

表3 バイオ医薬品の特徴

- 構造的に複雑
- 不均一性(heterogeneity)や高次構造
- ドメイン構造とそのドメイン構造に特有の生物機能
- 品質特性が非常に複雑な構成要素からなる
- 最終製品で品質特性をとらえることが困難
- 製造工程により品質の恒常性を担保する必要
- 免疫学的特性=抗原性

バイオ医薬品はさまざまなドメイン構造を持ち、かつそのドメイン構造に対応する生物活性等の機能を有している。したがって、1つの生物活性だけでバイオ医薬品の機能をとらえることが難しい場合も多く、生物活性をさまざまな角度から評価することがバイオ後続品開発において必要であることが多いと考えられる。

バイオ医薬品の品質特性に関しても、化学薬品に比べて非常に複雑な要素から構成されている。例えば、目的とする有効成分(desired product)そのものが糖タンパク質医薬品のように不均一な集団から構成されている場合がある。また、デスアミド体や酸化体などの目的物質関連物質(Product-related substances)と規定できる場合もあり、一方で活性や有効性の観点から目的物質関連物質と考えることのできない分解物などの目的物質由来不純物(Product-related impurities)や工程由来不純物が含まれ、これらを含めた品質特性を明らかにすることが求められる。

一般的に、バイオ医薬品の持つ不均一性や構造の複雑さゆえに、最終製品の特性解析のみによってその品質特性やその恒常性を立証することが困難であるという特徴を持ち、バイオ医薬品の開発においては、恒常性と頑健性のある製法を確立しておくことが製品の品質恒常性確保に重要と考えられている。バイオ後続品の開発においても、1次構造や物理的特性などについて、目的物質との同一性の立証が可能な場合であっても、多様な要素からなるバイオ後続品の品質特性全体について参照する先行バイオ医薬品との「同一性」を示すことは非常に困難である。さらには、品質特性に関する比較では、類似性、さらに言えば高い類似性を示すことが現実的であろうと考えられる。したがって、バイオ後続品開発においては、このような複雑な品質特性をどのように評価し、かつ目的とする先行バイオ医薬品との類似性(similarity)を明らかにしていくかが重要とされた。

(2) 抗原性

抗原性に関連して、EUで承認されたバイオ後続品で

あるオムニトロープの臨床開発初期段階での抗hGH抗体が、高頻度で発現されたことが注目されている。また、エリスロポエチンの製法変更に伴う抗体産生による有害事象発症などもあり、抗原性がバイオ後続品開発での大きな懸念である。オムニトロープの開発では、製法の見直しにより抗原性は十分にコントロールできるようになったことが審査報告書で示されており、混入していた宿主タンパク質によるアジュバント効果が開発初期での抗体産生につながったと推定されている。したがって、バイオ後続品開発では、抗原性についての懸念を常に念頭に置いておくことが求められる。

バイオ医薬品の抗原性は、目的物質や重合体等の目的物質由来物質による場合と、不純物によるアジュバント効果などさまざまな要因によって引き起こされる可能性がある。エリスロポエチンの事例のように、容器等の基材が抗原性に影響を与えることも念頭に置く必要がある。さらには、生産細胞基材によってはGal α 1-3Galのように異種糖鎖抗原が製品に含まれる可能性や工程由来不純物に対する抗原性も考慮する必要がある。しかしヒトでの抗原性を、動物を用いた非臨床試験で予測することは困難であり、抗原性の有無については臨床研究で評価せざるを得ない場合も多いと考えられる。しかし、従来の経験から製造細胞由来タンパク質の混入が多い場合や重合体比率などは抗原性の観点から事前に評価を行う必要があると思われる。

(3) 製造用細胞基材や製造工程

バイオ後続品開発において、先行バイオ医薬品に用いられている生産基材である宿主細胞や遺伝子発現構成体、組換え体の選択方法などの情報は公開されていないと想定される。CHO細胞やSP1細胞などを用いていることが明らかになっている場合もあるが、サブストレインまでの情報が公開されることはないと思われる。

一般に、生産基材である細胞が異なれば、宿主由来不純物のプロファイルや翻訳後修飾に伴う不均一性などが異なってくる可能性が高いと想定される。また、それぞれ用いる細胞の特性に応じて培養液や血清等の要求性が異なり、このような培養条件の差異が製品における不純物の差異につながる可能性もある。一般に、糖タンパク質バイオ医薬品の開発においても、宿主細胞の選択によって糖鎖等の不均一性が大きく異なることがよく知られており、バイオ後続品の開発においても同様の差異が生じることが想定される。

さらに宿主細胞選択や培養条件等の製造方法ばかりでなく、精製工程が製品の品質特性に大きく影響することが知られている。例えば、カラムクロマトグラフィー工程で分取方法によっては、特定の不均一性を持つ製品を製造することも可能である。したがって、先行バイオ医薬品と高い類似性を持つバイオ後続品を開発するために、先行バイオ医薬品の採用している生産細胞基材、培養方法、精製工程を通じての違いによる不均一性もバイオ後続品開発の大きな課題と考えられる。

(4) 先行バイオ医薬品が持つ複数の効能

バイオ後続品の開発において、先行バイオ医薬品が複数の効能を持つ場合、それぞれの効能について承認を得るにはどのようなデータが求められるのかが大きな論点の1つとなった。先行バイオ医薬品の開発では、それぞれの効能ごとに臨床試験を行い、効能ごとの承認を得ている。しかし、バイオ後続品の開発においても、先行バイオ医薬品と同様にそれぞれの効能ごとに臨床試験を行うべきか、科学的な合理性があれば効能ごとに臨床試験を行わなくてもそれぞれの効能に対する承認が得られるかということが議論された。わが国では、5つのヒト成長ホルモン製品(ソマトロピン)が承認されているが、それぞれ複数の効能を持っており、また、その効能は製品ごとに異なっている。バイオ後続品の位置づけは、先行バイオ医薬品と高い類似性を持つことが前提であり、高い類似性があれば先行バイオ医薬品と同じ効能を持つことが期待されるという側面がある。これは複数の承認品目がある製品の取り扱いを含め十分な議論が必要とされた点である。また対照とする先行バイオ医薬品にない効能に関して、他の同種同効医薬品において認められている効能が取得できるかという点も大きな論点であった。

(5) 先行バイオ医薬品との類似性評価一同等性・同質性評価

前述したようにバイオ後続品の開発において、その品質特性が先行バイオ医薬品と高い類似性を持つことを前提として、安全性や有効性に関する試験の一部を省略できるのではないかとされている。EMAのバイオシミラーガイドラインにおいても、この先行バイオ医薬品との同等性・同質性を評価する手法としては、ICH Q5Eの製法変更における同等性・同質性のアプローチが適用できるとしている。

しかし、ICH Q5Eガイドラインは自社において製法

を変更した場合に、旧製法の製品と新製法の製品を比較するための技術的要件について書かれたガイドラインである。すなわち、ICH Q5Eガイドラインでは、比較する旧製法の製品のすべての情報にアクセス可能であり、比較試験においても製剤・原薬ばかりでなく中間工程製品についても直接比較することが可能であることを前提に書かれている。したがって製法変更の場合には、現在の最先端の科学技術を駆使すれば、品質特性について両製品の類似性を評価することはそれほど困難でないと思定される。

一方、バイオ後続品の開発では、ほとんどの先行バイオ医薬品の情報は非公開であると想定されることより、これらの情報を得ることなく先行バイオ医薬品との類似性を評価していくことが求められる。さらに、比較に用いる先行バイオ医薬品の原薬の入手は困難である可能性が高い。したがって自社製品の製法変更時の試験が必ずしも適用できるというわけではない。また、製剤での比較試験を実施する際には、製剤に添加されている添加剤を除去する必要がある場合も考えられ、そのような除去工程による操作が、目的とする先行バイオ医薬品の本来の品質特性に影響を与える可能性も考慮しなければならない。

ただし、現在の科学の進歩はめざましく、さまざまな手法を駆使し、創意工夫を凝らすことにより製剤を用いても物理化学的性質や生物活性などの解析が可能になりつつある。また、新たな技術開発により、従来の手法とはまったく異なるアプローチから先行バイオ医薬品との類似性を評価することも可能になってくることもありえる。

さらに、バイオ後続品開発を開発する企業が持つさまざまな経験をどこまで活用するかなども含め、さまざまな制限のある中でいかにして、品質・安全性・有効性において対照とする既承認バイオ医薬品との同等性・同質性を明らかにできるかが、バイオ後続品の評価において重要であると考えられる。

2. バイオ後続品指針について

バイオ後続品指針案作成のための専門家会議では前述の論点について議論を重ね、指針に盛り込むべき観点を整理した。まずバイオ後続品をどのような医薬品として規制するかについて議論を行った。

バイオ後続品の品質・安全性・有効性確保のための指針では、「国内で既に新有効成分含有医薬品として承認されたバイオテクノロジー応用医薬品(以下「先行バイオ

医薬品」という。)と同等/同質の品質、安全性、有効性を有する医薬品として、異なる製造販売業者により開発される医薬品である。」と定義し、その開発は「安全性及び有効性について、先行バイオ医薬品との比較から得られた同等性/同質性を示すデータ等に基づき開発できる。」とした。すなわち、バイオ後続品は国内ですでに承認を受けている先行バイオ医薬品があり、独自に実施する試験に加えて、先行バイオ医薬品を参照しながら、品質のみならず非臨床試験や臨床試験を含めて合理的な比較試験を実施し、先行バイオ医薬品との同等性・同質性を示すことにより開発が可能であると結論した。

バイオ後続品に関する指針としては、現時点でわが国でのバイオ後続品の審査に関する経験がほとんどないことから、品質から非臨床開発、臨床開発、市販後調査を含めた包括的な指針とすることとした(表4)。より詳細な指針や個々の製品ごとの評価指針の作成の可能性については将来的な課題とした。また、バイオ後続品指針の対象としては、理論的にはすべてのバイオ製品について後続品の開発が可能と考えられるが、目的有効成分を明確に規定することが比較的容易で、かつ高度に精製され十分な品質特性解析が容易な組換えタンパク質を取り上げることにした。

また、先行バイオ医薬品が国内で承認されていない場

表4 バイオ後続品の品質・安全性・有効性確保のための指針の構成

1. 始めに
2. 適用範囲(対象)
3. バイオ後続品開発における一般原則
4. バイオ後続品の製法・品質特性解析
4.1. 製法開発
宿主・ベクター系
セルバンクシステム
培養・精製工程
4.2. 特性解析(構造解析, 物理的・化学的性質, 生物活性等)
4.3. 製剤設計
4.4. 安定性試験
5. 品質特性に関する同等性/同質性の評価試験
①構造解析, 物理的・化学的性質に関する比較試験
②生物活性に関する比較試験
③免疫原性等に関する比較試験
6. 規格及び試験方法
7. 非臨床試験
7.1. 毒性試験
7.2. 薬理試験
8. 臨床試験
8.1. 臨床薬物動態(PK)試験, 薬力学(PD)試験及びPK/PD試験
8.2. 臨床的有効性の比較
8.3. 臨床的安全性の確認
9. 製造販売後調査
用語集・定義

合にはバイオ後続品の開発は認められないとした。開発途上国では国内市場の大きさや経済的な理由から先行バイオ医薬品の開発が行われていない製品も多く、自国内で承認されていないバイオ医薬品に対するバイオ後続品が承認されている製品もあることが報告されている。指針の開始に当たってもわが国で承認されていないバイオ医薬品に対するバイオ後続品が可能かについて議論を行った。その結果、わが国で治験が行われておらず、バイオ医薬品に対する審査の経験もない製品に対するバイオ後続品を十分評価可能であるか不明なことなどから現時点では認められないとした。

バイオ後続品の開発に当たって考慮すべきこととして、バイオ後続品は先行バイオ医薬品の開発から10~15年の科学の進歩があることから、その進歩を取り入れた安全性対策を模索し、また最新の評価手法を適用することが求められると考えられる。

図1に、バイオ後続品の承認申請に当たって必要なデータについての概念をまとめてみた。すなわち先行医薬品のすべてのデータは公開されておらず、得られる情報は限られていることを前提とすることから、通常の新規バイオ医薬品と同様に独自に恒常性と頑健性のある製法を確立することが求められる。さらに、確立した製法で製造されたバイオ後続品について新薬と同様に品質特性解析データを取得し、これらのデータを提出することが求められる。その上で、先行バイオ医薬品との品質特性の高い類似性を示し、さらに合理的かつ必要と考えられる非臨床試験、臨床試験をデザインし、これらの試験を総合して先行バイオ医薬品との同等性・同質性を示すことによりバイオ後続品としての開発が可能と考えられる。したがって、非臨床試験や臨床試験については、一部のデータの取得は不要であるとしても、後発品と異なり承認申請資料としてはCTDの第3部までのデータが求められる。

理論的には、すべてのバイオ後続品が開発可能と考えられるが、バイオ医薬品には組換えタンパク質性医薬品から生体由来製品やワクチンなど多様な製品が含まれ、すべての製品に対する指針を作成することは現実的ではないと考えられた。したがって、バイオ後続品指針の適用範囲としては高度に精製され、十分な特性解析が可能な組換えDNAタンパク質性医薬品/ポリペプチドとした。ただ、ICH Q6Bで書かれているように、細胞培養由来/組織・体液から分離されるタンパク質性医薬品で、高度に精製され、特性解析可能な医薬品にも本指針で示

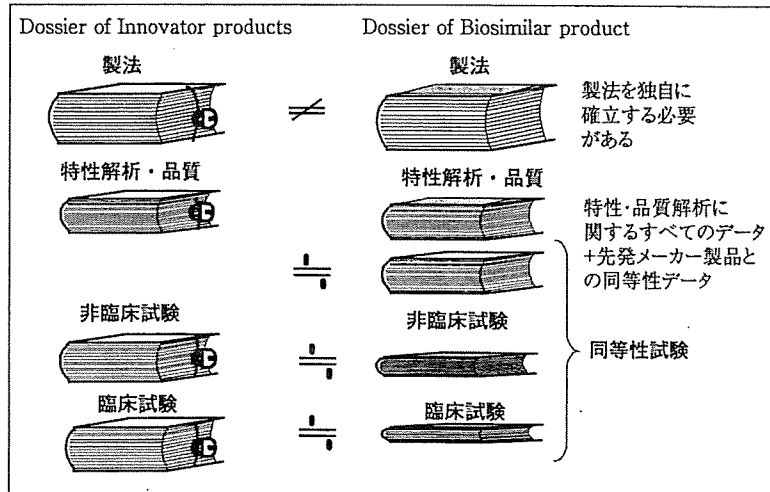


図1 バイオ後続品に求められるデータ

す基本原則が適用可能な場合があるとしている。

上述したようにバイオ後続品開発においては、新規遺伝子組換えタンパク質性医薬品と同様に、独自に確立した頑健性と恒常性のある製法により製造したバイオ後続品候補品について、新規バイオ医薬品と同様に徹底した品質特性解析を実施し、品質特性を明らかにすることが求められる。その上で、対照とする先行バイオ医薬品と品質特性における類似性を示す必要がある。

製法と品質特性に関しては、まず新薬同様に開発を行っていくことが必要となるが、先行バイオ医薬品との品質特性の比較試験の結果、品質特性に差異が認められた場合には2つの選択肢が想定される。

1つには、明らかに不均一性等に差異がある場合に、それが培養工程や精製工程等を工夫することにより品質特性がより類似した製品の開発が可能と予測される場合には、製造工程を再デザインし、目的とする先行バイオ医薬品と類似性の高い品質特性を持つバイオ後続品を得られるように製法を最適化していくことである。

一方、先行バイオ医薬品との品質特性の差異が大きい場合や、品質特性の高い類似性を示すためにはデータが不足しているものの有効性や安全性に影響を与えないと推定される場合には、非臨床試験や臨床試験での先行バイオ医薬品との比較試験を実施し、同等性・同質性を実証していくことも考えられる。品質特性での類似性をどの程度立証できたかにより、必要となる非臨床試験や臨床試験のデータが異なってくると想定される。

対照とする先行バイオ医薬品とバイオ後続品との比較試験の意義や実証方法については次のように考えることができるであろう。先行バイオ医薬品との類似性を実証

するということは、先行バイオ医薬品に対して、バイオ後続品の品質特性がまったく同一であることを意味するのではなく、品質特性において類似性が高く、かつ、品質特性に何らかの差異があったとしても、最終製品の安全性や有効性に有害な影響を及ぼさないと科学的に判断できることを意味する。

しかし、バイオ後続品の開発では、ほとんどの先行バイオ医薬品の承認申請に関連するデータがブラックボックスの中にあることから、手法的にはさまざまな工夫が必要となるであろう。また、試験に用いることのできる先行バイオ医薬品の入手の可能性や試験法の限界などにより、実施が困難な試験も想定される。したがって、すべての特性を比較することを求めているのではなく、合理的にアプローチ可能な範囲での比較試験を実施することが求められている。

(1) 製造方法

バイオ後続品の製法開発では、対照とする先行バイオ医薬品の製法に関しては限定的な情報しか得られないと考えられることを前提とせざるを得ず、宿主・ベクター系、セルバンクシステム、培養・精製工程を同一とすることは不要とされている。しかし、対照とした先行バイオ医薬品で用いられている生産細胞基材が明らかになっている場合には、同一の細胞基材を用いることが望ましいと考えられる。この場合でも、CHO細胞やBHK細胞を用いているといった情報が得られても、そのサブタイプまでは不明の場合が多いと想定される。したがって、細胞のサブタイプまでの一致を望ましいとしているわけではない。一方、異なる生産細胞基材を用いることを否

定しているわけではないが、異なる生産細胞を用いる場合には、宿主由来タンパク質や目的とする細胞基材に用いられる培養液や添加剤などに由来する工程由来不純物が異なってくると想定され、このような工程由来不純物が異なっていることを前提に不純物の安全性評価を行うことが求められる。

樹立した細胞の特性解析、ウイルス安全性評価、遺伝的安定性評価については、ICH Q5D, Q5A, Q5Bガイドラインに従って評価を行うことが求められる。

一方で、安全性確保の観点から有効性に影響を与えない範囲でより安全な製法を模索することはむしろ推奨される場合もある。特に、先行バイオ医薬品や関連するバイオ医薬品に関するさまざまな安全情報に基づき、適切な製法をデザインしていくことは有用と考えられ、例えばウシ血清を用いた培地よりも無血清培地の開発を行うことなどがあげられる。しかし、目的とする有効成分が糖タンパク質であった場合に、血清の有無はその糖鎖の不均一性に大きな影響を与える可能性もあり、ひいては糖鎖の違いによる生物活性や体内動態の大きな差異になる可能性もあるため、品質に対する影響も十分考慮する必要がある。

(2) 品質特性解析

特性解析では、新規組換えタンパク質医薬品と同様に最新の科学技術を用いて、①構造・組成、②物理的・化学的性質、③生物活性、④免疫化学的性質、⑤不純物等について十分なデータを得る必要がある。この特性解析の結果に基づき、ICH Q6Bに準じて規格及び試験方法の設定を行うことが求められる。

製剤設計に当たっては、原則的に対照バイオ医薬品と剤形や投与経路は同一であることが必要となるが、臨床での利便性から対照バイオ医薬品と異なる剤形を選択することが可能な場合も想定される。例えば、先行バイオ医薬品が凍結乾燥製剤であったとしても、品質特性や安定性に影響を与えないことが実証された液剤での開発はむしろ臨床現場での利便性が増す可能性もある。ただ、剤形によっては貯法が大きく異なる場合も想定され、臨床で混乱が起きる可能性も考えられるので、総合的に判断する必要もある。

また、保存条件および有効期間が対照バイオ医薬品と同一であることは必ずしも必須条件ではないが、極端に有効期間が短い場合には臨床での誤用も考えられることから、妥当とされない場合もありえる。対照バイオ医薬

品と異なる添加剤を用いることが可能な場合もある。特に安全性の観点から、異なる添加剤の選択が有用な場合もありえる。

安定性試験では、ICH Q5Cに準じて実保存条件・実時間での長期保存試験の実施が求められており、この試験結果に基づき有効期間を設定することが必要となる。ただし、承認申請時には少なくとも6か月以上の試験データを提出することでよいとされている。

保存条件および有効期間が対照バイオ医薬品と同一であることは必須条件でないことから、安定性について対照バイオ医薬品と必ずしも比較する必要はないと考えられる。加速試験および苛酷試験により、最終的な保存期間の推定や、想定される分解物についての有用な情報を得られることもある。

(3) 品質特性に関する同等性/同質性の評価試験

可能かつ合理的な範囲で、目的物質、目的物質関連物質、目的物質由来不純物、工程由来不純物を含めた品質特性に関して、対照とするバイオ医薬品とバイオ後続品との類似性の評価を行うことが求められる。ただし、工程由来不純物については製法そのものが異なると想定されることから、差異があることが前提とした対応が必要となるであろう。したがって、工程由来不純物については類似性を評価するというよりも、どのような差異があるのかを明らかにし、その差異が臨床上的安全性や有効性に及ぼすインパクトを評価するためのデータとも考えられる。

構造解析では一般に原薬を用いた試験が行われるが、原薬の入手は困難な場合が多く、多くの場合、製剤そのものを用いて試験を行うか、あるいは製剤から抽出した検体を原薬の代わりに用いることが想定され、抽出精製を行う場合にはその抽出法の有用性を評価しておく必要がある。

製剤を市場から入手する場合、可能であれば複数ロットを用いて品質特性に関する比較試験を行うことが有用と考えられるが、市場での入手には限界があり、必ずしも代表的なロットが得られるわけでもないことに注意を払う必要がある。

原薬の入手の困難さや製剤ではすべての試験が適用できるわけではないと考えられることから、物理的・化学的な比較試験の実施には限界があると想定される。物理化学的試験と異なり生物活性は、製剤でも十分適用可能であり、かつ高次構造の差異の検出にも有用であることが

ら可能なかぎり複数の方法を用いて生物活性の比較試験を実施することが望ましい。特にドメインごとに機能の異なる生物活性を持つ製品では、それぞれの生物活性を比較評価することにより、構造上の類似性を評価できる可能性もある。

また、免疫反応性の比較を行うことにより構造的差異(重合体や分解の差異)についての情報やアジュバント効果のある不純物の評価に有用な情報が得られることもある。

品質特性解析の比較試験を通じて認められた先行バイオ医薬品との差異が、有効性や安全性にどのような影響があるかを評価し、その結果に基づいて非臨床・臨床で実施すべき試験をデザインすることが求められる。

(4) 規格試験法

規格及び試験方法は、ICH Q6Bガイドラインを参考に、特性解析結果および対照とした先行バイオ医薬品との同等性/同質性評価データに基づき、適切に設定することが求められる。規格試験法の設定では、製造工程管理試験との相互補完性にも考慮し、合理的に試験を設定することが有用である。

(5) 非臨床試験

臨床試験の開始に先立って、安全性についての非臨床試験は終了しておく必要がある。非臨床試験では、対照バイオ医薬品と比較評価することが適切と考えられる薬理学試験と、不純物の安全性評価のようにバイオ後続品のみを対象として試験を行うのが合理的な場合がある。すなわち、目的物質関連不純物や工程由来不純物に関しては、製造工程が異なっていることから不純物プロファイルも異なっていると考えられ、対照バイオ医薬品との比較を行うことよりも、確立した製法の特徴を考慮し、不純物そのものの特性を評価した上で独自に安全性評価を行うことが合理的と考えられる。

ただし、不純物プロファイルに差異があることを前提として、対照バイオ医薬品との安全性を比較するというアプローチを選択することも開発戦略としては考えられる。

(6) 臨床試験

バイオ後続品は、品質特性および非臨床試験結果のみによって、対照バイオ医薬品との同等性/同質性を評価することは困難と考えられ、原則的には、臨床試験により同等性/同質性を評価することが必要となる。また臨床試験の実施に当たっては、品質特性、非臨床試験結果、

ならびに対照バイオ医薬品との比較データ等に基づき、さらに対照とするバイオ医薬品に関する種々の知見も考慮して、必要かつ合理的な試験をデザインすることが適切である。

臨床試験では、まず、薬力学的試験(PK試験)、薬物動態試験(PD試験)またはPK/PD試験を対照とする先行バイオ医薬品と比較して行うことになる。原則的に、対照とする先行バイオ医薬品との薬物動態の同等性/同質性を適切にデザインされたクロスオーバー試験により確認することが必要になると思われるが、長半減期あるいは抗体産生が起こる可能性のあるバイオ後続品ではクロスオーバー試験が不適切である場合もある。したがって、健常人ではなく患者を対象とするほうが適切な場合もありえる。

このような試験により目的とする臨床エンドポイントにおける同等性/同質性を保証できる十分なデータが得られた場合には、それ以上の臨床試験を省略できる場合もありえる。

有効性に関する比較臨床試験を実施する場合には、必要かつ妥当な症例数を設定するとともに、臨床的に確立されたエンドポイントを用い、どの程度の差異の範囲であれば同等/同質とするのかについての許容域をあらかじめ規定しておく必要がある。適切な代替エンドポイントがある場合には、必ずしも真のエンドポイントを用いる必要はないが、その妥当性を裏付けるデータや文献等により十分な説明が必要とされる。

さらに、対照バイオ医薬品が複数の効能・効果を持つ場合、他の効能・効果においても薬理学的に同様の作用が期待できることが説明できるのであれば、対照バイオ医薬品の他の効能・効果をバイオ後続品に外挿することが可能となる場合もあると考えられる。

(7) 製造販売後調査

バイオ後続品の開発では、臨床試験を含めてすべての情報が得られるわけではないと想定される。特に、バイオ後続品は不純物プロファイルの差異や免疫原性の差異がある可能性があり、製造販売後に安全性プロファイル等について引き続き調査することが必要と考えられる。

また、このために有害事象のトレーサビリティを確保することが重要である。対照とする先行バイオ医薬品やそれと同一の有効成分を持ついわゆる同種・同効医薬品をバイオ後続品に変更することは可能とされるが、一連の治療期間内に混用することは基本的に避けることが望ましいと考えられる。

Chapter 18

LC/MSⁿ for Glycoprotein Analysis: N-Linked Glycosylation Analysis and Peptide Sequencing of Glycopeptides

Nana Kawasaki, Satsuki Itoh, and Teruhide Yamaguchi

Summary

Liquid chromatography/multiple-stage mass spectrometry (LC/MSⁿ) is an effective means for the site-specific glycosylation analysis of a limited quantity of glycoproteins, such as gel-separated proteins. Generally, a tryptic digest of the glycoprotein is separated by reversed-phase LC, and peptide sequencing and glycosylation analysis are achieved with on-line MSⁿ. In this chapter, a protocol for the LC/MS/MS/MS of a proteolytic digest of a gel-separated glycoprotein is described.

Key words: Glycosylation analysis, Peptide sequencing, LC/MSⁿ.

1. Introduction

A glycoprotein consists of many different glycoforms due to their carbohydrate heterogeneity at multiple glycosylation sites. Glycosylation analysis at each glycosylation site is important for gaining an understanding of the function of the glycoproteins. Mass spectrometric peptide mapping by liquid chromatography/mass spectrometry (LC/MS) is one of the effective means in use for the site-specific glycosylation analysis of glycoproteins (1, 2). Glycosylation at each site can be elucidated by the MS of glycopeptides which are separated from the proteolytic digests by on-line LC (3, 4).

LC/multiple-stage mass spectrometry (MSⁿ) has become a powerful tool for the site-specific glycosylation analysis of a limited quantity of glycoproteins, such as gel-separated proteins (5). A mixture of cellular proteins is reduced and alkylated, and then separated by SDS-polyacrylamide gel electrophoresis (PAGE) or

2D-electrophoresis. The gel band or spot that seems to contain a target glycoprotein is excised and crushed. The proteins are extracted from the gel in an intact form by shaking in 1% SDS and precipitation with acetone. The extracted proteins are digested with an appropriate enzyme which ideally provides glycopeptides with one attached sugar chain. The complex mixture of peptides is separated by reversed-phase LC, and glycosylation analysis and peptide sequencing are achieved by MS/MS and MS/MS/MS of the glycopeptides. The carbohydrate structures are fragmented and may be deduced from the MS/MS spectra whilst the peptide often remains with only a single GlcNAc attached. The peptide sequence of the glycopeptide can then be confirmed by further MS/MS/MS spectra of the [peptide + GlcNAc + H]⁺ as a parent ion.

In this chapter we describe the protocol for the LC/MS/MS of a gel-separated glycoprotein for the identification and analysis of the site-specific N-glycosylation. Thy-1, which is one of the GPI-anchored proteins and glycosylated at Asn23, 74, and 98, is used as a typical model glycoprotein (6).

2. Materials

2.1. Reduction and Carboxymethylation

1. *Buffer.* 0.5 M Tris-HCl (pH 8.6) containing 8 M guanidine hydrochloride and 5 mM EDTA.
2. *Reducing reagent.* 2-mercaptoethanol. [Dithiothreitol, and tris(carboxyethyl)phosphin (TCEP) can be used instead.]
3. *Alkylation reagent.* Sodium iodoacetate. [Iodoacetamide can be used instead].
4. *Gel filtration column for desalting.* Sephadex G-25 column (e.g., PD-10 column, GE Healthcare).

2.2. SDS-Polyacrylamide Gel Electrophoresis

1. *Gel.* A precast polyacrylamide gel (20 × 80 × 1 mm 4% stacking gel, and 80 × 80 × 1 mm 12.5% separating gel) with ten wells. Commercially available from BIO CRAFT Co. Ltd., Tokyo, Japan.
2. *Running buffer.* 25 mM Tris-HCl, 192 mM glycine and 0.1% (w/v) SDS.
3. *Sample buffer.* 50 mM Tris-HCl (pH6.8), 10% (v/v) Glycerol, 2% (w/v) SDS, 5% (w/v) 2-mercaptoethanol, and 0.0025% (w/v) bromophenol blue.
4. *Dye reagents.* Coomassie[®] G-250 (Simple Blue SafeStain). Commercially available from Invitrogen (Carlsbad, CA, USA).

5. *Prestained molecular weight markers.* Precision Plus Protein™ All Blue Standards (Bio-Rad Laboratories, Inc.).
6. *Instrument.* Vertical Slab gel electrophoresis unit (BIO CRAFT). Power supply.

2.3. Extraction of Glycoproteins and Proteolytic Digestion

1. *Extraction buffer.* 20 mM Tris-HCl (pH 8.0) containing 1% (w/v) SDS. [1% (w/v) octylglucoside can be used instead.]
2. Solvent for glycoprotein precipitation: Cold acetone.
3. *Filter for removal of gel pieces.* Ultrafree-MC (0.22 μm, Millipore Corporation, Bedford, USA).
4. *Protease.* Trypsin-Gold (Promega, Madison, WI, USA) (*see Note 1*).
5. *Proteolytic digestion buffer.* 0.1 M Tris-HCl (pH 8.0).
6. *Vacuum centrifuge.* SpeedVac concentrator (Thermo Fisher Scientific, Waltham, MA, USA).

2.4. LC/MSⁿ and Database Search Analysis

1. *LC equipment.* HPLC system capable of binary gradient formation at a flow rate of μl/min or nl/min (e.g., Paradigm MS4, Michrom BioResources Inc., Auburn, CA, USA).
2. *Column.* A reversed-phase (RP) column (0.2 mm i. d. × 50 mm) (*see Note 2*). A C18 or C30 column is commonly used for the peptide/glycopeptide mapping. Graphitized carbon column (Thermo Fisher Scientific) can be used for the analysis of small hydrophilic glycopeptides.
3. *Tip.* Nanospray tip and *x-y-z* translational stage. Commercially available from AMR (Tokyo, Japan).
4. An ion trap type mass spectrometer (e. g., LTQ, Thermo Fisher Scientific) equipped with ESI.
5. *LC mobile phase.* HPLC grade acetonitrile and formic acid. Solvent A: 0.1% formic acid in 2% acetonitrile. Solvent B: 0.1% formic acid in 90% acetonitrile.
6. *Search engine.* TurboSEQUENT (Thermo Fisher Scientific) or Mascot (Matrix Sciences, London UK).
7. *Database.* Swiss-Prot or NCBI-nr.

3. Methods

3.1. Reduction and Alkylation of Glycoproteins (See Note 3)

1. Dissolve the crude fraction (360 μg) in 360 μl of reduction and carboxymethylation buffer.
2. Incubate the glycoprotein sample with 2-mercaptethanol (2.6 μl) at room temperature for 2 h.

3. To the reaction mixture, add 7.56 mg of sodium iodoacetate and incubate the mixture at room temperature for 2 h in the dark.
4. Apply the mixture (2.5 ml) to the PD-10 column equilibrated with water. Pass through the column to remove the excess reagent.
5. Wash the column with 3.5 ml of water and collect the eluant.
6. Lyophilize the eluant.

3.2. SDS-Polyacrylamide Gel Electrophoresis

1. Set a polyacrylamide gel in the slab gel electrophoresis unit.
2. Pour the running buffer into the upper and lower chambers of the gel unit.
3. Load the alkylated proteins ($\sim 25 \mu\text{g}/\sim 25 \mu\text{l}$) in a well. Include one well for a prestained molecular weight marker.
4. Connect the assembly of the gel tank to a power supply. Run at a 20 mA constant current until the dye front reaches the end of the gel.
5. Stain the proteins in the gel with Coomassie[®] G-250. The typical electrophoretic separation is shown in **Fig. 1**.

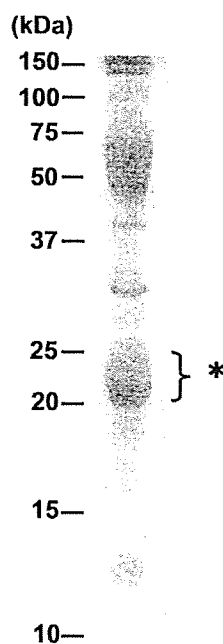


Fig. 1. SDS-PAGE of crude fraction containing the target glycoprotein (Thy-1). GPI-anchored proteins were prepared by phosphatidylinositol-specific phospholipase C treatment of a rat brain membrane fraction solubilized with Triton X-114. The crude fraction containing lipid-free soluble GPI-anchored proteins were separated by SDS-PAGE. The protein at 20–25 kDa indicated by an asterisk is a target glycoprotein, Thy-1. This band was excised, and the proteins extracted from the gel were digested with trypsin.

3.3. Extraction of the Glycoprotein from the Gel and Proteolytic Digestion (See Note 4)

1. Excise the gel bands of Coomassie-stained proteins and crush the gel pieces.
2. Extract the glycoproteins by shaking the crushed gel vigorously in 500 μ l of extraction buffer overnight.
3. Filter the extracted solution with Ultrafree-MC to remove the gel pieces
4. To the filtrated solution, add four volumes of cold acetone (2 ml) to precipitate the proteins.
5. Centrifuge the sample and remove the supernatant.
6. Incubate the precipitated glycoproteins with trypsin-gold (1 μ g) in 20 μ l of proteolytic digestion buffer at 37°C overnight.

3.4. LC/MSⁿ

1. Equilibrate the RP-column by running through 95% of solvent A at a flow rate of 3 μ l/min for 20–30 min.
2. Attach the nanospray tip to the outlet of the column, and set it in the *x-y-z* translational stage. Align the spray tip with the capillary of the mass spectrometer for accurate spray direction by moving the *x-y-z* translational stage.
3. Set up the mass spectrometer (e. g. LTQ: ESI voltage, 2.0 kV; capillary temperature, 200°C, etc.) and check the stable spray with elution buffer.
4. Inject samples (5–10 μ l) onto the column and simultaneously start both the chromatography gradient and the mass spectrometry data collection. The column is eluted typically by a linear gradient from 2% of B to 90% of solvent B in 60 min. A full MS scan (*m/z* 300–2,000) followed by MS/MS and MS/MS/MS scans are acquired in a data-dependent manner (see Note 5). Simultaneous in-source collision-induced dissociation (CID) (*m/z* 100–500) is recommended for locating the glycopeptides in the mass spectrometric peptide map (see Note 6) (3). As a typical separation of peptides, the total ion chromatogram (TIC) obtained by a full MS scan of the tryptic digested Thy-1 is illustrated in Fig. 2A.

3.5. Assignment of Glycopeptides

3.5.1. Database Search Analysis

1. Begin the database search analysis using a search engine such as TurboSEQUENT or Mascot to match the acquired spectra with predicted spectra for the known sequence in a database for protein identification. For the identification of glycopeptides which have a single remaining GlcNAc attached in the MS/MS spectrum, use the following parameters: carboxymethylation (58 Da) at Cys, and a possible modification of HexNAc (203 Da) at Asn (see Note 7).
2. After the glycopeptides have been identified by the database searching, pick out the MS/MS spectra of identified

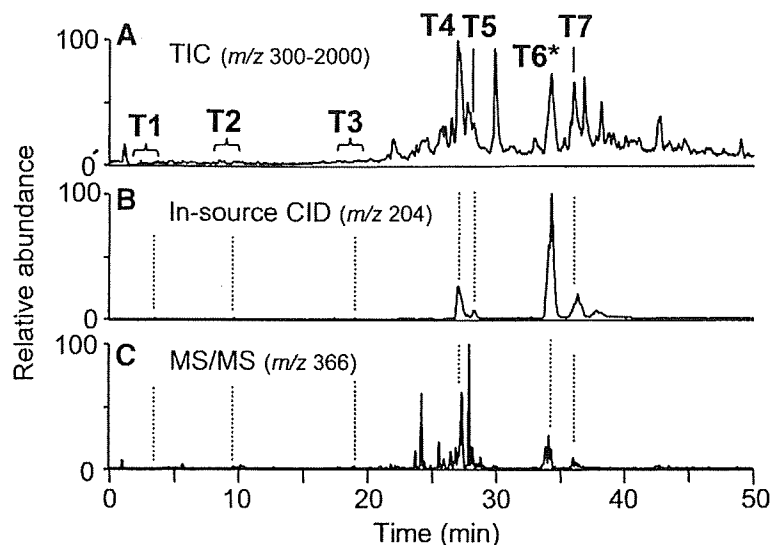


Fig. 2. Separation of tryptic digest of extracted Thy-1. (A) TIC obtained by a full MS scan (m/z 300–2,000) in the positive ion mode. (B) EIC of m/z 204 obtained by in-source CID. (C) EIC of m/z 366 obtained by data-dependent MS/MS. LC: instrument, Paradigm MS4; column, C18 (Magic C18, 0.2×50 mm, Michrom Bioresources); MS: instrument, LTQ (Thermo Fisher Scientific); ion mode, positive. T1 – T7 and the *dotted lines* depict the elution position of glycopeptides. Tryptic digestion of Thy-1 provides three major glycopeptides and some minor glycopeptides arising from Thy-1 by failed digestion.

glycopeptides from all product ion spectra. The carbohydrate sequence can sometimes be deduced from the fragment ions, such as the B and Y ions in the MS/MS spectra depending on the complexity of the glycan structure (7, 8). The peptide sequence can be confirmed by the b- and y-ions produced by MS/MS/MS (5, 9).

As the typical product ion spectra, MS/MS and MS/MS/MS spectra of glycopeptide from Thy-1 are indicated in Fig. 3. Figure 3A shows the MS/MS spectrum acquired from $[M + 2H]^{2+}$ (m/z 1,512.2) as a precursor ion. The most intense ion $[\text{peptide} + \text{GlcNAc} + H]^+$ (m/z 1,310.7) was subjected to MS/MS/MS data-dependently and identified as VLTLA(GlcNAc-)N FTTK based on the b- and y-ions by the database search analysis (Fig. 3B). From the calculated glycopeptide mass (3,022.4 Da) and the peptide theoretical mass (1,106.6), the mass of N-glycan moiety was calculated as 1,933.8 Da (dHex₂Hex₅HexNAc₄). One of two dHex (Fuc) is suggested to be attached to Gal-GlcNAc at the non-reducing end as the Lewis a/x antigen (Gal-(Fuc-)GlcNAc-) or the blood group H-determinant (Fuc-Gal-GlcNAc-) by the presence of distinctive ions, such as dHex₁Hex₁HexNAc₁⁺ (B_{2 α} , m/z 512) and dHex₁Hex₂HexNAc₁⁺ (B_{3 α} , m/z 674). The presence of Y_{1 α} ⁺ (m/z 1,456.7) reveals that the other Fuc is attached to the inner trimannosyl core GlcNAc. The

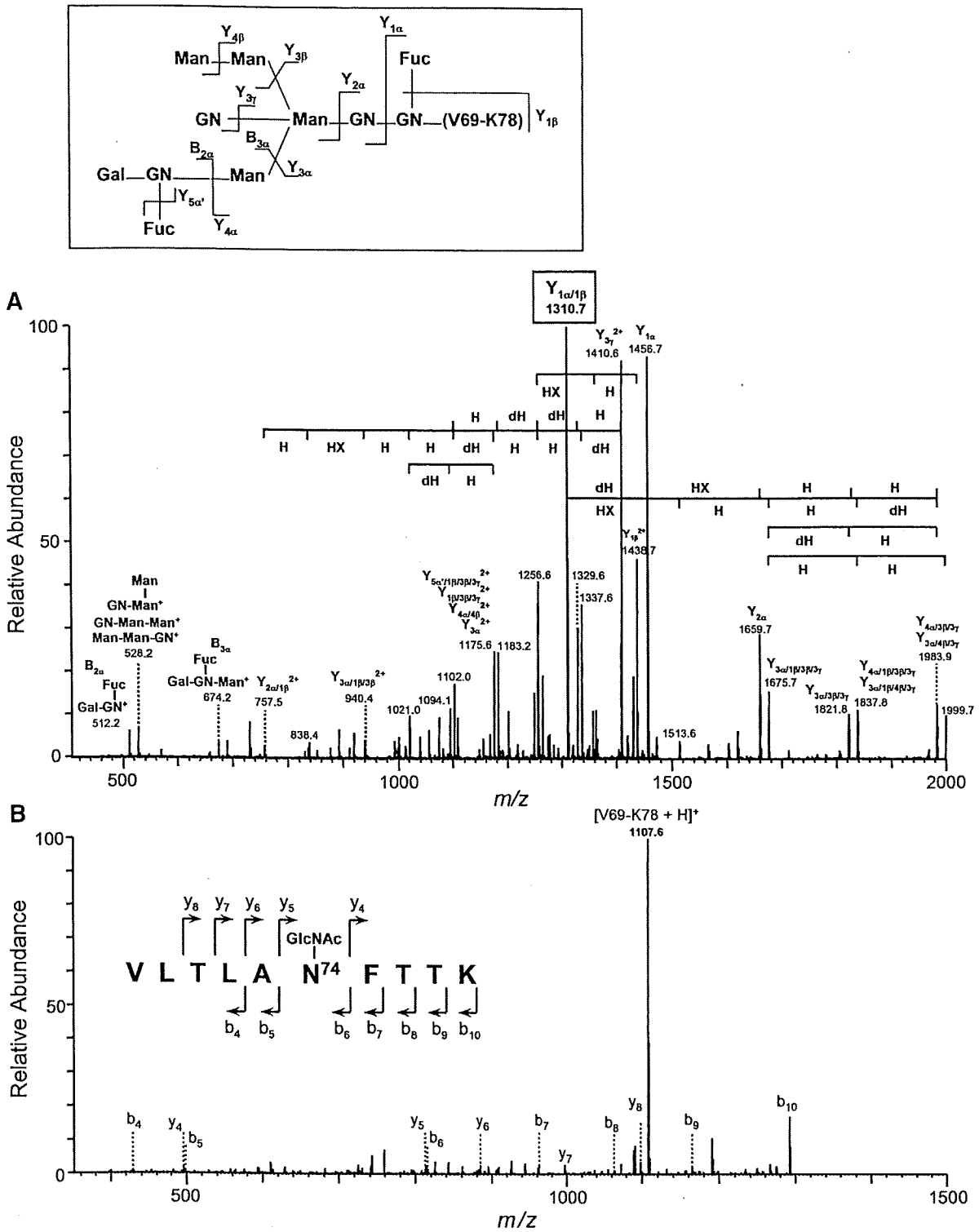


Fig. 3. Product ion spectra of Thy-1 glycopeptide eluted at 34.52 min, (A) MS/MS spectrum of the glycopeptide precursor ion: m/z 1,512.2 (2+), (B) MS/MS/MS spectrum of the glycopeptide precursor ion: m/z 1,310.7 (1+). A fucosylated hybrid type oligosaccharide was deduced from the MS/MS/MS spectrum, and Val69-Lys78 in Thy-1 was identified by database search analysis using the MS/MS/MS spectrum. *H* Hexose; *HX* *N*-acetylhexosamine; *dH* Deoxyhexose.

product ion (m/z 1,410.6) that arose from the precursor ion (m/z 1,512.2) by loss of 101.6 u (HexNAc) and [GlcNAc-Man-GlcNAc-GlcNAc-peptide (V69-K78) + H]⁺ ($Y_{3\alpha/1\beta/3\beta}$, m/z 940.4) suggests a linkage of a bisecting GlcNAc. From these product ions, it was deduced that the structure of the N-glycan moiety is a hybrid type oligosaccharide as indicated in Fig. 3, inset.

3.5.2. Assignment of Glycopeptides Not Identified by Database Search Analysis

1. Most of the glycopeptides can be located in the chromatogram using oxonium ions generated by CID. To estimate the elution positions of glycopeptides, including unidentified glycopeptides by database searching, draw the extracted ion chromatogram (EIC) of m/z 204 (HexNAc⁺) or m/z 366 (HexHexNAc⁺) produced by in-source CID or CID-MS/MS. The peaks appearing in the EIC can be assumed as glycopeptide peaks. Fig. 2B, C show the EICs of HexNAc⁺ (m/z 204) produced by in-source CID of the tryptic digest of Thy-1 and HexHexNAc⁺ (m/z 366) that were obtained by data-dependent CID-MS/MS of the digest, respectively (10–12). After the elution times of glycopeptides have been assumed, pick out the MS/MS spectra acquired when the glycopeptides were eluted. Ascertain whether the MS/MS spectra were really acquired from glycopeptides or not, by making a check on the presence of carbohydrate distinctive ions, such as HexHexNAc⁺ (m/z 366) and NeuAcHexHexNAc⁺ (m/z 657) (see Note 8).
2. Peptides which have a single remaining GlcNAc are often detected as intense ions in the picked MS/MS spectra. The peptide moiety can be assigned by matching the m/z value of the intense ion with the theoretical value of the peptide bearing GlcNAc. The peptide sequence could be ascertained if an additional LC/MS/MS/MS would be performed by setting the m/z values of peptide-related ions as precursors. Once the peptide can be assigned, the carbohydrate structure may be deduced from the MS/MS spectra.

4. Notes

1. *Enzyme specificity is as follows.* Trypsin, Lys/Arg-↓-X; Lys-C, Lys-↓-X; Glu-C, Glu (Asp)-↓-X; AspN, X-↓-Asp (Glu). Glu-C digests without cleavage of Asp-X at a 1:50 enzyme-to-substrate ratio at pH 8.0, and Asp-X can be hydrolyzed by Glu-C at a 1:4 enzyme-to-substrate ratio.

2. *Typical flow rate against the column diameter is as follows.* 0.3 mm, 5–10 $\mu\text{l}/\text{min}$; 0.2 mm, 3 $\mu\text{l}/\text{min}$; 0.1 mm, 0.5 $\mu\text{l}/\text{min}$; 0.075 mm, 0.2 $\mu\text{l}/\text{min}$.
3. In-gel reduction and carboxymethylation could be employed instead of solution reduction and carboxymethylation as follows (13):
 - (a) Incubate destained and dried gel pieces with 10 mM dithiothreitol in 25 mM NH_4HCO_3 at 56°C for 1 h.
 - (b) After washing gel pieces with 25 mM NH_4HCO_3 , incubate gel pieces with 55 mM sodium monoiodoacetic acid in 25 mM NH_4HCO_3 at room temperature in the dark for 45 min.
4. SDS/acetone is used for the extraction and precipitation of most glycoproteins. For soluble glycoproteins, extraction with 1% octylglucoside and gel filtration with Sephadex G-25 for the removal of octylglucoside can be used instead.
5. A data-dependent scan is performed for the most intense ion at each scan. The positive ion mode is used for neutral and mono/disialylated glycopeptides, and the additional negative ion mode is recommended for tri/tetrasialylated and sulfated glycopeptides.
6. Monitoring of the oxonium ions at m/z 204 (HexNAc^+) and 366 (HexHexNAc^+) is effective to locate both *N*- and many but not all *O*-glycosylated peptides (14). In most cases, an ion trap type mass spectrometer cannot detect small fragment ions, such as HexNAc^+ , due to its low mass cut-off system. In-source CID is recommended for glycopeptide mapping when an ion trap type mass spectrometer is used.
7. For the identification of peptides, care should be taken that an *N*-terminal glutamine residue is converted to pyroglutamic acid, which leads to a 17 Da mass loss from the predicted peptide. A similar mass loss is observed on an *N*-terminal carboxyamidomethylated cysteine residue (alkylated with iodoacetamide) (15).
8. Addition to the oxonium ions at m/z 204 and 366, monitoring of m/z 292 (NeuAc^+) and m/z 512 ($\text{Hex}-(\text{dHex})\text{HexNAc}^+$, Lewis a/x motif) is also useful to locate the glycopeptides. For sulfated, sialylated, and fucosylated oligosaccharides, a neutral loss of 80, 291, and 146, respectively, are often used to assume their elution position in the chromatogram.

References

1. Wuhler, M., A.M. Deelder, and C.H. Hokke (2005) Protein glycosylation analysis by liquid chromatography-mass spectrometry: *J Chromatogr B Analyt Technol Biomed Life Sci* 825, 124–33.
2. Zaia, J. (2004) Mass spectrometry of oligosaccharides: *Mass Spectrom Rev* 23, 161–227.
3. Guzzetta, A.W., et al. (1993) Identification of carbohydrate structures in glycoprotein peptide maps by the use of LC/MS with selected ion extraction with special reference to tissue plasminogen activator and a glycosylation variant produced by site directed mutagenesis: *Anal Chem* 65, 2953–62.
4. Ling, V., et al. (1991) Characterization of the tryptic map of recombinant DNA derived tissue plasminogen activator by high-performance liquid chromatography-electrospray ionization mass spectrometry: *Anal Chem* 63, 2909–15.
5. Itoh, S., et al. (2005) Characterization of a gel-separated unknown glycoprotein by liquid chromatography/multistage tandem mass spectrometry: analysis of rat brain Thy-1 separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis: *J Chromatogr A* 1094, 105–17.
6. Parekh, R.B., et al. (1987) Tissue-specific N-glycosylation, site-specific oligosaccharide patterns and lentil lectin recognition of rat Thy-1: *Embo J* 6, 1233–44.
7. Domon, B. and C.E. Costello (1988) A systematic nomenclature for carbohydrate fragmentation in FAB-MS/MS spectra of glycoconjugates: *Glycoconjugate J* 5, 397–409.
8. Demelbauer, U.M., et al. (2004) Determination of glycopeptide structures by multistage mass spectrometry with low-energy collision-induced dissociation: comparison of electrospray ionization quadrupole ion trap and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron time-of-flight approaches: *Rapid Commun Mass Spectrom* 18, 1575–82.
9. Zhang, S. and D. Chelius (2004) Characterization of protein glycosylation using chip-based infusion nanoelectrospray linear ion trap tandem mass spectrometry: *J Biomol Tech* 15, 120–33.
10. Harazono, A., et al. (2005) Site-specific glycosylation analysis of human apolipoprotein B100 using LC/ESI MS/MS: *Glycobiology* 15, 447–62.
11. Sullivan, B., T.A. Addona, and S.A. Carr (2004) Selective detection of glycopeptides on ion trap mass spectrometers: *Anal Chem* 76, 3112–8.
12. Wuhler, M., et al. (2005) Protein glycosylation analyzed by normal-phase nano-liquid chromatography – mass spectrometry of glycopeptides: *Anal Chem* 77, 886–94.
13. Kikuchi, M., et al. (2004) Proteomic analysis of rat liver peroxisome: presence of peroxisome-specific isozyme of Lon protease: *J Biol Chem* 279, 421–8.
14. Huddleston, M.J., M.F. Bean, and S.A. Carr (1993) Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective detection of glycopeptides in protein digests: *Anal Chem* 65, 877–84.
15. Krokhin, O.V., W. Ens, and K.G. Standing (2003) Characterizing degradation products of peptides containing N-terminal Cys residues by (off-line high-performance liquid chromatography)/matrix-assisted laser desorption/ionization quadrupole time-of-flight measurements: *Rapid Commun Mass Spectrom* 17, 2528–34.

Identification of Glycoproteins Carrying a Target Glycan-Motif by Liquid Chromatography/Multiple-Stage Mass Spectrometry: Identification of Lewis x-Conjugated Glycoproteins in Mouse Kidney

Noritaka Hashii,^{†,‡} Nana Kawasaki,^{*,†,‡} Satsuki Itoh,[†] Yukari Nakajima,^{†,‡} Akira Harazono,[†] Toru Kawanishi,[§] and Teruhide Yamaguchi[†]

Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan, and Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Agency (JST), 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012 Japan

Received January 20, 2009

Certain glycan motifs in glycoproteins are involved in several biological events and diseases. To understand the roles of these motifs, a method is needed to identify the glycoproteins that carry them. We previously demonstrated that liquid chromatography–multiple-stage mass spectrometry (LC–MSⁿ) allowed for differentiation of oligosaccharides attached to Lewis-motifs, such as Lewis x (Le^x, Galβ1–4(Fucα1–3)GlcNAc) from other glycans. We successfully discriminated Le^x-conjugated oligosaccharides from other N-linked oligosaccharides derived from mouse kidney proteins by using Lewis-motif-distinctive ions, a deoxyhexose (dHex) + hexose (Hex) + *N*-acetylhexosamine (HexNAc) fragment (*m/z* 512), and a Hex + HexNAc fragment (*m/z* 366). In the present study, we demonstrated that this method could be used to identify the Le^x-conjugated glycoproteins. All proteins in the mouse kidney were digested into peptides, and the fucosylated glycopeptides were enriched by lectin-affinity chromatography. The resulting fucosylated glycopeptides were subjected to two different runs of LC–MSⁿ using a Fourier-transform ion cyclotron resonance mass spectrometer (FTICR–MS) and an ion trap-type mass spectrometer. After the first run, we picked out product ion spectra of the expected Le^x-conjugated glycopeptides based on the presence of Lewis-motif-distinctive ions and assigned a peptide + HexNAc or peptide + (dHex)HexNAc fragment in each spectrum. Then the fucosylated glycopeptides were subjected to a second run in which the peptide-related fragments were set as precursor ions. We successfully identified γ -glutamyl transpeptidase 1 (γ -GTP1), low-density lipoprotein receptor-related protein 2 (LRP2), and a cubilin precursor as Le^x-conjugated glycoproteins by sequencing of 2–5 glycopeptides. In addition, it was deduced that cadherin 16, dipeptidase 1, H-2 class I histocompatibility antigen, K–K alpha precursor (H2–K(k)), and alanyl (membrane) aminopeptidase could be Le^x-conjugated glycoproteins from the good agreement between the experimental and theoretical masses and fragment patterns. The results indicated that our method could be applicable for the identification and screening of glycoproteins carrying target glycan-motifs, such as Lewis epitopes.

Keywords: liquid chromatography/multiple-stage mass spectrometry • specific detection • database search analysis • Lewis x-conjugated oligosaccharides

Introduction

Glycosylation is one of the most common post-translational modifications of proteins.^{1,2} Certain glycan motifs on glycoproteins are involved in several biological events, including cell adhesion,³ differentiation and development. They are also known to be closely associated with some diseases, such as

tumors and hepatic diseases.^{4–6} Glycomics, the study of all glycoconjugates in a cell type or in an organism, is crucial to understanding the mechanisms of glycan-mediated biological events and diseases.^{7,8} Mass spectrometry (MS) and multiple-stage mass spectrometry (MSⁿ) in combination with several types of chromatography are known to be the most powerful tools of structural glycomics.^{9–14} There are two major approaches to mass spectrometric glycome analysis. One is mass spectrometric glycan profiling, which is achieved by online or off-line liquid chromatography/mass spectrometry (LC–MS) of oligosaccharides enzymatically or chemically released from proteins.^{15–22} This technique has advantages for conducting a detailed structural analysis and a quantitative analysis of

* To whom correspondence should be addressed. Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Fax: +81-3-3700-9084. E-mail: nana@nihs.go.jp.

[†] National Institute of Health Sciences.

[‡] Japan Science and Technology Agency.

[§] Division of Drugs, National Institute of Health Sciences.

oligosaccharides,^{23–25} but it does not provide any information on protein sources that carry the glycans of interest. The other approach is the mass spectrometric mapping of proteolytic digests. This method enables us to characterize glycan structures based on fragment ions,^{26–31} and to deduce the peptide sequence from b- and y-ions that arise from the peptide backbone.²⁷ In addition to these approaches, glycomics require more advanced methods that can identify target proteins carrying a glycan motif of interest; that is to say, a technique for focused glycomics.

Lectin and immunological-based approaches have been widely used for the specific detection of target glycans and glycoproteins.^{32–34} There have been numerous reports on the use of MS in combination with affinity chromatography and Western blotting with lectins or glyco-epitope-specific antibodies.^{35–38} In a previous study, we demonstrated that LC–MSⁿ is also useful for the analysis of target glycans in a complex mixture.³⁹ Several glycan motifs often yield motif-specific ions by MSⁿ, along with common glycan-related ions such as the *N*-acetylhexosamine (HexNAc) fragment (*m/z* 204) and hexose (Hex) + HexNAc fragment (*m/z* 366).^{40,41} For example, Lewis-motifs that consist of fucose (Fuc), galactose (Gal) and *N*-acetylglucosamine (GlcNAc) yield a distinctive ion at *m/z* 512 that corresponds to the B-type ion of deoxyhexose (dHex) + Hex + HexNAc. This B-type ion subsequently provides the product ion at *m/z* 366 ([Hex + HexNAc]⁺) by MS/MS/MS. Using these Lewis-motif-distinctive ions, we successfully differentiated the oligosaccharides bearing the Lewis-motifs from many other oligosaccharides in mouse kidneys.³⁹ This method could be used to differentiate the Lewis-conjugated glycopeptides from a proteolytic digestion of proteins. Furthermore, the protein sources of the Lewis-motif-conjugated glycopeptides could be identified by further MSⁿ of peptide-related ions.

In this study we demonstrated a method for the identification of Lewis x (Le^x, Galβ1–4(Fuca1–3)GlcNAc)-conjugated glycoproteins in tissue by LC–MSⁿ. Our method consists of two different runs of LC–MSⁿ using a Fourier-transform ion cyclotron resonance mass spectrometer (FTICR–MS) and ion trap-type mass spectrometer (IT–MS). After the first run, we sorted out the product ion spectra of expected Le^x-conjugated glycopeptides based on the presence of Lewis-motif-distinctive ions and assigned a peptide + HexNAc or peptide + (dHex)HexNAc fragment in each spectrum. Then the fucosylated glycopeptides were subjected to a second run in which the peptide-related fragments were set as precursor ions (Figure 1). As a model tissue, we used a mouse kidney in which we previously confirmed the presence of the Galβ1–4(Fuca1–3)GlcNAc motif (Lewis x and y) as well as the absence of the Galβ1–3(Fuca1–4)GlcNAc motif (Lewis a and b) and Fuca1–2Galβ1–3/4GlcNAc motif (blood group H).³⁹

Experimental Section

Materials. Trypsin (Trypsin Gold, mass spectrometry grade), *Aleuria aurantia* lectin (AAL)-immobilized agarose column and Peptide-*N*-glycosidase F (PNGase F) were purchased from Promega (Madison, WI), Honen (Tokyo, Japan) and Roche (Mannheim, Germany), respectively. Murine kidneys (MRL/MpJ-lpr/lpr) were purchased from Japan SLC Inc. (Hamamatsu, Japan).

Sample Preparation. Murine kidney cells were filtrated by a cell strainer (70 μm; BD Biosciences, San Jose, CA) and solubilized in lysis buffer (7 M urea, 2 M thiourea, 2% CHAPS, and 30 mM Tris-HCl) containing a protease inhibitor mixture (Wako, Tokyo,

Japan) by vortexing at 4 °C. After quantifying the proteins, cold acetone (final concentration, 80% (v/v)) was added to the protein solution (500 μg protein). The precipitated protein was dissolved again in 100 μL of 0.5 M Tris-HCl, pH 7.0, and precipitated with an 8-fold volume of acetone. The precipitated protein was dissolved in 810 μL of 0.5 M Tris-HCl (pH 8.6) containing 8 M guanidine-HCl and 5 mM EDTA, and the mixture was incubated with 6.0 μL of 2-mercaptoethanol at room temperature for 2 h. Freshly prepared 0.6 M sodium monoiodoacetate (135 μL) was added to the solution, and the mixture was incubated at room temperature for 2 h in the dark. The reaction mixture was desalted with a PD10 column (GE Healthcare Bio-Sciences, Uppsala, Sweden), and the solution containing proteins was freeze-dried. The carboxymethylated proteins were dissolved in 500 μL of bicarbonate buffer (pH 8.5) and incubated with 2 μg of trypsin at 37 °C for 16 h. After deactivation of trypsin by boiling for 3 min, a 10-fold volume of phosphate buffered saline (PBS) was added to the reaction mixture.

Lectin Affinity Chromatography. The sample solution was applied to the AAL-immobilized agarose column (1.45 mg of lectin, 1.5 × 1.0 cm) and washed with 2.5 mL of cold PBS at 4 °C (approximately one drop/s). The absorbed glycopeptides were eluted with PBS containing 0.2 M fucose (2.5 mL), and the fraction was desalted with a C18 cartridge (Micro Trap, 8.0 × 1.0 mm; Michrom BioResources, Auburn, CA). The absorbed glycopeptides in the cartridge were eluted with 2 mL of 0.1% trifluoroacetic acid containing 45% acetonitrile, and the fraction was dried, resuspended in 0.1% formic acid, and then analyzed by LC–MSⁿ.

PNGase F Treatment. Fucosylated glycopeptides enriched by lectin affinity chromatography were treated with 10 units of PNGase F in 50 μL of 50 mM phosphate buffer (pH 8.0) at 37 °C for 48 h to release *N*-linked oligosaccharides. After terminating the reaction by boiling, the reaction mixture was evaporated to dryness, resuspended in 0.1% formic acid, and then analyzed by LC–MS/MS.

Online Liquid Chromatography/Mass Spectrometry (LC–MS). Chromatographic separation of the fucosylated glycopeptide was performed using the Paradigm MS4 HPLC system (Michrom BioResources). The fucosylated glycopeptides were dissolved in 25 μL of 0.1% formic acid, and 2 μL of the sample solution was injected into 2 μL capillary loop. The analytical column was a reversed-phase capillary column (Magic C18, 50 × 0.2 mm, 5 μm; Michrom BioResources). The mobile phase was 0.1% formic acid containing 2% acetonitrile (A buffer) and 0.1% formic acid containing 90% acetonitrile (B buffer). The fucosylated glycopeptides were eluted at a flow rate of 2 μL/min with a gradient of 5–65% of B buffer in 90 min.

Mass spectrometric analysis of fucosylated glycopeptides was performed using a FTICR/IT–MS (LTQ-FT; Thermo Fisher Scientific, Waltham, MA) equipped with a nano-electrospray ion source (AMR, Tokyo, Japan). The conditions for FTICR–MS and IT–MS were as follows: an electrospray voltage of 2.0 kV in positive ion mode, a capillary temperature of 200 °C, a tube lens offset of 140 V, a collision energy of 35% for the MSⁿ experiment, maximum injection times (FTICR–MS and IT–MS) of 1250 and 50 ms, respectively, a resolution of FTICR–MS of 50 000, a scan time of approximately 0.2 s, dynamic exclusion of 18 s, and an isolation width of 3.0 u (range of precursor ion ±1.5).

First Run. The mass spectrometric mapping of fucosylated glycopeptides was performed by a sequential scan: full mass scan using FTICR–MS (*m/z* 1000–2000), data-dependent

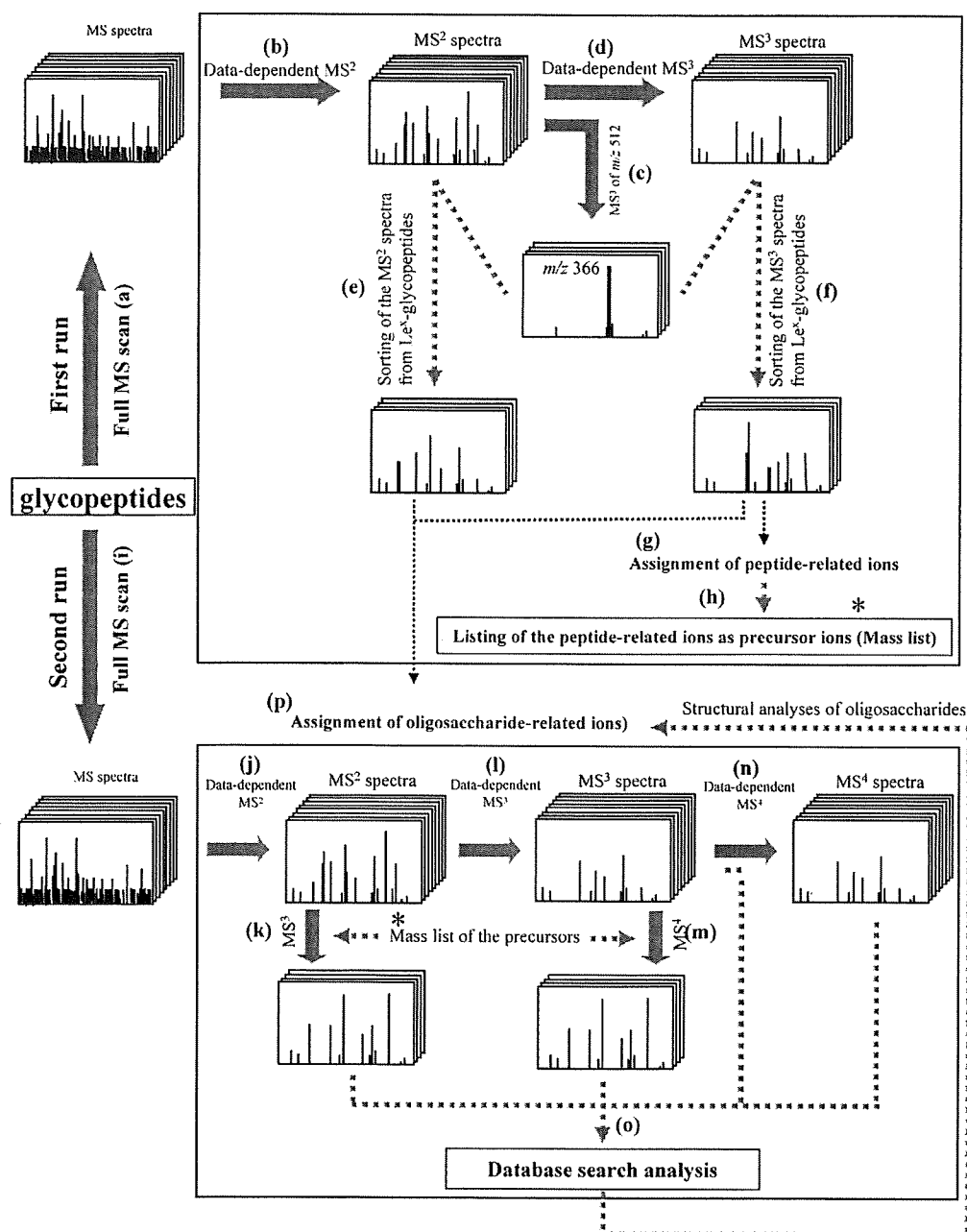


Figure 1. Strategy for the identification of Le^x-conjugated glycopeptides by LC-MS. The fucosylated glycopeptides were subjected to two different runs. In the first run, fucosylated glycopeptides were analyzed by (a) a full mass scan using the FTICR-MS, (b) data-dependent MS/MS, (c) MS/MS/MS of the Le^x-motif-distinctive ion (m/z 512), and (d) data-dependent MS/MS/MS. (e, f) On the basis of the presence of the product ion at m/z 512, which subsequently yielded the ion at m/z 366 by the MS/MS/MS, MS/MS and MS/MS/MS spectra of Le^x-conjugated glycopeptides were picked out from all the acquisition data. (g) Peptide-related ions were ascertained in the MS/MS/MS spectra. (h) The peptide-related ions were listed as precursor ions. In the second run, the fucosylated glycopeptides were identified by a full MS scan using the (i) FTICR-MS, (j) data-dependent MS/MS, (k) MS/MS/MS of the peptide-related ions listed as precursor ions, (l) data-dependent MS/MS/MS, (m) MS/MS/MS/MS of the peptide-related ions, and (n) data-dependent MS/MS/MS/MS. (o) Product ion spectra were submitted to database search analysis with a static modification of Cys with carboxymethyl (58.0 u) and possible modification of Asn with HexNAc (203.1 u) and dHex + HexNAc (349.1 u). MS², MS/MS; MS³, MS/MS/MS; MS⁴, MS/MS/MS/MS. (p) Carbohydrate structures were deduced from the oligosaccharide-related ions in the MS/MS and MS/MS/MS spectra.

MS/MS, MS/MS/MS of the Le^x-motif-distinctive ion (m/z 512) generated in data-dependent MS/MS, and MS/MS/MS of the most intense ions generated in data-dependent MS/MS by MSⁿ using IT-MS. Based on the presence of the product ion at m/z 512 that subsequently yielded the ion at

m/z 366 by the MS/MS/MS, the MS/MS and MS/MS/MS spectra of expected Le^x-conjugated glycopeptides were picked out from all the acquisition data. Peptide-related ions were ascertained in the MS/MS/MS spectra, and were listed as precursor ions for the second run.

Second Run. Peptide sequencing was performed in a sequential scan: a full mass scan using FTICR-MS (m/z 1000–2000), data-dependent MS/MS, MS/MS/MS of the peptide-related ions listed as precursor ions, data-dependent MS/MS/MS, MS/MS/MS/MS of the peptide-related ions, and data-dependent MS/MS/MS/MS using IT-MS.

Protein Identification by Database Search Analysis. The spectra data obtained by the data-dependent collision-induced dissociation (CID)s and predominant CIDs of peptide-related ions were subjected to database search analysis with the TurboSEQUEST algorithm (BioWorks 3.1; Thermo Fisher Scientific) by using the NCBI database (*Mus musculus*, 28–11–06). The static modification of carboxymethylation (58.0 u) at Cys, and the possible modification of HexNAc (203.1 u) and dHex + HexNAc (349.1 u) at Asn were used as the modified parameters of database search analysis. The SEQUEST criteria, known as cross correlation (Xcorr) scores, were set to 1.5/2.0/2.5 (charge states of +1/+2/+3) for the protein identifications. DTA files were generated for spectra with a threshold of 10 ions and a TIC of 100. Precursor and fragment ion mass tolerance in the MSⁿ spectra for database search analysis were set to 2.0 u and 1.0 u, respectively.

Results

Sorting out the Product Ion Spectra of the Le^x-Conjugated Glycopeptides by the First LC-MS/MS/MS Run. Proteins from mouse kidney cells were carboxymethylated and digested with trypsin. In the mass spectrometric mapping of the proteolytic digest, we often fail to acquire glycopeptide ions due to their lower ionization efficacy compared to coeluted unmodified peptides. To prevent interference by peptides in the ionization of glycopeptides, fucosylated glycopeptides including Lewis-motif-conjugated glycopeptides were enriched by affinity chromatography with an AAL-immobilized agarose column. The fucosylated glycopeptides were desalted by a C18 cartridge and injected into an LC-MS system equipped with FTICR-MS for a MS scan (m/z 1000–2000). The most intense ions on the MS scan were subjected to data-dependent MS/MS and MS/MS/MS, and an additional MS/MS/MS was performed when a dHex + Hex + HexNAc fragment (m/z 512) was detected on the MS/MS scan (Figure 1, first run). Figure 2A, B and C show the total ion chromatograms (TIC) obtained by the FTICR-MS scan of the fucosylated glycopeptides, the extracted ion chromatogram (EIC) of the dHex + Hex + HexNAc fragment acquired on the data-dependent MS/MS scan, and the EIC of the Hex + HexNAc fragment (m/z 366) that arose from the fragment at m/z 512 by the MS/MS/MS, respectively. We presumed that the Le^x-conjugated glycopeptides had been eluted around the peaks appearing in Figure 2C.

Assignments of the Peptide-Related Ions Derived from the Expected Le^x-Conjugated Glycopeptides. We picked out dozens of product ion spectra, and 22 precursors were determined to be the Le^x-conjugated glycopeptide-derived mass spectra based on the presence of Le^x-motif-distinctive ions. The elution positions of the glycopeptides are indicated in Figure 2C. In our previous report, most of the *N*-glycosylated peptides yielded peptide-related ions, such as [peptide + HexNAc + nH]ⁿ⁺ and [peptide + (dHex)HexNAc + nH]ⁿ⁺, which provided peptide fragment b- and y-ions by further CID, and subsequent database search analysis successfully revealed the peptide sequences of the glycopeptides.²⁷ In the present study, therefore, we examined the peptide-related ions in the product ion spectra of the glycopeptides 1–22 for peptide sequencing.

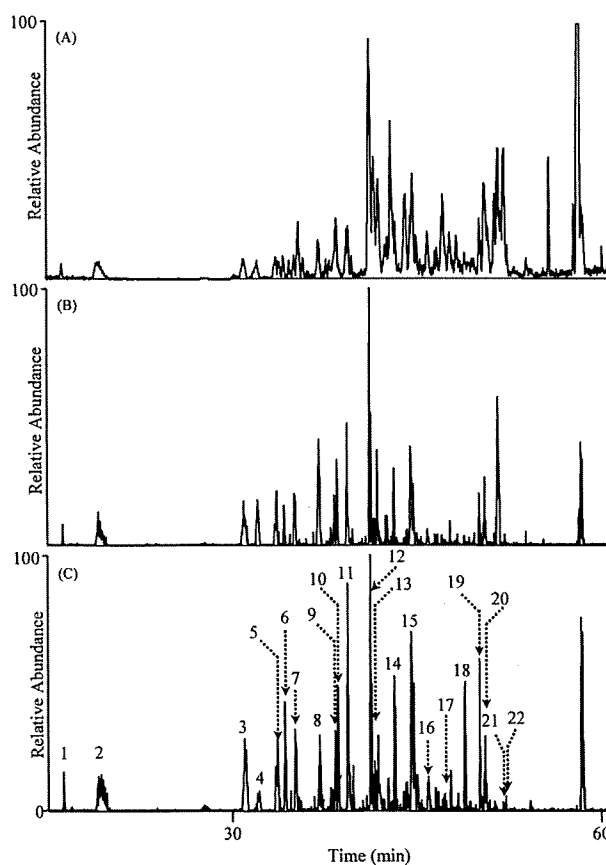


Figure 2. LC-MSⁿ of fucosylated glycopeptides. (A) Total ion chromatogram (TIC, m/z 1000–2000) of fucosylated glycopeptides. (B) Extracted ion chromatogram (EIC) of the ion at m/z 512 produced by data-dependent MS/MS. (C) EIC of the product ions at m/z 366 produced from the product ion at m/z 512 by MS/MS/MS. Expected Le^x-conjugated glycopeptides were designated as glycopeptides 1–22.

Figure 3A shows the integrated mass spectrum of the expected Le^x-glycopeptides eluted at 41–43 min. On the basis of an m/z spacing pattern that included m/z 67.69 (HexNAc³⁺), m/z 54.02 (Hex³⁺) and 48.69 (dHex³⁺), intense ions at m/z 1067.105, 1134.803, 1189.505, and 1237.508 were assigned to triply charged ions of glycopeptides differing in glycosylation. Two intense ions, glycopeptides 12 (m/z 1237.508) and 13 (m/z 1134.803), were further subjected to MS/MS and MS/MS/MS, and yielded identical peptide-related ions at m/z 1650 (Figure 3B, C). This result indicates that glycopeptides 12 and 13 are glycoforms containing Le^x-motifs, and implies that the carbohydrate heterogeneity of a peptide of interest could be deduced from the integrated mass spectra.

Figure 4A indicates the MS/MS spectrum of glycopeptide 8. The presence of a dHex + Hex + HexNAc fragment (m/z 512) suggests that this glycopeptide is one of the Le^x-conjugated glycopeptides. The most intense fragment (m/z 1681.0) was further subjected to MS/MS/MS and provided a series of doubly charged Y-ions with an m/z spacing pattern, including m/z 101 (HexNAc²⁺) and m/z 81 (Hex²⁺). It was revealed that the fragment at m/z 906.2 was our desired product ion, [peptide + HexNAc + 2H]²⁺ (Figure 4B).

Figure 5A and B show the MS/MS and MS/MS/MS spectra of glycopeptide 15, respectively. The triply charged ion (m/z 1414.4)