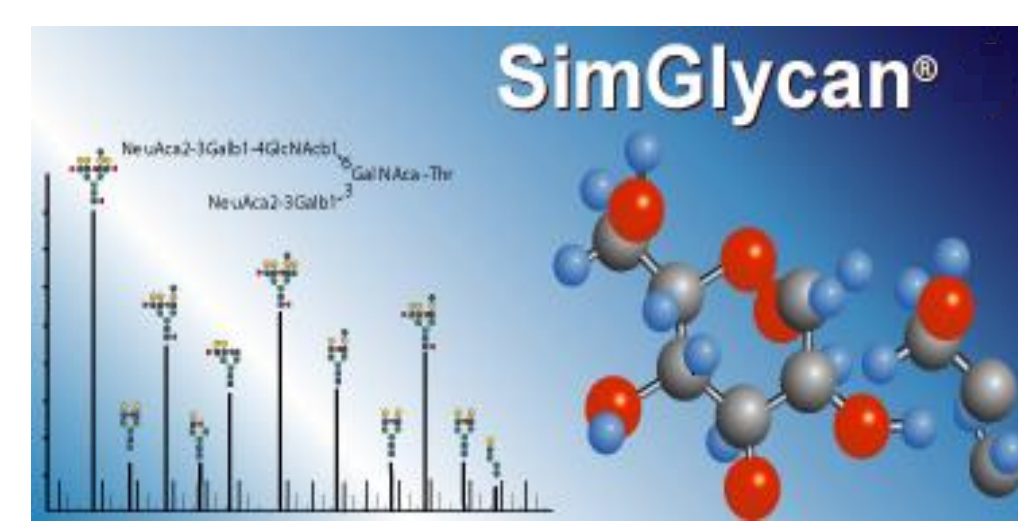


糖鎖解析ソフトウェア *SimGlycan* によるLC-MS/MSデータを利用したハイスループット構造予測



Ningombam Sanjib Meitei¹, Arun Apte², Rupanjan Goswami¹, Mrinal K. Chatterjee¹, 〇 朴 鍾 圭³

¹ PREMIER Biosoft, Indore, India, ² PREMIER Biosoft, Palo Alto, CA, USA, ³ 株式会社ネットウエルライフサイエンス営業部



Introduction

タンパク質の糖鎖付加は、良く知られた翻訳後修飾の一つであり、生体における生物学的プロセスにおいて重要な役割を担う。糖鎖の量と構造の変化をモニタリングすることで、タンパク質の糖鎖修飾と様々な疾患との関係を説明するために必要である。クライコミクス研究における新しい傾向の一つは、精密な質量分析法に基づいた糖鎖の定性および定量分析に関連したイノベーションである。近年、Thermo Fisher Scientific社から、糖質の効率的な相対定量を可能にするaminooxyTMT (Tandem Mass Tag) Reagentsが紹介された。当ラベリングキットを用いることによって、糖鎖修飾のイオン化効率が増進され、解析処理能力が増大する。しかしながら、質量分析分野でテクノロジーが急速に発展しているにもかかわらず、ハイスループットな糖鎖構造の解明が今もなおチャレンジングな課題であることは、迅速な同定を可能にするような糖鎖の構造テンプレートが欠落していることに起因する。ハイスループット糖鎖同定とそれに続く糖鎖の定量の両方を行うことができるソフトウェアツールが、クライコミクス研究を促進するために必要である。

Methods

SimGlycan® Enterprise Edition (現行version 5.4.1)は、定性および定量的な糖鎖データの分析を目的として、ESI-MS/MSやLC-ESI-MS/MSから得られたデータの解析を効率化するために開発された。Figure 1は、*SimGlycan*® のワークフローの略図を示す。

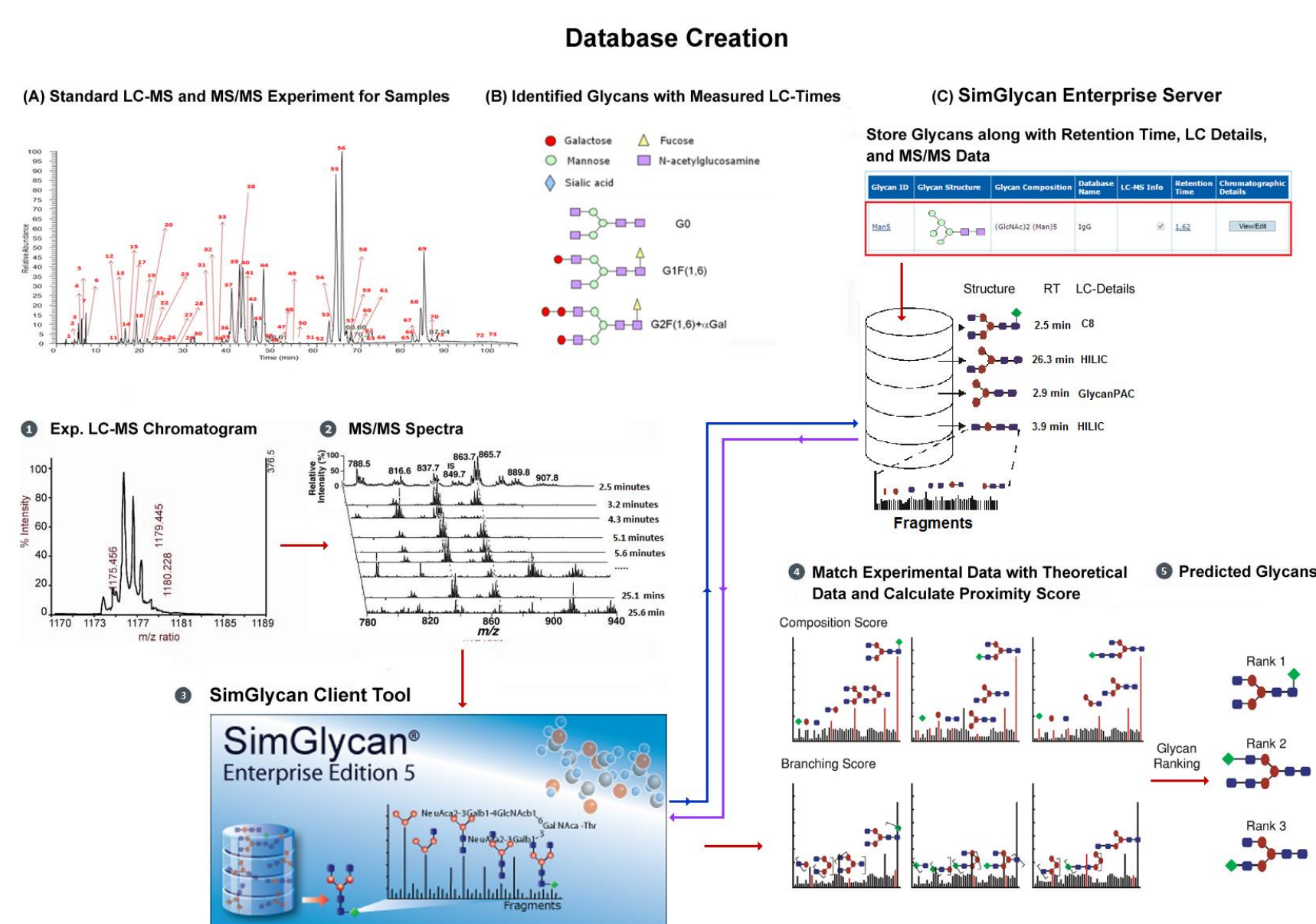


Figure 1. *SimGlycan*® のワークフロー

次に、データ解析ワークフローを自動化するために開発されたソフトウェアモジュールの詳細を紹介する。

➤ Create glycan templates

糖鎖の迅速なプロファイリングと同定をより簡単に行えるようにするために、ウェブベースのカスタムデータベースモジュールを開発した。研究者自身で糖鎖構造やLC保持時間、付随するカラム詳細、そして他の関連情報、例えば糖鎖クラスや生物試料由来、パスウェイ、生化学反応、酵素、他の公共データベースへのリンクなどを簡単に保存し、データベースのカスタマイズを容易にできるようにした。KEGG Chemical Function (KCF) 形式のファイルをバッチモードでインポートすることによって、ユーザー自身で糖鎖をデータベースへ追加できる (Figure 2)。KCFファイルがすぐに入手できない場合は、*SimGlycan*® の“Draw”モジュールでその構造を描いて、該当するKCFデータを保存することもできる (Figure 3)。

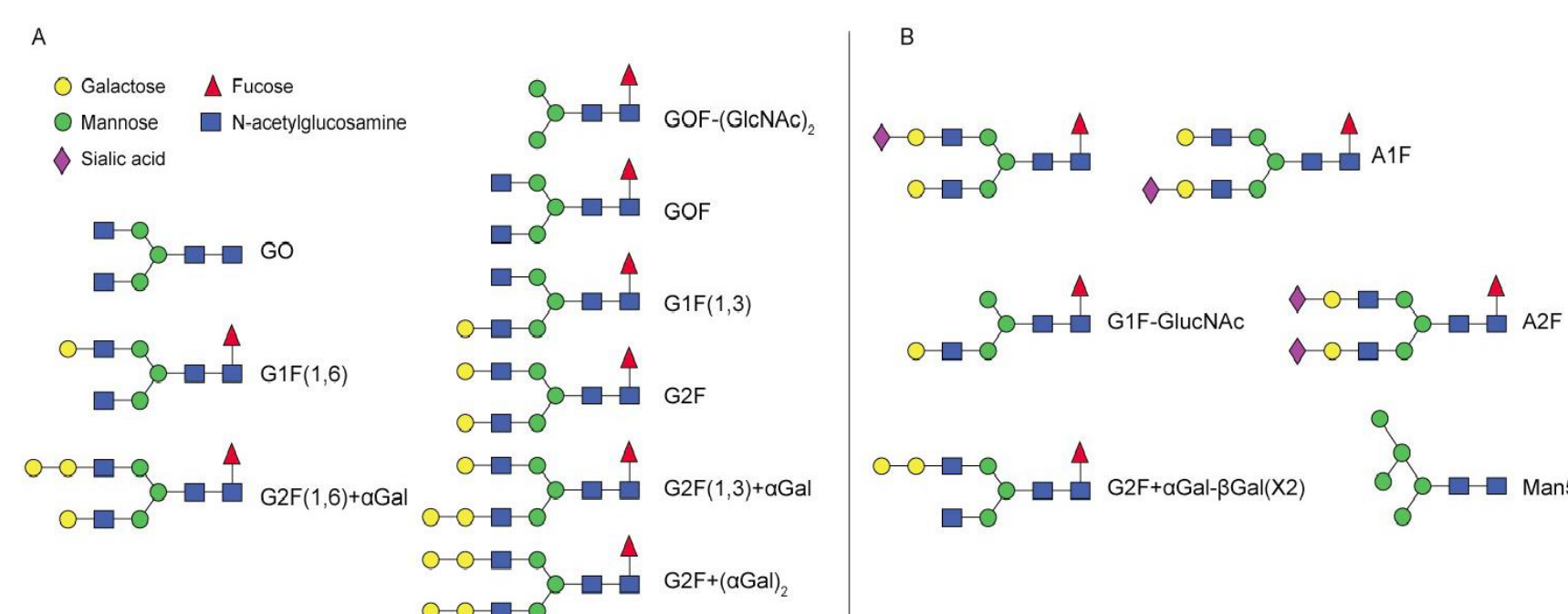


Figure 2. 糖鎖のプロトタイプテンプレート

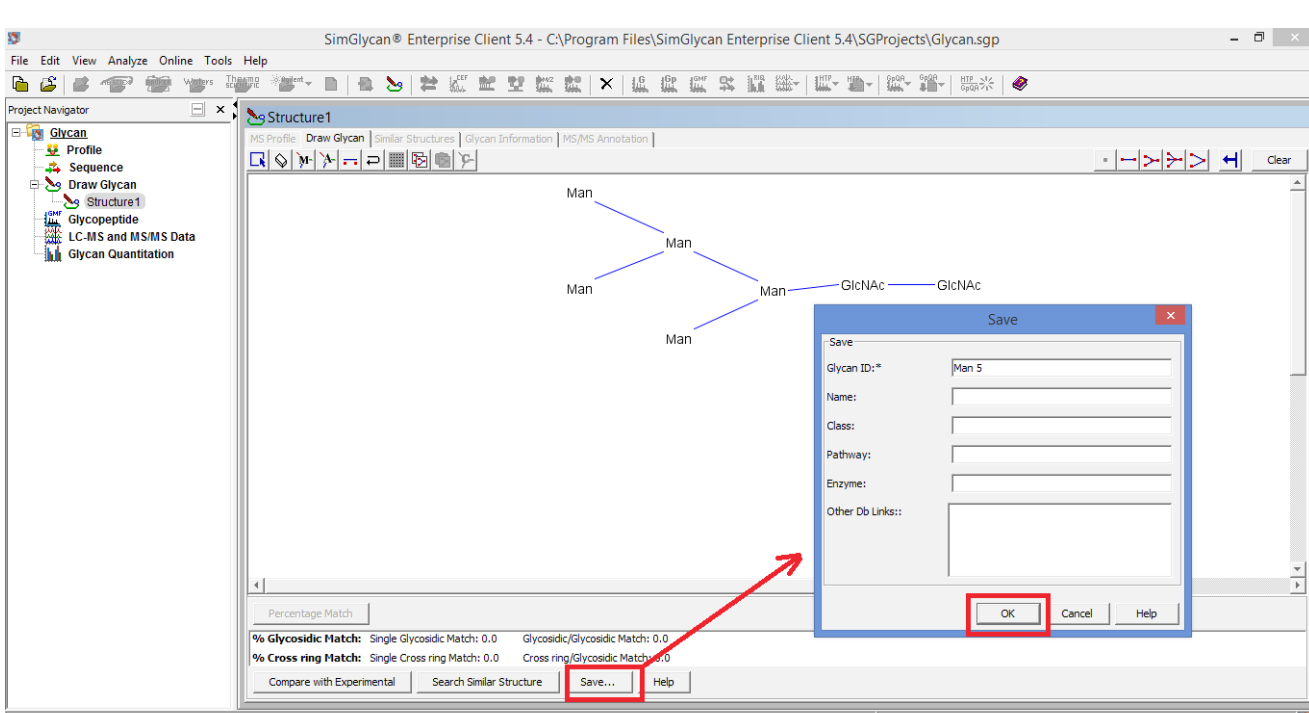


Figure 3. *SimGlycan*® の“Draw”モジュールのインターフェイス。糖鎖を描画し、KCFファイル形式で保存することができる。

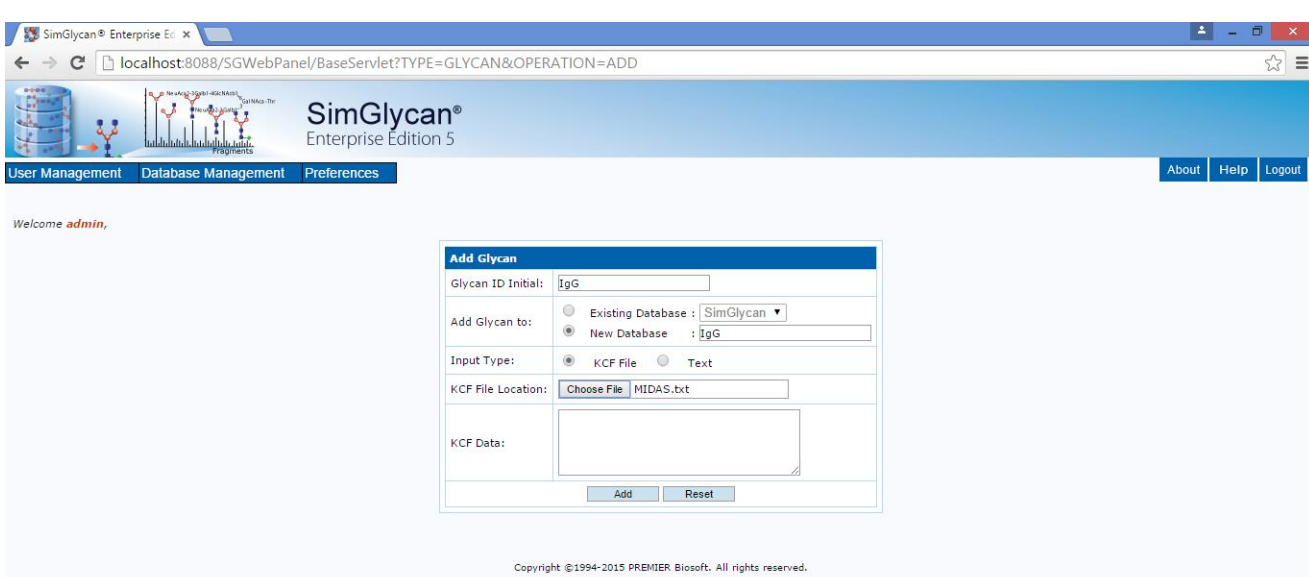


Figure 4. *SimGlycan*® で既存または新規データベースに糖鎖を追加する際のインターフェイス

他の機能として、分類群や構造情報でデータベース内を検索し該当する糖鎖を閲覧したり、あるデータベースから他のデータベースへデータを移動/コピーしたり、既存の糖鎖情報を編集することができる。Figure 4は、既存または新規データベースに糖鎖を追加する際の *SimGlycan*® のウェブブラウザインターフェイスの一例を示す。該当する糖鎖の保持時間を保存するためには、“Add” (Figure 5) をクリックし、LCの実験から得られる (複数の) 詳細情報を入力するだけ (Figure 6)。

Figure 5. 検索結果の糖鎖を表示する *SimGlycan*® のウェブページ一例

Figure 6. LC実験から得られた詳細情報を保存する際の *SimGlycan*® のインターフェイス

➤ High Throughput glycan identification

直感的なGUIに基づいたプログラムが、MS/MSデータを用いたハイスループットな糖鎖同定を可能にする (Figure 7)。バッチモードで同定を行う際には、複数のrawファイルやピークリストを指定することができる。MS1データで観測された同位体クラスターを利用して、プレカーサー *m/z* 値がMS/MSデータに対して修正されるが、ピークリストには同定された分子構造と関連MS/MSデータと一緒にLC時間で検出された化合物情報が含まれる。誘導体化や還元末端修飾、実験条件に基づいた糖鎖のフラグメンテーションパターンのような設定は、Figure 7のように指定される。

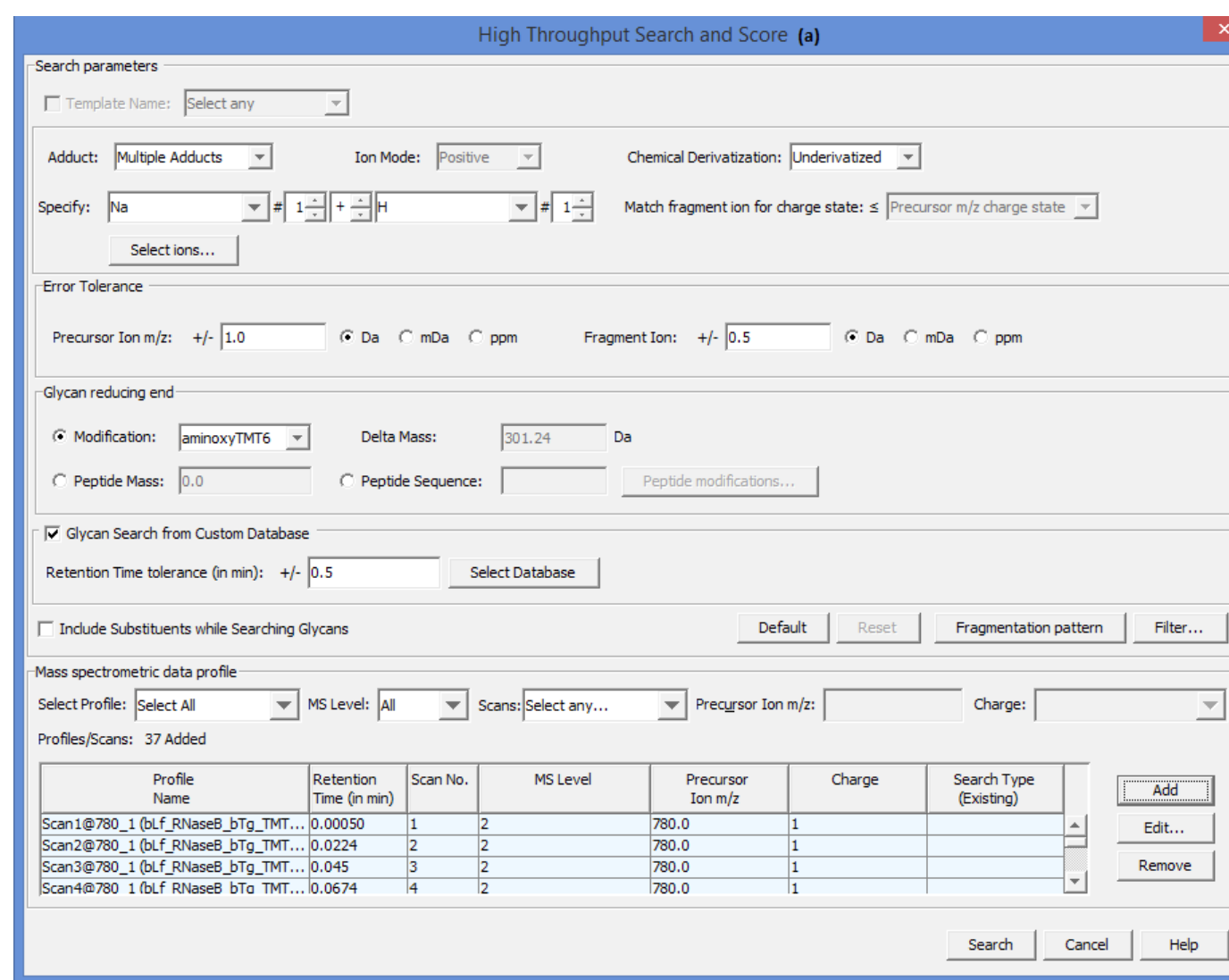


Figure 7. *SimGlycan*® のクライアントウィンドウの一例。構造検索およびスコアリングをハイスループットで行うための各種設定はダイアログボックスに従って行う。

➤ Data Filters

SimGlycan® は様々なデータフィルターが装備されており、同定した糖鎖候補のリスト中で感度を減らすことができる。クラスやサブクラス、生物試料由来、パスウェイなどに基づいて、候補構造を検査する際のデータベースに制限を設けることができる (Figure 8)。ユーザー自身でキュレーションしたカスタムデータベースが作成してある場合は、そのデータベースに限定して構造検索ができるので、同定の信頼性がより高くなる (Figure 7 & Figure 8)。

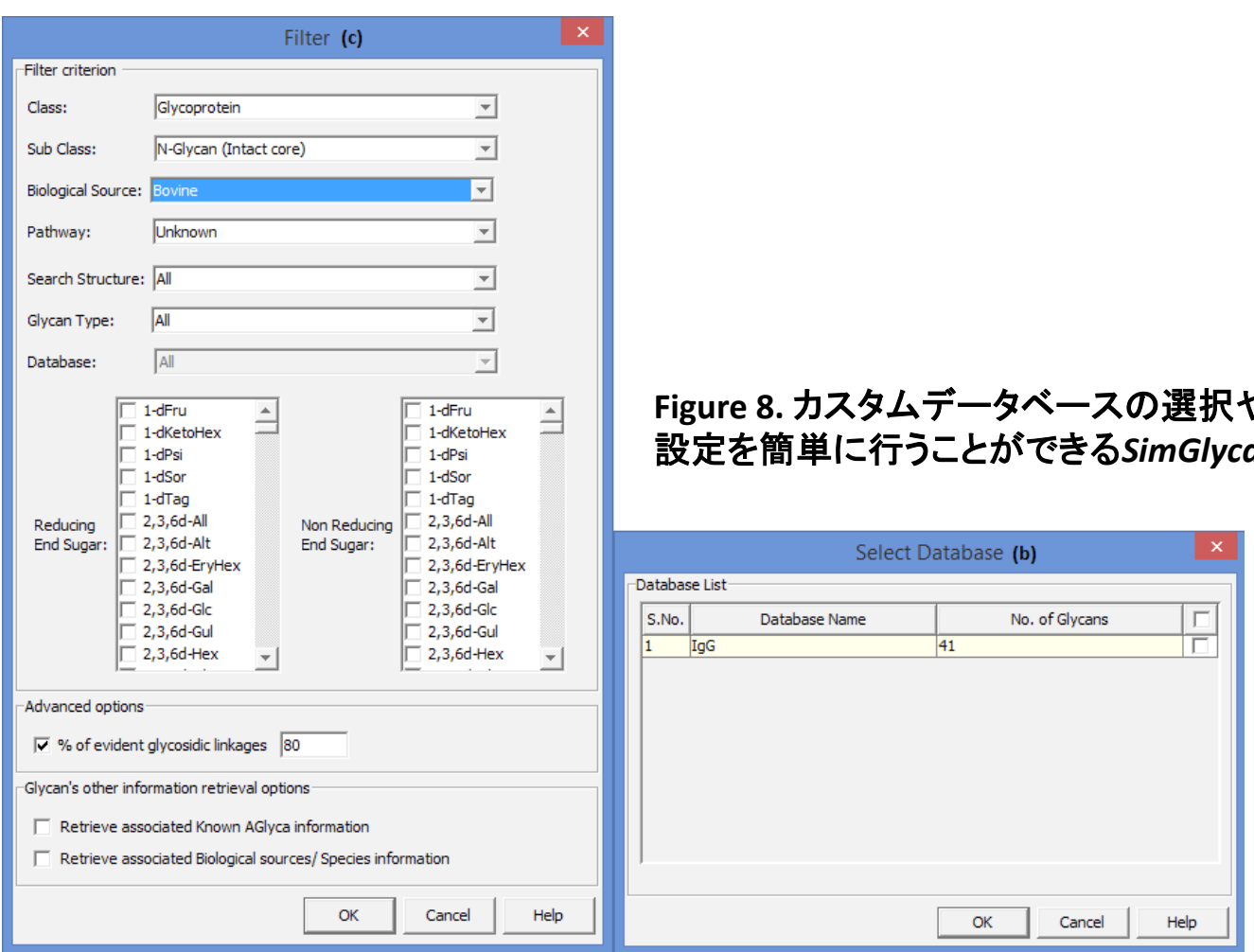


Figure 8. カスタムデータベースの選択とフィルター設定を簡単に行うことができる *SimGlycan*® のGUI

➤ Scoring Mechanism

SimGlycan® は各MS/MSデータセットに対して、観測された構造特異的な診断イオンに基づき、独自開発したスコアリングアルゴリズムを用いて最も可能性の高いマッチング順に候補糖鎖にスコアを付ける [5]。実際の手順としては、*SimGlycan*® は各MS/MSデータに対して、はじめに検索条件の述部として (保持時間も可能)、プレカーサー *m/z* を利用して候補構造のリストを作成する。プログラムは、スコアリングアルゴリズムのベースとなる観測されたフラグメントに対して、各候補のin silicoフラグメントを比較する。このアルゴリズムは、高いIntensityを持ったピーク (典型的には、単一または複数の糖残基の診断イオンに一致) と低いIntensityを持ったピーク (典型的には、糖残基のクロスリンクフラグメントの診断イオンに一致) の特徴を良くする単糖残基と分岐パターンの候補に対して、高いスコアを付ける。

SimGlycan® によるウシ糖たん白質から得られた糖鎖の同定と定量

➤ Sample Preparation

サンプルには3種の標準的な糖タンパク質 (各々ウシ由来のRNase B、ラクトフェリン、サイログロブリン) を用いた。サンプルをPNGase F (New England Biolabs, Ipswich, MA) でグリコシダーゼ処理しN型糖鎖を遊離させた。逆相クロマトグラフィーによる精製を経て、aminooxyTMT reagents (Thermo Scientific, Rockford, IL) を用い、メーカプロトコルに従って糖鎖を標識した。次に、標識されたサンプルをアセトンで洗浄し、過剰な未反応のaminooxyTMT reagentsを除去するためもう一度クリーンアップ後、20 μM NaOHを含む50%アセトリル水溶液に溶解させた。サンプルは、デュアルプレッシャーリーニアイオントラップ型の質量分析計LTQ Velos Proに直接導入し、ポジティブイオンモードで測定した。2価イオン (M+H+Na)²⁺プレカーサーは、コリジョンエネルギーの設定値を45とし、トラップ-高エネルギー衝突誘起解離法 (trap-HCD) によるフラグメンテーションによって分離および断片化させた [4]。HCDスペクトルは、5つの高マウス糖鎖のプレカーサーの *m/z*、すなわち780、861、942、1023、1104の値で取得した。

Data Analysis

SimGlycan® Enterprise Edition 5.4 (PREMIER Biosoft, USA) は、データベース作成、糖鎖同定、定量を包括的に行うことができるソフトウェアである。当プログラムでは、デフォルトの設定リストから還元末端修飾として “aminooxyTMT6” を選択することによって、ハイスループットな糖鎖同定を行うことができる。

Results and Discussion

Figure 9の左側にあるProject Navigatorには、ファイル名とMS/MSスキャンが表示される。検索結果は、2つのウィンドウ (Search Results & Annotated Peaklist) で表現される。下図では、プレカーサー *m/z* 779.8のMS/MSに対する結果がSearch Resultsウィンドウで示されている。同定された糖鎖に対して、Name、Sequence、Composition、Glycan Mass、Carbohydrate Mass、Class、ReactionsやPathway、他のデータベースへのリンクといった情報と一緒に構造が提示される。Annotated Peaklistウィンドウでは、マッチしたフラグメント、荷電状態、アダクトイオンなどの情報を持ったMS/MSスペクトルデータリスト (*m/z*, Intensity) を含むテーブルが表示される。マッチしたフラグメントの構造も確認することができる。

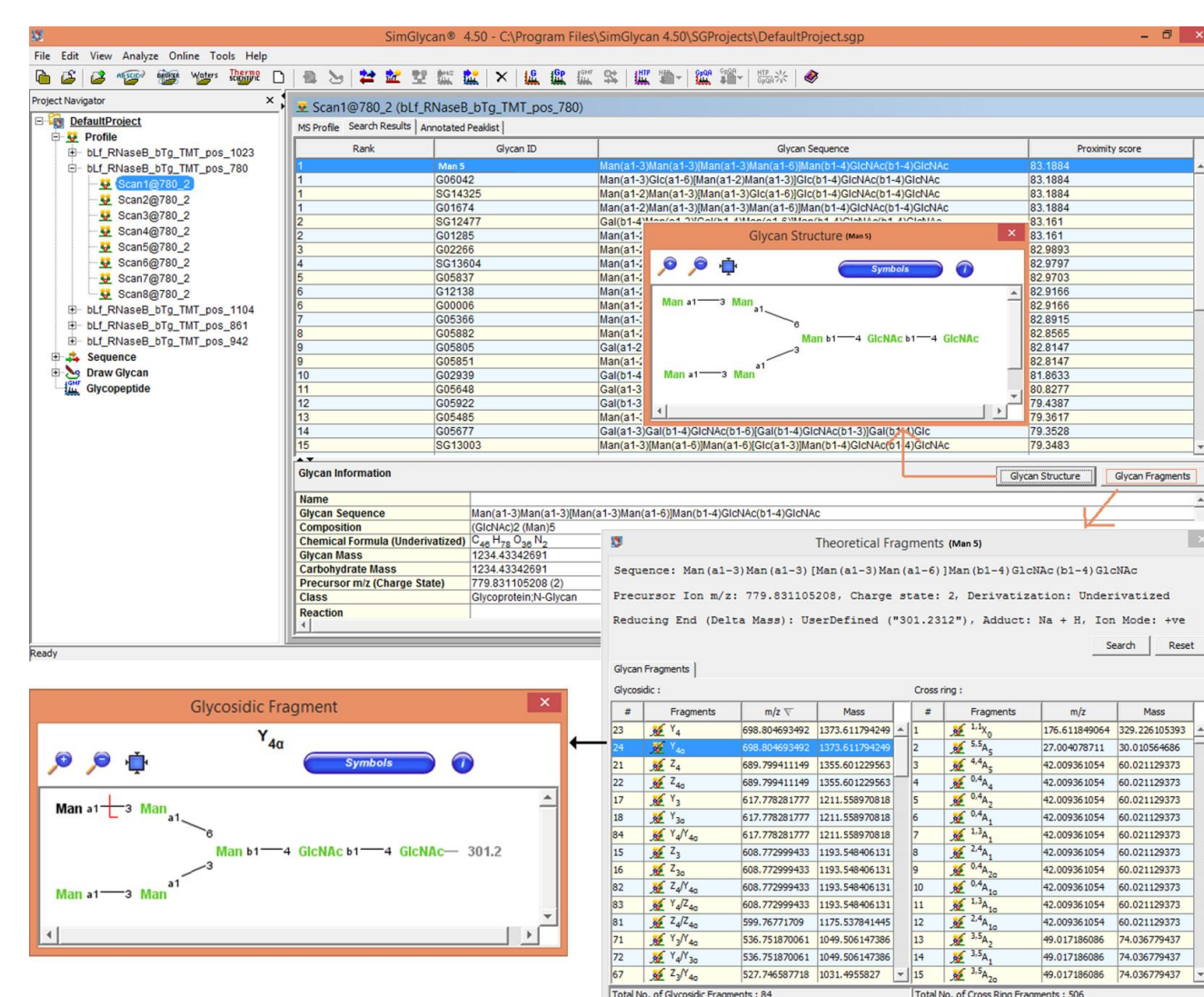


Figure 9. *SimGlycan*® の検索結果インターフェイスの一例。プレカーサー *m/z* 779.8の糖鎖 (GlcNAc)2(Man)5の構造が示されている。In silicoフラグメントとそれに一致する構造も確認できる。

➤ Quantification

SimGlycan® では、ハイスループット糖鎖同定に続いて、定量分析のためのプロジェクトを作成することができる。Figures 10とFigure 11では、*SimGlycan*® による定量分析のための設定画面を示す。プロジェクト名の指定、同重量法法の選択、MS/MSスキャンから正確なレポーターイオンIntensityを見極めるための質量許容誤差設定、Biological/Technical replicatesに応じたrawファイルのグループ化、コントロールとしてのレポーターイオンチャンネルの指定などが可能である。また、*SimGlycan*® では、同重量レポーターイオン間のオーバーラップを防ぐために、各レポーターイオンに対して許容誤差を0.01までの検索幅で設定することができる。

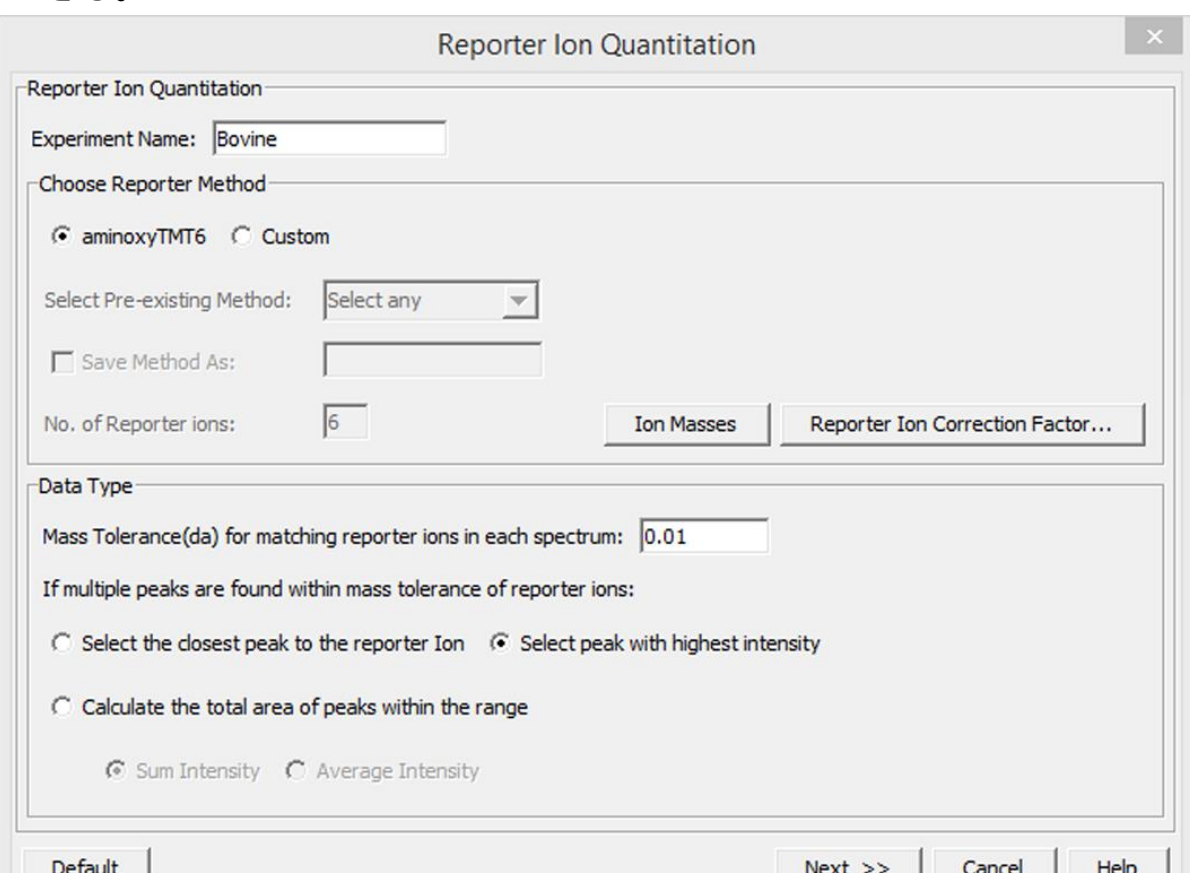


Figure 10. *SimGlycan*® における定量実験モデルのインターフェイス

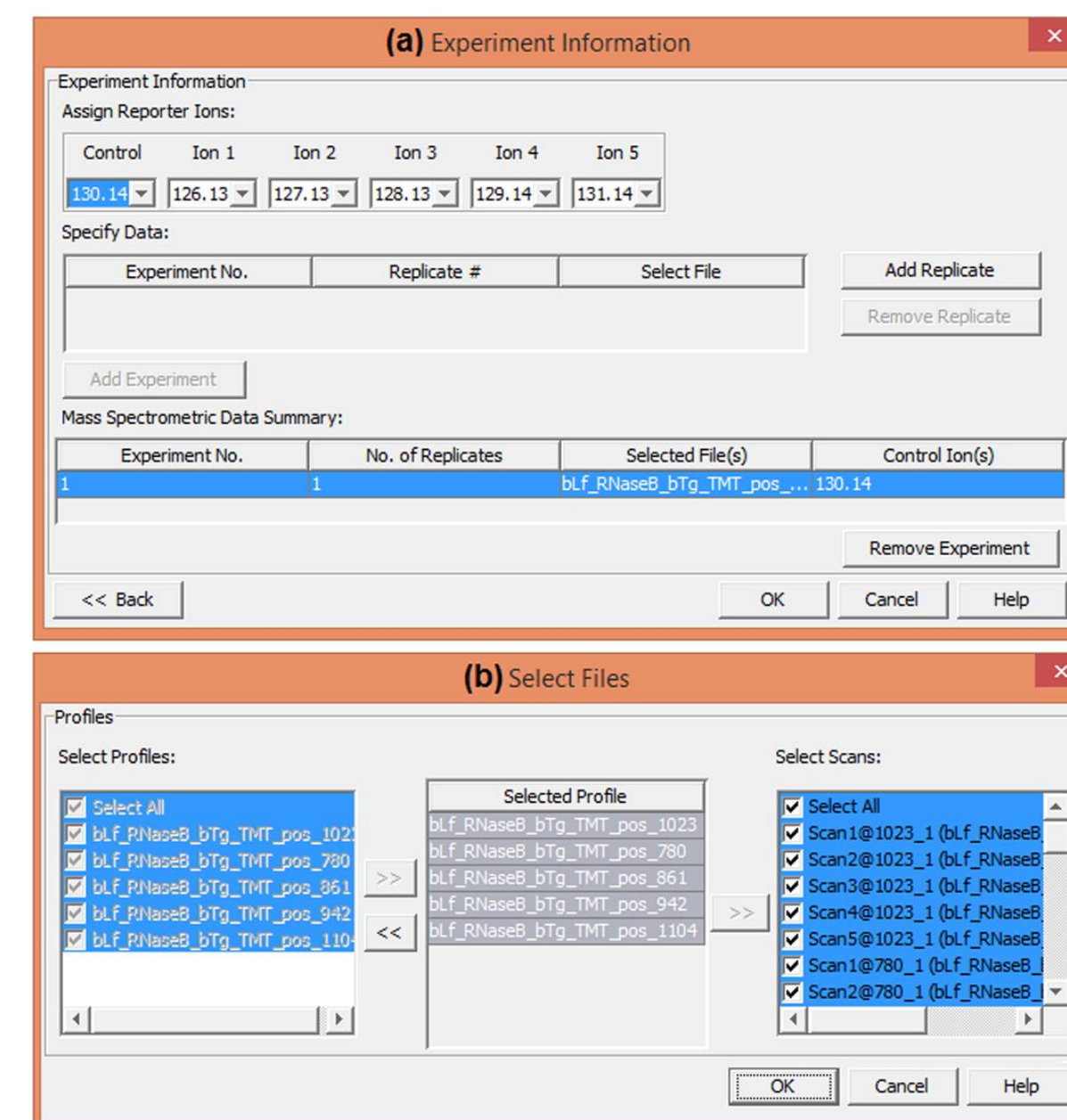


Figure 11. *SimGlycan*® における定量実験モデルのインターフェイス (a) コントロールサンプル (またはレポーターイオン) の指定、Technical replicatesの追加 (b) Technical replicatesに対するrawファイルを選択したMS/MSスキャンのアサイン

SimGlycan® が定量分析を完了すると、解析結果の統計サマリーを表示する (Figure 12)。統計サマリーでは、レポーターイオンのIntensityに関する合計値/平均値/中央値/標準偏差、コントロールとレポーターイオンのIntensity比に関する合計値/平均値/中央値、さらにそれらの比のLog2 相対発現量、各糖鎖を同定するMS/MSスペクトルの数 (glycan spectral match/ count) から構成される。

Figure 12. 同定した糖鎖と該当するレポーターイオンのIntensityを示す *SimGlycan*® の“Glycan View”

Bar Chart (Figure 13) では、各TMTチャンネルでの糖鎖毎のトータルイオンIntensityを表示する。TMT 131では、(GlcNAc)2(Man)5が最も多い量の糖鎖で、ついで(GlcNAc)2(Man)6、(GlcNAc)2(Man)8、(GlcNAc)2(Man)7、(GlcNAc)2(Man)9の順に続く。同様に、他のTMTイオンに関して、Bar Chartから直感的な把握が可能である。

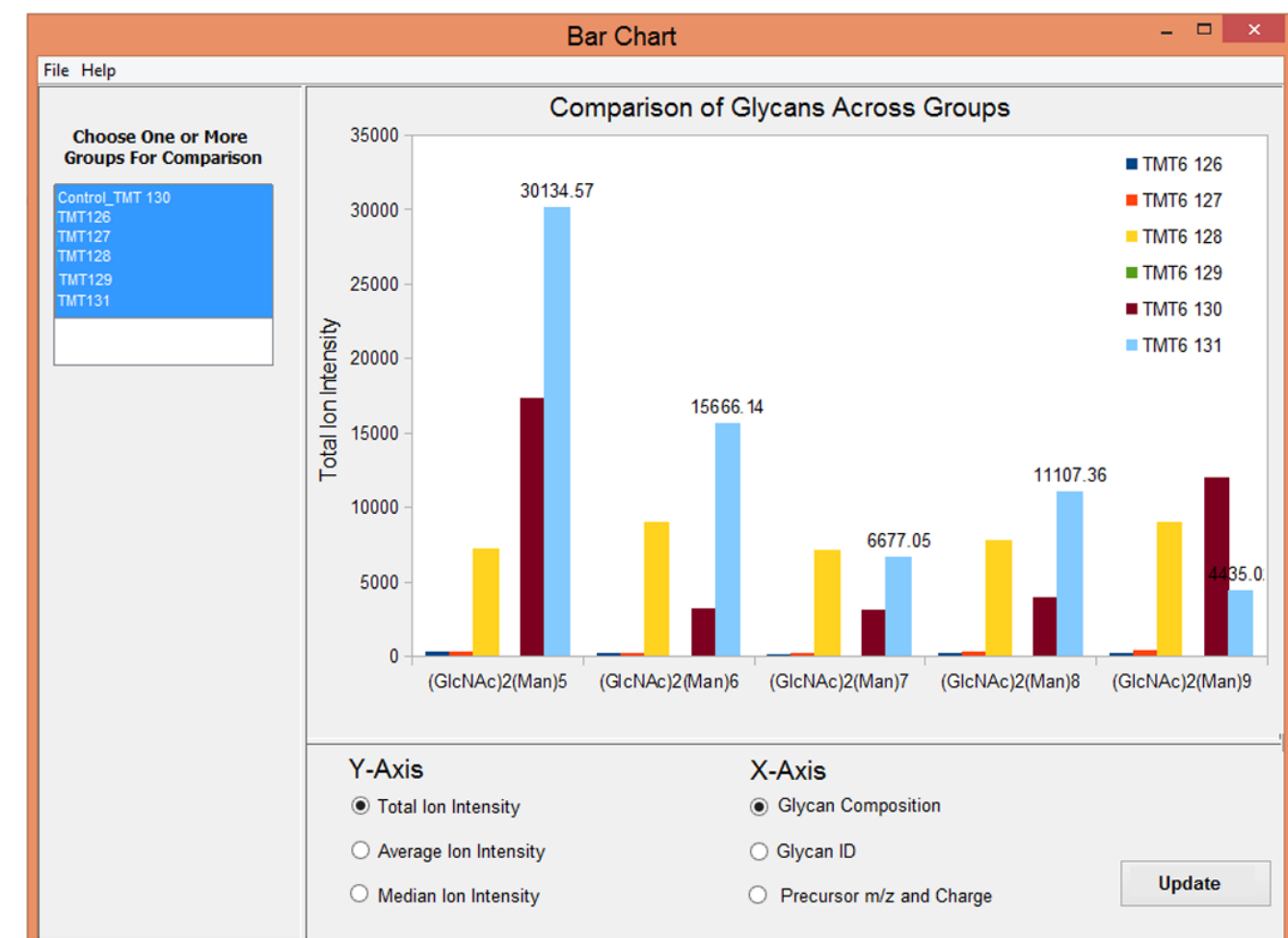


Figure 13. 各TMTレポーターイオンについて糖鎖毎のトータルイオンIntensityを表示する *SimGlycan*® のBar Chart

Cluster Dot Plot (Figure 14) では、TMTレポーターイオンをコントロールレポーターイオンで正規化したIntensity比をlog2スケールで表示する。プロットでは、TMT130をコントロールとして、各TMTレポーターイオンチャンネルに対する糖鎖の発現レベルの変化を表示した。

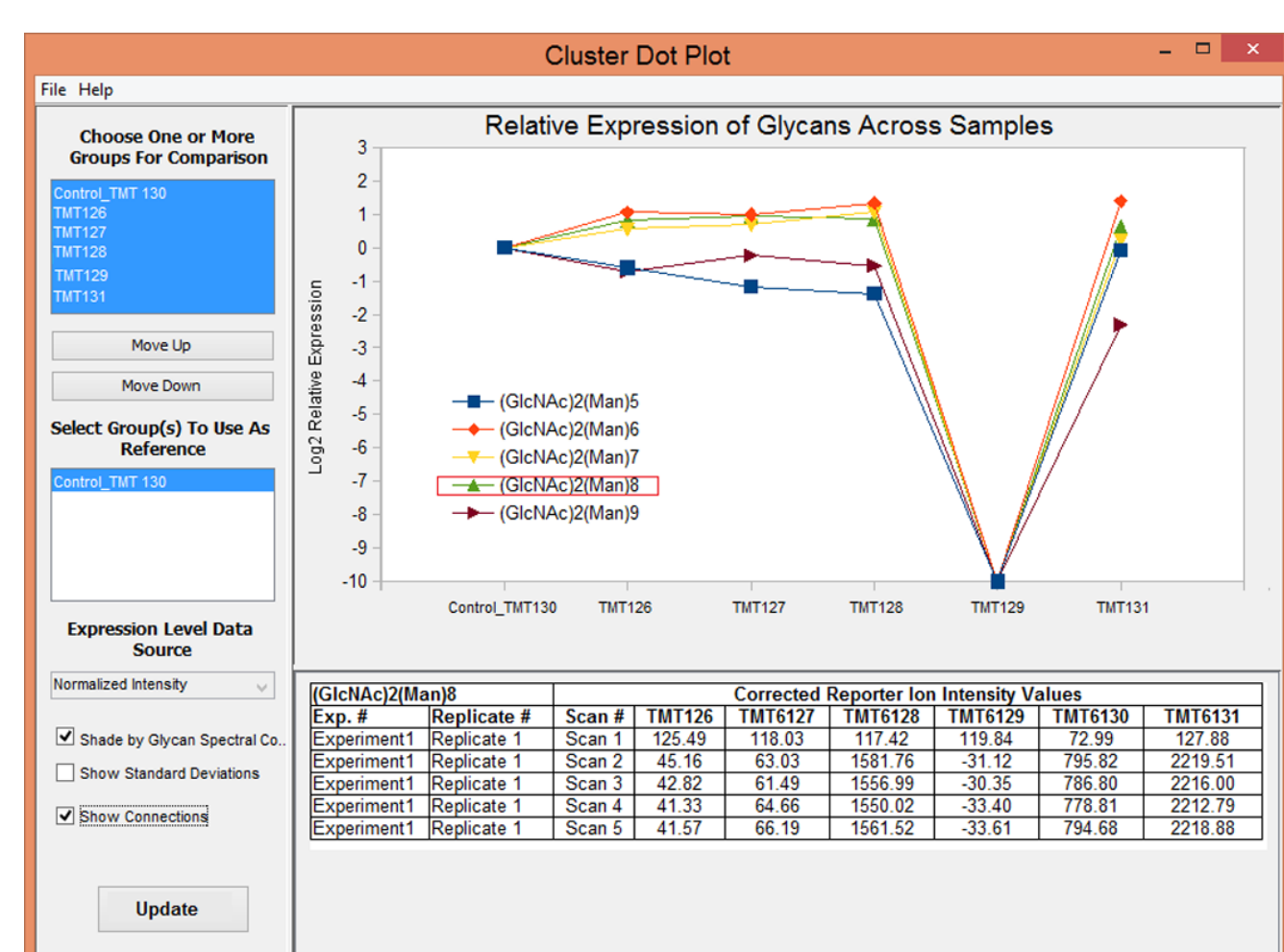


Figure 14. コントロールで正規化した各TMTチャンネルのfold-changeを表示する *SimGlycan*® のDot Plot

➤ Report Generation

ユーザーは、同定した糖鎖構造をカスタマイズ可能なMS Excel形式でレポートを出力して、共同研究者と分析結果を簡単に共有することができる。Users can generate a customizable MS excel report with identified glycan structures, for easy sharing of their findings.

Figure 15. *SimGlycan*® によるMS Excelベースのレポート出力

Conclusion

SimGlycan® Enterprise Edition 5.4のGUIで、簡単にカスタムデータベースを作成することができる。*SimGlycan*® Enterprise Edition 5.4は、シンプルなバッチ設定 (10,000 MS/MSスキャンまで) で糖鎖の構造同定を行うことができる。複数のバッチ検索も同時処理することができる。さらに、糖鎖構造を含む結果は、情報共有しやすい形式のMS Excelファイルで出力されるので、構造同定後の更なるデータ解析を行いやすい。*SimGlycan*® Enterprise Edition 5.4では、Thermo Fisher Scientific社のaminooxyTMT Reagentsを用いて定量分析を簡単に行うことができる。Bar Chart, Dot Plotといった直感的でインタラクティブなプロットによって、各TMTレポーターイオンチャンネルにある糖鎖の比較解析や差分解析を容易に行うことができる。

References

- Bigge, J. C., Patel, T. P., Bruce, J. A., Goulding, P. N., Charles, S. M., & Parekh, R. B. (1995). Nonspecific and efficient fluorescent labeling of glycans using 2-amino benzamide and anthrnic acid. *Analytical biochemistry*, 230(2), 229-238.
- Chai, W., Piskarev, V., & Lawson, A. M. (2001). Negative-ion electrospray mass spectrometry of neutral underivatized oligosaccharides. *Analytical chemistry*, 73(3), 651-657.
- Harvey, D. J. (2005). Fragmentation of negative ions from carbohydrates: part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from N-linked carbohydrates. *Journal of the American Society for Mass Spectrometry*, 16(5), 622-630.
- Meitei, N. S., Apte, A., Snovida, S. I., Rogers, J. C., & Saba, J. (2015). Automating Mass Spectrometry-Based Quantitative Glycomics using Aminooxy Tandem Mass Tag Reagents with *SimGlycan*®. *Journal of proteomics*. doi:10.1016/j.jpro.2015.05.015
- Apte, A., Meitei, N. S. (2010) Bioinformatics in glycomics: Glycan characterization with mass spectrometry data using *SimGlycan*®. In *Functional Glycomics*, Humana Press; p. 269-81.