

Fukozilált humán glikoproteinek szerkezetvizsgálata tandem tömegspektrometriával

Doktori tézisek

Ács András

Semmelweis Egyetem
Gyógyszertudományok Doktori Iskola



Témavezető: Dr. Vékey Károly, D.Sc., professor emeritus
Dr. Telekes András, C.Sc., címzetes egyetemi tanár

Hivatalos bírálók: Dr. Alberti-Dér Ágnes, Ph.D., egyetemi docens
Dr. Takátsy Anikó, Ph.D., egyetemi docens

Szigorlati bizottság elnöke: Dr. Zelkó Romána, D.Sc., egyetemi tanár

Szigorlati bizottság tagjai: Dr. Háda Viktor, Ph.D., csoportvezető
Dr. Riba Pál, Ph.D., egyetemi docens

Budapest
2019

I. Bevezetés

Az elmúlt évtizedek kutatómunkája rávilágított, hogy a gének által kódolt fehérjék a biológiai funkciók sokszínűségét talán még jobban reprezentálják, mint maguk a gének. A fehérjék szerkezetének, funkciójának tanulmányozásával és mennyiségi meghatározásával foglalkozó tudományág a proteomika. A humán genomot körülbelül 25.000 gén alkotja, míg a fehérjék számát több mint 1.000.000-ra becsülik. A fehérjék szerkezete és funkciója még tovább módosulhat különböző poszt-transzlációs módosítások hatására. Jelenleg több mint 200 különböző poszt-transzlációs módosítást ismerünk, ezek közül az egyik leggyakoribb típus a glikoziláció. Fehérjék szerkezetvizsgálatának fontos eleme a glikoziláció jellemzése. A modern technológiai fejlesztések lehetővé teszik az adott fehérjéhez kapcsolódó oligoszacharid láncok szerkezetének és pozíciójának pontos felderítését. Kutatómunkám során a fukóz szubsztituens elhelyezkedését vizsgáltam *N*-glikozilált fehérjéken.

Egy fehérje glikozilációs mintázatának vizsgálatára számos, széles körben elterjedt módszer létezik. Ezen technikákat rendszeresen alkalmazzák biológiai eredetű gyógyszerkészítmények minőségellenőrzésére és a biomarker kutatásában is. A fukozilált glikoformok mennyiségi változását számos biológiai folyamatban leírták, de a fukóz pozícióját csak néhány esetben határozták meg. Arról is csak nagyon kevés információ található az irodalomban, hogy a fukóz mindig ugyanazon helyen található, vagy ez változhat biológiai folyamatok során. Ennek fő oka, hogy a fukóz helyének meghatározása nem triviális, nincsenek rutinszerűen alkalmazott módszerek. A fukóz pozíciója alapján megkülönböztethetünk core- és antenna-fukozilációt. Munkám során core- és antenna-fukozilált modellvegyületek energiafüggő fragmentációját vizsgáltam tömegspektrometriával. A core-fukoziláció azonosításához példavegyületnek a PSA (Prosztata Specifikus Antigén) fehérjét választottam, mely az eddigi ismeretek szerint kizárólag core-fukozilált glikán-oldalláncokat tartalmaz. Az antenna-fukoziláció jellemzésére az AGP-t (α 1-savas glikoprotein) választottam, mely az eddigi ismeretek szerint csak az antennán elhelyezkedő fukóz szubsztituens tartalmaz. A modellvegyületeket enzimatiszta elemésztettem, és az így keletkező glikopeptideket vizsgáltam tovább. Az egyes fragmentációs folyamatokat letörési görbékkel (breakdown

curves) és egyenletekkel jellemeztem. A spektrumok elemzésével sikerült mindkét típusú fukoziláció esetében olyan ún. diagnosztikus fragmenseket azonosítani, melyek jelenlétéből és egymáshoz viszonyított arányából egyértelműen következtetni lehet a fukóz pozíciójára. A kidolgozott módszer segítségével azonosítottam egy ritka, bifukozilált glikoform szerkezetét. A módszer könnyedén beilleszthető az elterjedt proteomikai munkafolyamatokba, nem igényel plusz mintaelőkészítést. Emellett a technika gyakorlatilag bármely fehérje esetében alkalmazható.

II. Célkitűzések

A proteomika fejlődésével és a terápiás célú monoklonális antitestek elterjedésével párhuzamosan glikoproteinek vizsgálata egyre nagyobb jelentőségre tett szert az elmúlt évtizedben. A cukor-oldalláncok változatosságának és komplexitásának megismerésével megjelent az igény olyan minőségi paraméterek meghatározására, melyekre jelenleg nincsenek rutinszerűen alkalmazott vizsgálati módszerek. Az egyik ilyen paraméter a core- és antenna-fukoziláció elkülönítése.

1. Munkám során célul tűztem ki egy olyan tömegspektrometrián alapuló módszer kifejlesztését, mely egyértelműen megkülönbözteti a core- és antenna-fukozilációt. A vizsgálatokhoz modellvegyületként az eddigi irodalmi ismeretek alapján a tisztán core-fukozilált PSA és a tisztán antenna-fukozilált AGP glikoproteineket választottam.

2. Célom volt ezen proteinekből glikopeptidek előállítása, majd ezek részletes, tandem tömegspektrometrián alapuló jellemzése. Az első lépés néhány olyan glikopeptid kiválasztása volt, melyek összetételük és intenzitásuk alapján várhatóan jól vizsgálhatók, és alkalmasak a core- és antenna fukoziláció tanulmányozására. Mivel a kiválasztott glikopeptidek bonyolult összetételű fehérjeemésztményben vannak, vizsgálatukat nano-HPLC-MS/MS módszerrel végeztem.

3. Cél volt a kiválasztott glikopeptidek energiafüggő MS/MS jellemzése. Az ütközési energiát a proteomikában használt ütközési energia 30% és 150%-a közötti tartományban vizsgáltam.

4. Az energiafüggés értelmezésére a letörési görbe (breakdown curve) módszert alkalmaztam. Célul tűztem ki olyan diagnosztikus fragmensionok és olyan energiaintervallum meghatározását, mely jól alkalmazható a core- és az antenna-fukóz megkülönböztetésére.

5. További célom volt a modellfehérjék vizsgálatával nyert ismeretek alapján ritka, többszörösen fukozilált AGP glikoformok szerkezetének meghatározása.

III. Módszerek

III.1. Enzimatis emésztés

A liofilizált állapotban tárolt AGP és PSA standardokból 1 nmol anyagmennyiséget 30 μ l víz + 5% metanolban oldottam fel és használtam az emésztéshez. Az enzimatis emésztéshez a kutatócsoportunk által korábban, kis mennyiségű fehérjék emésztéséhez kifejlesztett protokollt alkalmaztam. A fehérjék szolubilizálásához és a diszulfidhidak redukálásához a mintákat 5 μ l 0,5% Rapigest-tel és 2 μ l 200 mM DTT-vel (1,4-ditiotreitol) inkubáltam 60 °C-on, 30 percig. Ezután a diszulfidhidak újbóli kialakulásának megakadályozására és az emésztéshez használt enzimek számára optimális 7,8 pH kialakítására 2,5 μ l 200 mM IAA-t (jódacetamid) és 5 μ l 200 mM NH_4HCO_3 -ot (ammónium-bikarbonát) adtam hozzá a mintákhoz és fénytől elzárt helyen, szobahőmérsékleten 30 percig inkubáltam. A tripszin enzim önmagában alkalmazva a lizin aminosav mellett gyakran nem hasítja el a peptidláncot (ún. missed cleavage). Ennek kiküszöbölésére előemésztésként 1:100 enzim:fehérje arányban, 1 μ l térfogatban Lys-C és tripszin enzim keveréket adtam a mintákhoz (Lys-C/Tripszin), és 1 órán keresztül 37 °C-on inkubáltam. A Lys-C a tripszinhez viszonyítva nagyobb hatékonysággal hasít a lizin mellett, így használata javítja az emésztés hatékonyságát. Ezt követően 1:25 enzim:fehérje arányban, 1 μ l térfogatban hozzáadtam a tripszint és további 2 órán át 37 °C-on inkubáltam. Az enzimek inaktiválásához, így az emésztés leállításához 1 μ l FA-t (hangyasav) használtam. A PSA standard esetén a Lys-C/Tripszin és tripszines emésztés egy két aminosavból álló (NK, monoizotópos tömeg 260,148) glikopeptidet eredményezett, amely kis tömege miatt nem volt alkalmas a további vizsgálatokra. Az Arg-C enzimmel való emésztés során egy 9 aminosav hosszúságú glikopeptid (NKSVILLGR, monoizotópos tömeg 998,624) keletkezett, mely jól vizsgálható a további kísérletekben. Ezért a PSA standard esetén a fentiekben említett protokoll úgy módosult, hogy Lys-C/Tripszin és tripszin helyett 1:25 enzim:fehérje arányban, 1 μ l térfogatban Arg-C enzimmel emésztettem. Az Arg-C optimális működéséhez 1 μ l 200 mM DTT-t és 1 μ l 15 mM CaCl_2 -ot (kalcium-klorid) adtam hozzá a mintához az enzim hozzáadása előtt és 10 percig szobahőmérsékleten inkubáltam.

III.2. Nano LC-MS/MS

A folyadékkromatográfiás elválasztás egy Ultimate 3000 nanoRSLC system (Dionex, Sunnyvale, CA, Egyesült Államok) típusú készüléken történt. A minták sómentesítéséhez egy Acclaim PepMap100 C-18 csapdázó oszlopot (trap column, 100 μm \times 20 mm; Thermo Scientific, Sunnyvale, CA, Egyesült Államok), a peptidek elválasztásához pedig egy Acquity UPLC M-Class Peptide BEH C18 oszlopot (1,7 μm , 130 \AA , 75 μm \times 250 mm; Waters, Milford, MA, Egyesült Államok) használtam. Az elválasztás során 48 $^{\circ}\text{C}$ -on, 300 nl/perc áramlási sebességen 60 vagy 90 perces gradiens elúciót alkalmaztam, az A eluens összetétele víz + 0,1% FA, a B eluens összetétele pedig acetonitril + 0,1% FA volt. A folyadékkromatográfiás rendszert egy CaptiveSpray nanoBooster ionforrással szerelt Maxis II ETD Q-TOF (Bruker Daltonics, Bremen, Németország) tömegspektrométerhez csatlakoztattam.

III.3. Glikozilációs analízis

Az MS1 spektrumok a 150-3000 m/z tömegtartományban kerültek felvételre 5 Hz frekvencián. A kiválasztott prekursor ionokról ütközés indukált disszociáció (CID) segítségével MS/MS felvételek készültek, az intenzív ionokról (>25000 beütés/másodperc) 4 Hz, a kis intenzitású ionokról (>5000 beütés/másodperc) 1 Hz frekvencián. A tetraantennás AGP glikopeptidek csak alacsony intenzitással jelennek meg a spektrumban, ezért e mérések esetében az MS1 felvételek frekvenciáját 3 Hz-re, az intenzív ionokról készült MS/MS felvételek esetén pedig 1 Hz-re állítottam. A "standard ütközési energia" a készülégyártó ajánlásainak megfelelően, az izolációs tömegtartomány szélesség, az izolációs m/z és az ion töltésszáma alapján lett megállapítva. Az energiafüggés vizsgálatakor ennek az energiának a meghatározott százalékával fragmentáltam az adott prekursor iont, a 30%-150%-os tartományban. A nyers mérési adatokat a Compass DataAnalysis 4.3 (Bruker Daltonics, Bremen, Németország) szoftverrel értékeltem. A program segítségével meghatároztam a prekursor ion és a fragmentumok intenzitását. Az adott ion intenzitásaként mindig a legintenzívebb izotópcsúcsot vettem figyelembe.

IV. Eredmények

(1) Kidolgoztam egy tandem tömegspektrometrián alapuló módszert, mely egyértelműen elkülöníti a glikopeptideken található core- és antenna-fukozilációt.

Ennek első lépéseként kiválasztottam a core- és antenna-fukoziláció jellemzéséhez optimális glikopeptideket, melyek nagy intenzitású komponensek, tömegspektrometriával megbízhatóan vizsgálhatóak. Korábbi eredményeink és irodalmi adatok alapján a kiválasztott glikopeptidek kizárólag core- vagy antenna-fukoziláltak. A core-fukoziláció jellemzésére a PSA-eredetű NKSVILLGR-N4H5S2F1 glikopeptidet, az antenna-fukoziláció jellemzésére az AGP-eredetű NEEYNK-N4H5S2F1 glikopeptidet választottam. A kiválasztott glikopeptidek energiafüggő MS/MS spektrumainak vizsgálatával meghatároztam a fragmensionok összetételét.

(2) Diagnosztikus fragmensionokat azonosítottam az MS/MS spektrumokban, melyek jellemzik a core-, ill. az antenna-fukozilációt.

Az antenna-fukoziláció jelenlétét egyértelműen igazolja az alacsony ütközési energián jól észlelhető NHSF⁺ ion. Emellett megfigyelhető a komplementer [M-NHSF]²⁺ ion is. Alacsony ütközési energián a core-fukoziláció jelenlétére indirekt módon, az NHSF⁺ és [M-NHSF]²⁺ ionok hiányából lehet következtetni. A fukóz elhelyezkedésének azonosítása magasabb ütközési energiákon is lehetséges. Ilyen esetben a szekvenciális fragmentációs folyamatokban keletkezett [M-NHS-NHSF]⁺/[M-NHS-NHS]⁺ vagy [peptid+NF]⁺/[peptid+N]⁺ ionpárok arányait célszerű megvizsgálni.

(3) Az antenna-fukoziláció mértékének jellemzésére a diagnosztikus fragmensionok egymáshoz viszonyított arányán alapuló egyenleteket dolgoztam ki.

Ezen egyenletek alkalmasak az antenna-fukoziláció mértékének számszerűsítésére. Meghatároztam ezen egyenletek segítségével számolt értékek energiafüggését. Az antenna-fukoziláció mértékének megbecsülésére az F_I értéket javaslom ($F_I = \frac{[M-NHSF]^{2+}}{[M-NHS]^{2+} + [M-NHSF]^{2+}}$). Fontos megjegyezni azonban, hogy core-fukóz esetén 2%, antenna-fukóz esetén mintegy 14% mértékű interfencia várható még kis

ütközési energián is, részben fukóz vándorlás, részben pedig többlépcsős fragmentációs folyamatok eredményeképp.

(4) A fragmentációs folyamatok energiafüggését letörési görbe (breakdown curve) segítségével jellemeztem.

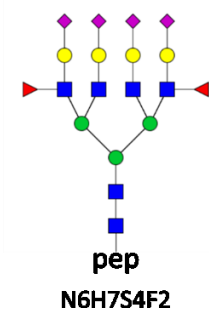
A letörési görbék segítségével jól szemléltethető az egyes fragmentációs mechanizmusok energiafüggése. Megfigyeltem, hogy a semleges fukóz veszteség az egylépcsős, töltésszeparációs folyamatokhoz képest alacsonyabb aktiválási energiát igényel. A core-fukóz erősen kötődik a core-szerkezethez. Ezzel szemben az antenna-fukóz lazán kötött, fragmentációra és vándorlásra hajlamos. Ez utóbbi jelenség az eredmények értékelését nehezíti.

(5) Meghatároztam az adott fukóz-pozíció jellemzésére optimális ütközési energiát.

Az optimális ütközési energiát minden vizsgált glikopeptid esetében egyedileg kell meghatározni. Olyan spektrum elérésére kell törekedni, ahol a survival yield értéke több, mint 50%. Ilyenkor többségében egylépcsős fragmentációs folyamatok figyelhetők meg, melyek könnyen értelmezhetőek. Emellett a diagnosztikus ionok intenzívek, megbízhatóan azonosíthatóak. A fukóz vándorlás mértéke alacsony, a kiértékeléskor nem jelent problémát.

(6) Ritka bifukozilált glikoformot azonosítottam az AGP fehérje A1AG2_HUMAN genetikai variánsában.

A 3+ töltésű, ENGTVSR-N6H7S4F2 glikopeptid ($m/z= 1524,577$) egy tetraantennás, minden antennán szializált, két fukóz szubsztituens tartalmazó szerkezetet jelöl. Meghatároztam, hogy mindkét fukóz szubsztituens az antennán helyezkedik el. A vizsgált glikoform szerkezetét az 1. ábra mutatja be.



1. ábra A tetraantennás ENGTVSR-N6H7S4F2 glikopeptid szerkezete

V. Következtetések

Vizsgálataim során az *N*-glikoziláció egy fontos jellemzője, a fukoziláció szerkezeti jellemzésével és glikoformok szerkezetazonosításával foglalkoztam. A fukoziláció jelentőségét sokféle biológiai és élettani folyamatban leírták, jelentősége a biológikumok elterjedésével a gyógyszeriparban is megnövekedett. A fukóz pozíciójának meghatározására jelenleg nincsen széles körben elfogadott analitikai módszer. Az értekezésben egy olyan technikát mutatok be, ahol alacsony ütközési energián felvett MS/MS spektrumokból a fukóz pozíciója megállapítható és mennyisége kvantifikálható. A gyakori core-fukoziláció mellett már kis mennyiségű antenna-fukóz is jól kimutatható. A rutinszerűen, magas ütközési energián felvett spektrumokból is levonhatók kvalitatív következtetések, illetve a különböző minták még magas ütközési energián felvett spektrumok alapján is összehasonlíthatók. A módszer nagy előnye, hogy nem igényel származékképzést, és a hagyományos proteomikai munkafolyamatba könnyen beilleszthető. A módszerfejlesztés során *N*-glikopeptideket vizsgáltam, azonban nincs elméleti akadálya a módszer kiterjesztésének *O*-glikopeptidek analízisére, ez további kutatási terveink közt szerepel. A korábbi szakirodalmi adatok alapján ismert, hogy a fukóz elhelyezkedésének azonosítását megzavarhatja a fukóz vándorlás folyamata. Az értekezésben bemutatom, hogy a fukóz vándorlás ténylegesen jelen van, de az ütközési energia csökkentésével mértéke jelentősen lecsökken és a gyakorlatban nem jelent problémát. A semleges fukóz veszteséssel járó szekvenciális folyamatok megnehezítik az értékelést, azonban az ütközési energia csökkentésével e folyamatok jelentősége is csökken. Alacsony energiájú CID fragmentáció esetén, mikor a molekulaion a legintenzívebb csúcs a spektrumban (pl. a survival yield 50% feletti), a fukóz vándorlás és a többi zavaró folyamat minimálisra csökkenthető. A fragmentációs vizsgálatokhoz a 3+ töltésű glikopeptidek használatát javaslom, mivel ezek spektruma a leginformatívabb és az interferenciák előfordulása is kevésbé valószínű. Lényeges megemlíteni, hogy a jelen vizsgálatokat egy kvadрупól és repülési idő analizátorral szerelt tömegspektrométeren (Q-TOF) végeztem. Ioncsapda típusú analizátorral rendelkező készüléken az analizátorban töltött hosszabb tartózkodási idő miatt a fragmentációs mechanizmusok eltérőek lehetnek, a fukóz vándorlás és az interferenciák előfordulásának

esélye megnő. Fontos kiemelni, hogy az alacsony ütközési energiájú CID használata nem csupán a fukóz vándorlás minimalizálása miatt előnyös, az így felvett MS/MS spektrumok rendkívül tiszták, nincs kémiai zaj és a diagnosztikus ionok a spektrum legintenzívebb csúcsai közé tartoznak. Alacsony ütközési energián, a többszörösen fukozilált glikoformok is jól azonosíthatóak és meghatározható a fukóz szubsztituensek elhelyezkedése. A glikán struktúrák részletes jellemzése hozzájárul a biológiai eredetű gyógyszerkészítmények minőségellenőrzésének fejlesztéséhez és a biomarker kutatást fontos új eszköze lehet.

VI. Saját publikációk jegyzéke

VI.1. Az értekezés témájában megjelent közlemények

1. Ács A, Ozohanics O, Vékey K, Drahos L, Turiák L. (2018) Distinguishing Core and Antenna Fucosylated Glycopeptides Based on Low-Energy Tandem Mass Spectra. *Anal Chem*, 90: 12776-12782.

IF: 6,350

2. Ács A, Turiák L, Révész Á, Vékey K, Drahos L. (2019) Identification of bifucosylated glycoforms using low-energy CID spectra. *J. Mass Spectrom.*

IF: 2,267

VI.2. Egyéb közlemények

1. Turiak L, Toth G, Ozohanics O, Revesz A, Acs A, Vekey K, Zaia J, Drahos L. (2018) Sensitive method for glycosaminoglycan analysis of tissue sections. *J Chromatogr A*, 1544: 41-48.

IF: 3,858

2. Gogl G, Biri-Kovacs B, Poti AL, Vadaszi H, Szeder B, Bodor A, Schlosser G, Acs A, Turiak L, Buday L, Alexa A, Nyitray L, Remenyi A. (2018) Dynamic control of RSK complexes by phosphoswitch-based regulation. *Febs J*, 285: 46-71.

IF: 4,739

3. Kovacs AF, Lang O, Turiak L, Acs A, Kohidai L, Fekete N, Alasztics B, Meszaros T, Buzas EI, Rigo J, Jr., Pallinger E. (2018) The impact of circulating preeclampsia-associated extracellular vesicles on the migratory activity and phenotype of THP-1 monocytic cells. *Sci Rep*, 8: 5426.

IF: 4,011

4. Turiak L, Ozohanics O, Toth G, Acs A, Revesz A, Vekey K, Telekes A, Drahos L. (2019) High sensitivity proteomics of prostate cancer tissue microarrays to discriminate between healthy and cancerous tissue. *J Proteomics*, 197: 82-91.

IF: 3,537

5. Kovács ÁF, Fekete N, Turiák L, Ács A, Kőhidai L, Buzás EI, Pállinger É. (2019) Unravelling the Role of Trophoblastic-Derived Extracellular Vesicles in Regulatory T Cell Differentiation. *Int. J. Mol. Sci.*, 20: 3457.

IF: 4,183

6. Turiak L, Sugar S, Acs A, Toth G, Gomory A, Telekes A, Vekey K, Drahos L. (2019) Site-specific N-glycosylation of HeLa cell glycoproteins. *Sci Rep*, 9: 14822.

IF: 4,011

7. Müller A, Langó T, Turiák L, Ács A, Várady G, Kucsma N, Drahos L, Tusnády GE. (2019) Covalently modified carboxyl side chains on cell surface leads to a novel method toward topology analysis of transmembrane proteins. *Sci Rep*, 9: 15729.

IF: 4,011