review and update. *J Drug Target* **2021,** *29* (4), 387-402.

331. Thom, J.; Anderson, D.; McGregor, J.; Cotton, G., Recombinant protein hydrazides: application to site-specific protein PEGylation. *Bioconjug Chem* **2011,** *22* (6), 1017-20.
332. Muir, T. W.; Sondhi, D.; Cole, P. A., Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **1998,** *95*, 705–6710.
333. Muralidharan, V.; Muir, T. W., Protein ligation: an enabling technology for the biophysical analysis of proteins. *NATURE METHODS* **2006,** *3* (6).
334. Brindha, S.; Kibria, M. G.; Saotome, T.; Unzai, S.; Kuroda, Y., EGFR extracellular domain III expressed in Escherichia coli with SEP tag shows improved biophysical and functional properties and generate anti-sera inhibiting cancer cell growth. *Biochem Biophys Res Commun* **2021,** *555*, 121-127.
335. Gibson, T. J.; Smyth, P.; McDaid, W. J.; Lavery, D.; Thom, J.; Cotton, G.; Scott, C. J.; Themistou, E., Single-Domain Antibody-Functionalized pH-Responsive Amphiphilic Block Copolymer Nanoparticles for Epidermal Growth Factor Receptor Targeted Cancer Therapy. *ACS Macro Lett* **2018,** *7* (8), 1010-1015.
336. Sigismund, S.; Avanzato, D.; Lanzetti, L., Emerging functions of the EGFR in cancer. *Mol Oncol* **2018,** *12* (1), 3-20.
337. Rosano, G. L.; Ceccarelli, E. A., Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol* **2014,** *5*, 172.
338. Mitchell, S. F.; Lorsch, J. R., Protein Affinity Purification using Intein/Chitin Binding Protein Tags. *Methods Enzymol* **2015,** *559*, 111-25.
339. Ta, D. T.; Redeker, E. S.; Billen, B.; Reekmans, G.; Sikulu, J.; Noben, J. P.; Guedens, W.; Adriaensens, P., An efficient protocol towards site-specifically clickable nanobodies in high yield: cytoplasmic expression in Escherichia coli combined with intein-mediated protein ligation. *Protein Eng Des Sel* **2015,** *28* (10), 351-63.
340. Aranko, A. S.; Iwai, H., The Inducible Intein-Mediated Self-Cleaving Tag (IIST) System: A Novel Purification and Amidation System for Peptides and Proteins. *Molecules* **2021,** *26* (19).
341. Nguyen, K. T.; Kim, J. M.; Park, S. E.; Hwang, C. S., N-terminal methionine excision of proteins creates tertiary destabilizing N-degrons of the Arg/N-end rule pathway. *J Biol Chem* **2019,** *294* (12), 4464-4476.
342. Macias, L. A.; Santos, I. C.; Brodbelt, J. S., Ion Activation Methods for Peptides and Proteins. *Anal Chem* **2020,** *92* (1), 227-251.
343. Mentinova, M.; Han, H.; McLuckey, S. A., Dissociation of disulfide-intact somatostatin ions: the roles of ion type and dissociation method. *Rapid Commun Mass Spectrom* **2009,** *23* (17), 2647-55.
344. Vandenberghe, L. H.; Wilson, J. M.; Gao, G., Tailoring the AAV vector capsid for gene therapy. *Gene Ther* **2009,** *16* (3), 311-9.
345. Cardoso, M. M.; Peça, I. N.; Roque, A. C. A., Antibody-Conjugated Nanoparticles for Therapeutic Applications. *Current Medicinal Chemistry* **2012**.
346. Strasser, L.; Morgan, T. E.; Guapo, F.; Fussl, F.; Forsey, D.; Anderson, I.; Bones, J., A Native Mass Spectrometry-Based Assay for Rapid Assessment of the Empty:Full Capsid Ratio in Adeno-Associated Virus Gene Therapy Products. *Anal Chem* **2021,** *93* (38), 12817-12821.
347. Jarrold, M. F., Applications of Charge Detection Mass Spectrometry in Molecular Biology and Biotechnology. *Chem Rev* **2022,** *122* (8), 7415-7441.
348. Strasser, L.; Fussl, F.; Morgan, T. E.; Carillo, S.; Bones, J., Exploring Charge-Detection Mass Spectrometry on Chromatographic Time Scales. *Anal Chem* **2023,** *95* (40), 15118-15124.
349. Kostelic, M. M.; Zak, C. K.; Liu, Y.; Chen, V. S.; Wu, Z.; Sivinski, J.; Chapman, E.; Marty, M. T., UniDecCD: Deconvolution of Charge Detection-Mass Spectrometry Data. *Anal Chem* **2021,** *93* (44), 14722-14729.
350. Papanastasiou, D.; Kounadis, D.; Lekkas, A.; Orfanopoulos, I.; Mpozatzidis, A.; Smyrnakis, A.; Panagiotopoulos, E.; Kosmopoulou, M.; Reinhardt-Szyba, M.; Fort, K.; Makarov, A.; Zubarev, R. A., The Omnitrap Platform: A Versatile Segmented Linear Ion Trap for Multidimensional Multiple-Stage Tandem Mass Spectrometry. *J Am Soc Mass Spectrom* **2022,** *33* (10), 1990-2007.
351. Smyrnakis, A.; Levin, N.; Kosmopoulou, M.; Jha, A.; Fort, K.; Makarov, A.; Papanastasiou, D.; Mohammed, S., Characterization of an Omnitrap-Orbitrap Platform Equipped with Infrared

Multiphoton Dissociation, Ultraviolet Photodissociation, and Electron Capture Dissociation for the Analysis of Peptides and Proteins. *Anal Chem* **2023,** *95* (32), 12039-12046.
352. T., K.; M., H.; D., P.; A., L.; D., K.; I., O., A Novel Instrument Platform for the Investigation of Particle Formation from the Gas Phase. 67th Conference on Mass Spectrometry and Allied Topics, WP466, Atlanta, GA. **2019**.
353. Waters, Enhanced biopharmaceutical characterization using the SELECT SERIES™ MRT. **2022**.
354. Heil, L. R.; Damoc, N. E.; Arrey, T. N.; Pashkova, A.; Denisov, E.; Petzoldt, J.; Peterson, A.; Hsu, C.; Searle, B. C.; Shulman, N.; Riffle, M.; Connolly, B.; MacLean, B. X.; Remes, P. M.; Senko, M.; Stewart, H.; Hock, C.; Makarov, A.; Hermanson, D.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J., Evaluating the performance of the Astral mass analyzer for quantitative proteomics using data independent acquisition. **2023**.

Université de Strasbourg

Rania BENAZZA

Nouvelles approches de spectrométrie de masse pour la caractérisation de protéines thérapeutiques bioconjuguées

École doctorale
Sciences chimiques | ED 222
Université de Strasbourg

Résumé

Ce travail de thèse se concentre des développements méthodologiques en spectrométrie de masse (MS) pour la caractérisation des protéines bioconjuguées, notamment les anticorps monoclonaux (mAb), ainsi que d'autre protéines thérapeutiques. Le premier axe majeur de cette thèse consiste à l'amélioration des techniques de couplage de chromatographie liquide (LC) à la MS dans des conditions natives (nMS) ainsi que dénaturantes afin de rendre ce couplage plus versatile et plus adaptée à l'utilisation en industrie pharmaceutique. Deuxièmement, le développement de nouvelles approches dites Top- ou middle-down MS (TD/MD-MS) a été étendu sur une large variété de biothérapeutiques, allant des protéines et de mAb de références jusqu'au protéines conjuguées. Les développements au niveau intact ainsi qu'au niveau de la structure primaire ont été réalisé sur des spectromètres de masse de pointe, ce qui a permis d'évalué ces instruments. De plus, cette thèse a permis de mettre en évidence la complémentarité des techniques nMS et TD/MD-MS offrant une meilleure compréhension des structures et des modifications post-traductionnelles des protéines bioconjuguées.

Mots-clés : spectrométrie de masse native, top-down, middle-down, anticorps monoclonaux, protéines bioconjuguées

Résumé en anglais

This PhD work focuses on methodological developments in mass spectrometry (MS) for the characterization of bioconjugated proteins, in particular monoclonal antibodies (mAb), as well as other therapeutic proteins. The first major thrust of this thesis is the improvement of the liquid chromatography (LC) coupling to MS under native (nMS) and denaturing conditions, in order to make this coupling more versatile and more suitable for use in the pharmaceutical industry. Secondly, the development of new Top- or middle-down MS (TD/MD-MS) approaches has been extended to a wide variety of biotherapeutics, from reference proteins and mAbs to conjugated proteins. Developments at both the intact and primary structure levels were carried out on state-of-the-art mass spectrometers, enabling the evaluation of these instruments. In addition, this thesis highlighted the complementary of nMS and TD/MD-MS techniques, offering a better understanding of the structures and post-translational modifications of bioconjugated proteins.

Keywords: native mass spectrometry, top-down, middle-down, monoclonal antibodies, bioconjugated proteins