Analytical characterization of NOTA-modified somatropins

Nathalie Bracke[a], Evelien Wynendaele[a], Matthias D'Hondt[a], Rob Haselberg[b], Govert W. Somsen[b], Ewald Pauwels[c], Christoph Van de Wiele[d], Bart De Spiegeleer[a1]

[a] Drug Quality and Registration (DruQuaR) group, Department of Pharmaceutical Analysis, Faculty of Pharmaceutical Sciences, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium

[b] Division of BioAnalytical Chemistry, AIMMS research group BioMolecular Analysis, Faculty of Sciences, VU University, De Boelelaan 1083, 1081 HV Amsterdam, the Netherlands

[c] Center for Molecular Modeling, Ghent University, Technologiepark 903, 9052 Zwijnaarde, Belgium

[d] Department of Nuclear Medicine, Ghent University Hospital, 9000 Ghent, Belgium

[1] To whom correspondence should be addressed: Bart De Spiegeleer

Tel: +32 9 264 81 00; Fax: +32 9 264 81 93; Email: Bart.DeSpiegeleer@UGent.be

(Our reference: 2014-034b)

Highlights

- NOTA-labeled somatropins were synthesized using different NOTA:protein ratios.
- Direct LC–MS and CE–MS approaches indicated multiple substitution degrees.
- Bottom-up approaches gave structural insights and information on the labeling yield.
- Lys-70 (*in silico* calculated p*K*a of 8.3) is the NOTA-modification hotspot.
- We report a synthesis procedure for the production of target-specific radiopharmaceuticals.

Abstract

Chemical modification of biomolecules like the introduction of metal-chelators into proteins can lead to heterogeneous product formation. The nature and extend of the modification is important in interpreting the biological properties of the bioconjugate, given their possible influence on the pharmacokinetics as well as on the binding affinity to the target. The present study describes the synthesis and analytical characterization of somatropin modified on its lysine's ε-amino groups with the acylating chelator *S*-2-(4-isothiocyanatobenzyl)-1,4,7-triazacyclononane-1,4,7-triacetic acid (*p*-SCN-Bn-NOTA). Direct separation and identification techniques (*i.e.* RP-MS and CE-MS) and peptide mapping after trypsin and chymotrypsin digestion demonstrated that the use of higher amounts of *p*-SCN-Bn-NOTA during synthesis leads to a complex product composition with higher order substitution degrees (*i.e.* multiple NOTA-moieties per somatropin molecule), as well as the presence of different position isomers. From the nine lysine (Lys) residues in somatropin, Lys-70 was experimentally found to be the modification hotspot under our synthesis conditions (pH=9.0). This was supported by the *in silico* calculated lowest pKa value of 8.3 for Lys-70. Based on the crystal structure of somatropin in complex with the extracellular parts of the growth hormone receptor, the Lys-70 residue is positioned outside the binding pockets and will therefore not directly interfere with receptor binding. Gallium chelation by NOTA-somatropin resulted in a 100% complexation. The synthesis of NOTA-somatropin using *p*-SCN-Bn-NOTA and somatropin under our operational conditions is therefore a suitable synthesis procedure for the production of a target-specific radiopharmaceutical for further investigation towards treatment and visualization of growth hormone-specific cancers.

Keywords

NOTA modification, somatropin, LC-MS, CE-MS, peptide mapping

1. INTRODUCTION

The recent successes of biopharmaceuticals are changing the focus of the pharmaceutical industry. The US Food and Drug Administration (FDA)'s Center for Drug Evaluation and Research (CDER) approved 39 new drugs in 2012, whereof six represent biologics license applications (BLAs) [1]. By 2015, it is even expected that 50% of the newly approved drugs will be biologics [2], which further illustrates the importance of the development of new biological entities (NBEs) and bioanalytical tools for the characterization of NBEs and biosimilars towards registration [3, 4].

Cancer is still the largest therapeutic area nowadays [1, 5]. Tumor targeting in nuclear medicine is based on a target-specific radiopharmaceutical ligand for selective receptor binding in the disease tissue [6-8]. The targeting biomolecules can be used as therapeutics to deliver a toxic radioactive payload selectively to a tumor site (*e.g.* I-131 tositumomab or Y-90 ibritumomab), as well as diagnostic agents for non-invasive single photon emission computed tomography (SPECT) or positron emission tomography (PET) imaging (*e.g.* In-111 Capromab pendetide or In-111 pentetreotide) [9, 10] and playing an important role in cancer management as a form of personalized medicine [9, 11].

One of the biomolecules of current interest is human growth hormone (hGH). In the late 50s, severe growth hormone deficiencies (GHD) in children were treated by cadaveric pituitary hGH or somatotropin extracts. However, several cases were reported which correlated the use of cadaveric somatotropin with Creutzfeldt-Jakob disease, leading to an abrupt stop of hGH extract treatment [12]. This led to the worldwide regulatory approval of recombinant hGH. Recombinant hGH or somatropin is nowadays used for the treatment of GHD, as well as the treatment of Turner Syndrome and AIDS associated catabolism [13]. Recently, biosimilar somatropin formulations were also globally approved [14-16]. Moreover, somatropin and analogues are often encountered as spurious/falsely-labelled/falsified/counterfeit (SFFC) medicines [17-19] and drug abuses in sports [20, 21], agriculture [22] and anti-aging [23] have been reported.

Somatropin can perform its actions by binding with high affinity to the extracellular domains of two identical molecules of human growth hormone receptors (hGHR) [24-27], which are widely expressed in liver tissue, but are also aberrantly overexpressed in numerous cancers such as prostate [28, 29], breast [30] and colon cancer [31]. The potential involvement of the GH system in tumor promotion

and progression, which has been critically reviewed in references [32] and [33], as well as the internalization of the receptor-ligand complex [34], makes hGHR a potential tumor target for the development of somatropin-based radiopharmaceuticals.

Modifications with bifunctional chelating agents (BFCA) like *S*-2-(4-isothiocyanatobenzyl)-1,4,7-triazacyclononane-1,4,7-triacetic acid (*p*-SCN-Bn-NOTA) allow the incorporation of radiometals for SPECT/PET-diagnostic (^{67}Ga, ^{68}Ga, ^{111}In) or therapeutic (^{90}Y) purposes [9, 35, 36]. However, because somatropin has nine potential lysine-amino binding sites for *p*-SCN-Bn-NOTA, it is important to characterize the obtained product under different synthesis procedures. We present the analytical characterization of NOTA-modified somatropins with special attention to the operational strategy and procedure which are widely applicable to other protein systems containing multiple modification sites towards other bifunctional chelators (*e.g.* DOTA, DTPA, MAMA) or fluorescent labels (*e.g.* fluorescein) [9].

2. MATERIALS AND METHODS

2.1 Materials and equipment

The *p*-SCN-Bn-NOTA was purchased from Macrocyclics Inc. (Dallas, TX, USA). Zomacton® 4 mg, (Ferring, somatropin Ph. Eur.) was purchased at the Ghent University Hospital (Ghent, Belgium). The enzymes for peptide mapping, L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin solution and immobilized chymotrypsin solution, were purchased at Pierce (Erembodegem, Belgium) and Sigma Aldrich (Diegem, Belgium), respectively. PD-10 sephadex G-25M columns were obtained from GE healthcare (Diegem, Belgium). Water was purified in-house using an Arium 611 purification system (Sartorius, Göttingen, Germany), yielding ≥ 18.2 MΩ × cm quality water. Other chemicals and solvents were purchased from Merck (Overijse, Belgium), Sigma Aldrich (Diegem, Belgium) or Fischer Scientific (Erembodegem, Belgium), all high quality (>98% purity) and/or HPLC/MS grade.

Freeze-drying was done using a Christ gamma 1-16 LSC freeze dryer (Qlab, Vilvoorde, Belgium). The HPLC-UV-MS apparatus consisted of a SpectraSystem separations module, a Finnigan LCQ Classic ion trap mass spectrometer in positive ion mode (all Thermo, San José, CA, USA) equipped

with a Waters 2487 dual wavelength absorbance UV detector (Waters, Milford, MA, USA) and Xcalibur 2.0 software (Thermo, San José, CA, USA) for data acquisition. For CE-MS, a P/ACE MDQ capillary electrophoresis instrument (Beckman Coulter Inc., Brea, CA, USA) was used for separations, whereas a micrOTOFQ orthogonal accelerated time-of-flight (TOF) mass spectrometer (Bruker Daltonics, Bremen, Germany) was used for detection and identification.

2.2 Synthesis

1 mg of somatropin (eq. to 45 nmol) was dissolved in 200 µL carbonate buffer (pH 9.0; 0.1 M), added to different volumes of 20 mM *p*-SCN-Bn-NOTA in carbonate buffer (pH 9.0; 0.1 M) and diluted to 350 µL with carbonate buffer (pH 9.0; 0.1 M) to obtain molar equivalents of 1:1 NOTA:somatropin, 3:1 NOTA:somatropin and 10:1 NOTA:somatropin. The mixtures were incubated for 20 hours at 20°C in the dark, while shaking at 300 rpm. The sample was loaded onto a PD-10 sephadex G-25M column (equilibrated with digestion buffer consisting of 10 mM ACES, 20 mM $CaCl_2$, pH 7.0). After collection of the protein fraction, the samples were freeze dried awaiting analytical characterization.

2.3 Direct analysis of NOTA-somatropins

NOTA-labeling of the 3:1 sample was monitored during the 20 hours incubation period, by taking 30 µL of sample after 1, 3, 5, 8, 10, 12.5, 17, 20 and 22 hours. The samples were diluted to 100 µL using carbonate buffer (pH 9.0; 0.1 M) and characterized using LC-MS. For the analysis of the different NOTA-somatropins, 0.1 mg of lyophilized 1:1 NOTA:somatropin, 3:1 NOTA:somatropin and 10:1 NOTA:somatropin (4.5 nmol protein) was dissolved in 100 µl water. Samples were analyzed via LC-MS. A Vydac Everest MS RP-C_4 (250 mm × 4.6 mm I.D., 5 µm particle size, 300 Å) column (Grace Vydac, Hesperia, CA, USA) was used and temperature controlled during analysis (35°C). The injection volume was 20 µL. The flow rate was set to 0.5 mL/min and the following gradient was used for separation of different somatropin derivatives (A: 50 mM ammonium bicarbonate pH 7.5 and B: 1-propanol): for 60 minutes, a linear gradient was applied from 70% A (v/v) + 30% B (v/v) to 50% A (v/v) + 50% B (v/v), followed by a 10 min isocratic section (50:50 A:B (v/v)). The method also included a rinsing step out of 60% 1-propanol, followed by returning to the initial conditions and re-

equilibration. ESI was conducted with a capillary voltage of 4.5 kV. Nitrogen was used as the sheath and auxiliary gas; the temperature of the heated capillary was set to 280°C. MS/MS spectra were used for identification and obtained by collision induced dissociation (CID) of the parent m/z, with the relative collision energy set to 35%. UV detection and quantification were performed at 220 nm.

The CE-MS analysis was performed using fused-silica capillaries (800 mm × 50 µm I.D., Polymicro Technologies, Phoenix, AZ, USA) coated with a bilayer of Polybrene and poly (vinyl sulfonic acid) as described previously [37]. Sample injections were performed at 1 psi for 12 s. The separation voltage was 30 kV and the capillary temperature was 20 °C. The background electrolyte (BGE) was 75 mM ammonium hydroxide adjusted to pH 8.5 with 1% (v/v) formic acid in deionized water. As sheath liquid, a mixture of isopropanol-water-acetic acid (75/22/3 v/v/v) was employed at 4 µL/min. ESI was conducted in positive ionization mode with a capillary voltage of 4.5 kV. To assure proper ion transfer, the analysis of somatropin was performed with a capillary exit and skimmer voltage of 250 and 83 V, respectively. CE-MS data were analyzed using Bruker Daltonics Data Analysis software. Molecular weight determinations of proteins were performed using the "Charge Deconvolution" utility of the Data Analysis software. Quantification of the conjugates was performed using the construction of an extracted-ion electropherogram. Each detected conjugate showed the same charge state distribution profile, with $(M+9H)^{9+}$ and $(M+10H)^{10+}$ as most abundant ions. The relative abundance of each conjugate in the analyzed preparation was established by calculating the ratio of the conjugate peak area to the total peak area.

2.4 Peptide mapping

The unmodified somatropin (*i.e.* control) and NOTA-modified somatropins (45 nmol protein) were dissolved in 1000 µL 6 M guanidine HCl, 35 mM tris, 20 mM DL-dithiothreitol at pH 7.5 and incubated for 30 min at 37°C, while shaking at 300 rpm. S-carboxymethylation of cysteine residues was performed by addition of 20 µL of iodoacetate (pH 7.2; 58 mM) and subsequent incubation for 30 min at 37°C (at 300 rpm). 100 µL of DL-dithiothreitol (1 M) was added and mixed. The sample was loaded onto a PD-10 column sephadex G-25M (equilibrated with digestion buffer consisting of 10 mM ACES, 20 mM $CaCl_2$, pH 7.0). After collection of the protein fraction in 2.0 mL, 1.0 mL was

transferred into 100 µL of immobilized TPCK-treated trypsin solution (20 TAME units) and mixed. The reaction mixture was incubated for 4 hours at 37°C (300 rpm). For the 3:1 protein sample, 1.0 mL of the protein fraction was transferred to 200 µL of immobilized chymotrypsin solution (8.3 ATEE units) as well, mixed, and incubated for 24 hours at 37°C (300 rpm). After incubation, the solutions were centrifuged at 100 g for 10 sec. 1.0 mL of the supernatant was transferred to 10 µL of formic acid (10% v/v), mixed and centrifuged at 20 000 g for 10 min. The supernatant was analyzed by LC-MS.

ESI was conducted using a needle voltage of 3 kV. Nitrogen was used as the sheath and auxiliary gas with the heated capillary set at 250°C. Positive mode mass spectra were obtained in the range of m/z 100 to 2000. MS/MS spectra were obtained by collision induced dissociation (CID) of the parent m/z, with the relative collision energy set to 35%. UV detection was performed at 195 nm. A Vydac Everest RP-C_{18} (250 mm × 2.1 mm I.D., 5 µm particle size, 300 Å) column (Grace Vydac, Hesperia, CA, USA) in an oven set at 45 °C, with a mobile phase consisting of (A) 0.1% (w/v) formic acid in water and (B) 0.1% (w/v) formic acid in acetonitrile was used in this experiment. The linear gradient program started with a 5 min isocratic hold at 96% (v/v) A and 4% (v/v) B, followed by a linear gradient to 60% (v/v) A + 40% (v/v) B at 113 min. The method also included a rinsing step at 60% B, followed by returning to the initial conditions and re-equilibration. The flow rate was set at 0.2 mL/min and a fixed injection volume of 20 µL was applied. Prediction of peak identity was performed upon comparison of m/z values with the SEQUEST algorithm of the Thermo BioWorks software (San José, CA, USA). Peptides with probability of more than 95% were reported and individually verified. Quantification occurred via ion-extracted chromatography (IEC) of the most abundant peptide charge responses. Peptide responses were normalized to the sum of all responses. The following calculations were made:

$$Sequence\ coverage\ (\%) = \frac{number\ of\ amino\ acid\ residues\ found}{total\ number\ of\ amino\ acid\ residues\ in\ somatropin\ (= 191)} \times 100$$

$$Lysine\ coverage\ (\%) = \frac{number\ of\ lysine\ residues\ found}{total\ number\ of\ lysine\ residues\ in\ somatropin\ (= 9)} \times 100$$

$$LysX\ coverage\ (\%) = \frac{number\ of\ peptides\ containing\ LysX}{number\ of\ peptides\ containing\ lysine\ residue} \times 100$$

$$NOTA - LysX\ modification\ yield\ (\%) = \frac{area\ of\ NOTA - LysX}{sum\ area\ of\ all\ LysX} \times 100$$

$$NOTA - LysX\ distribution\ yield\ (\%) = \frac{area\ of\ NOTA - LysX}{sum\ area\ of\ all\ NOTA\ modified\ lysines} \times 100$$

2.5 *In silico* pKa calculations

Structure-based pKa calculations were performed using the Adaptive Poisson−Boltzmann Solver (APBS version 1.1.0) [38], in which the pKa per titratable residue is determined as the sum of an unperturbed model value [39] and a perturbational shift reflecting the transfer of the amino acid from solution to the protein environment. The latter is calculated through a rigorous free energy cycle and numerical solution of the linearized Poisson−Boltzmann equation [40]. All calculations were carried out at 298.15 K with a solvent dielectric constant of 78.54 and a protein dielectric constant of 20. pKa calculations were performed on the protein structure only, taken from the last frame of a molecular dynamics (MD) simulation. Appropriate PQR files were generated with the aid of PDB2PQR version 1.4.0 [41, 42], employing the CHARMM [43, 44] parameterization.

Prior to the pKa calculations, structural calculations were performed using NAMD version 2.6 [45] and the CHARMM forcefield [43, 44], starting from the 1HGU crystal structure of the human somatotropin [46]. Optimal protonation states were assigned and missing atoms were added. In an ensuing 5000-step conjugate-gradient energy minimization only these atoms were allowed to move, while constraining all other atoms. Twelve amino acids were then mutated in accordance with the somatropin sequence (Q11D, E29Q, A47N, A50T, Q66E, A67T, Q75E, G91Q, D109N, A138I, A144S, A148T) followed by 5000 steps of conjugate-gradient minimization of the atoms of the mutated residues only. The resulting structure was solvated with 20.442 water molecules in an orthogonal box of size 83.1x86.5x95.1 $Å^3$ and made charge neutral by adding six sodium ions. The entire structure was subject to energy minimization (5000 conjugate gradient steps) with constraints on

all protein atoms. This was followed by an unconstrained MD equilibration run of 50 ps (1 fs time step) in the NVT ensemble at 300 K, employing Langevin dynamics with a damping coefficient of 1 ps^{-1} to control temperature. Electrostatics were treated with particle-mesh Ewald (PME) [47]. A short-range cutoff of 14 Å was maintained, and electrostatic and van der Waals interactions were gradually switched off starting from 10 Å. Neighbor lists were updated every 2 fs using a cutoff of 16.5 Å. The final production MD simulation totaled 1 ns with identical parameters.

2.6 Gallium labeling and quality control

For the labeling of 3:1 NOTA-somatropin with gallium, 45 nmol lyophilized protein sample was dissolved in 450 µL of 0.6 mM $GaCl_3$, 0.1 M HCl to obtain a 2:1 molar excess compared to the initial *p*-SCN-Bn-NOTA concentration. Then, 20 µL of 2 M NaOH and 40 µL of 0.1 M ammonium acetate, 0.2 mM acetylacetone buffer (pH 5.5) were added and the solution was mixed and incubated for 1 h at 20°C protected from light, while shaking at 300 rpm. The chelation efficiency with gallium was analyzed via peptide mapping as discussed in section 2.4.

3. RESULTS AND DISCUSSION

3.1 Production of NOTA-somatropin

Successful clinical use demands that a bifunctional chelating agent (BFCA) is both capable of maintaining a stable complex with a radionuclidic metal *in vivo*, *e.g.* Ga(III), while possessing a functional group which can be used for protein attachment. *p*-SCN-Bn-NOTA is a well-established hexadentate BFCA, forming an exceedingly stable complex with Ga(III) (log K = 30.98) [48]. The isothiocyanate function (R-NCS) allows formation of stable thiourea bonds at alkaline pH with free amines (Fig. 1). Somatropin has nine potential reaction sites (lysine's ε-amino residues) for the addition of *p*-SCN-Bn-NOTA. We have used a pH of 9.0 during synthesis: higher pH values will tend to accelerate the degradation of somatropin [49] and the R-NCS reagent, while lower pH values will lower the concentration of the free base form of the amines. These nine sites can lead to heterogeneous product formation consisting of different substitution degrees (*i.e.* the amount of bound NOTA-molecules per somatropin protein). Moreover, for a somatropin entity that has a single NOTA-

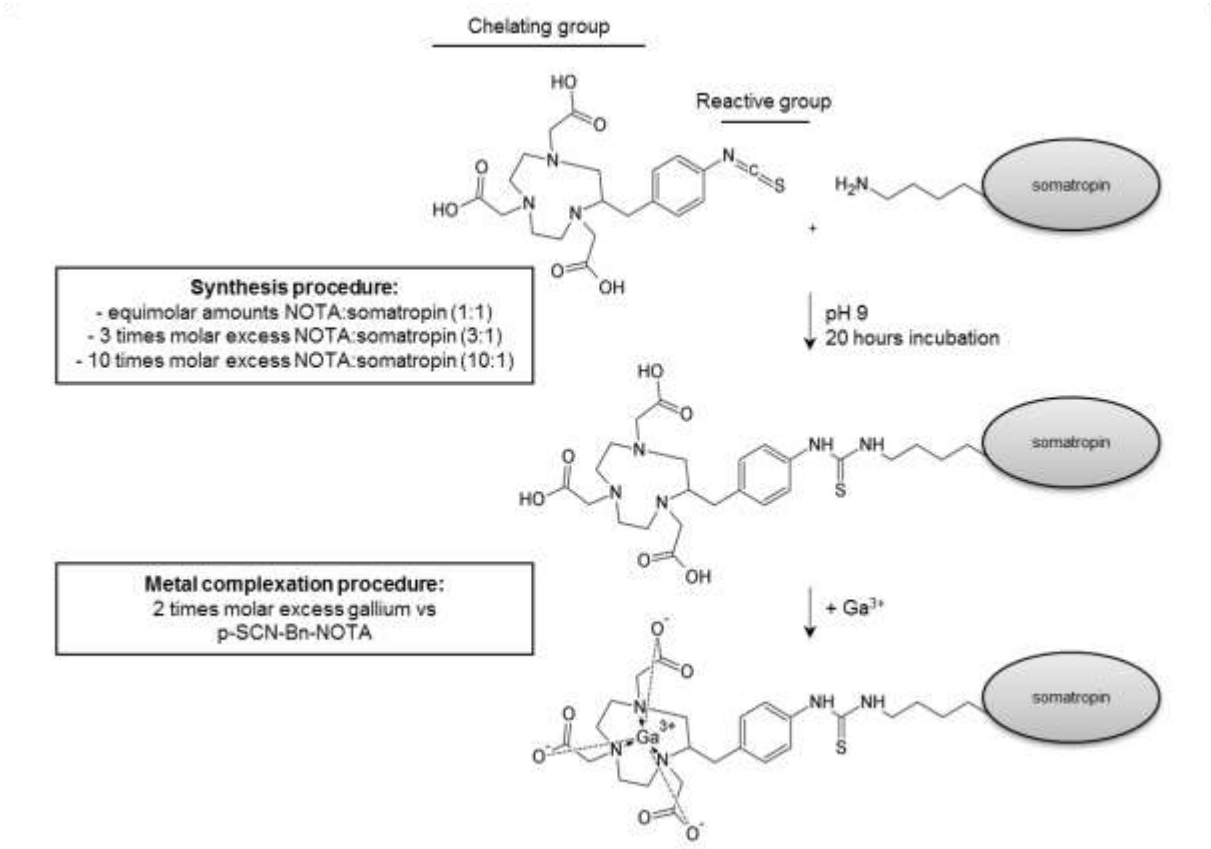

Fig. 1: Synthesis of gallium labeled NOTA-modified somatropin.

A

Substitution degree	Position isomers
1	9
2	36
3	84
4	126
5	126
6	84
7	36
8	9
9	1

B

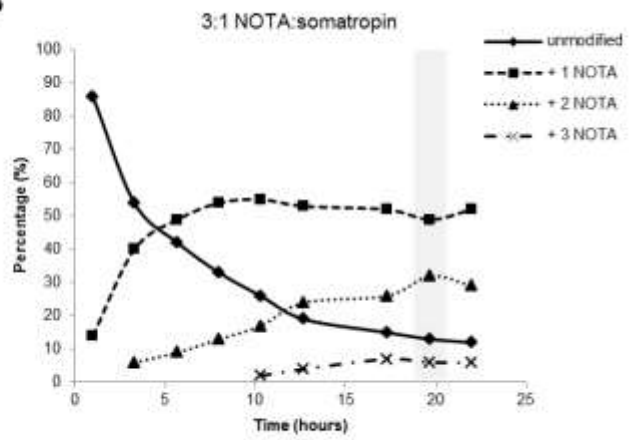

Fig 2: NOTA-modified somatropins product composition.. (A) Scheme of the possible products with different substitution degrees and position isomers. (B) Substitution degree of the 3:1 NOTA:somatropin preparation during 22 h of incubation, product composition after 20 h of incubation indicated in gray.

molecule attached, the NOTA-moiety may be attached at different amine sites. This creates the potential for a large number of position isomers as the degree of substitution increases (Fig. 2A) [50].

Analytical characterization of the heterogeneous production is very important as special attention must be paid to the lysine residues that are modified, since chemical modification can influence receptor binding and hence, the outcome of biological assays [51]. We applied three different synthesis ratios of *p*-SCN-Bn-NOTA based on references [52, 53]: equimolar amounts of *p*-SCN-Bn-NOTA and somatropin (1:1 NOTA:somatropin), three times molar excess of *p*-SCN-Bn-NOTA (3:1 NOTA:somatropin) and 10 times molar excess of *p*-SCN-Bn-NOTA (10:1 NOTA:somatropin) were applied.

3.2 Direct analytical characterization of the products

A single analytical technique for the characterization of biologicals is generally not sufficient [54, 55]. Therefore, a combination of LC-MS and CE-MS was applied to investigate the NOTA-somatropin products. In the biopharmaceutical field, LC is used for both the assessment of protein batch purity, protein modification (*e.g.* glycosylation) and aggregation [4, 56]. In combination with MS, precise and complementary information is generated. In this direct analytical approach, intact protein molecular ions generated by electrospray (ESI) or matrix-assisted laser desorption (MALDI) are introduced into the mass analyzer [57].

We have initially based our method on the related protein test described in the European Pharmacopoeia for somatropin (Ph. Eur. 8.0: 01/2008:0951) [58]: an isocratic LC-method using a mixture of 1-propanol and 0.05 M tris-hydrochloride buffer solution pH 7.5 (29:71 V/V) as the mobile phase. As the tris-hydrochloride buffer is not MS compatible, an ammonium bicarbonate buffer pH 7.5 was used. During pilot development, we have used isocratic methods with different mobile phase compositions (*i.e.* 20%, 30% and 35% organic mobile phase) and studied the somatropin retention time (RT: 49, 11 and 7.5 min, respectively). Based on the retention time of somatropin in the isocratic methods and the fact that the hydrophobicity of somatropin decreases upon NOTA labeling, we have used a 60 min linear gradient going from 30% to 50% 1-propanol. This method enabled us to characterize the NOTA:somatropin substitution degrees.

The NOTA-labeling was quantitatively monitored over time for 22 hours for the 3:1 NOTA:somatropin sample using our RP-C_4 method (Fig. 2B). The complexity of the product

composition was confirmed by a decrease of unmodified somatropin and an increase of higher order substitution degrees over time. Steady state was reached after 20 hours of incubation. Using this incubation period during synthesis, the unmodified somatropin was completely absent in the 10:1 NOTA:somatropin sample (Fig. 3A). The highest yield of the desired mono- and di-NOTA-somatropin (substitution degree +1 and +2) were obtained in the 3:1 sample (58%). Higher order modifications (substitution degree of more than 2) were heavily represented in sample 10:1 (98%). For some substitution degrees, different position isomers could be detected. The total peak recovery (peak balance) was found between 90-110%, confirming the analytical characterizing capacity of the method.

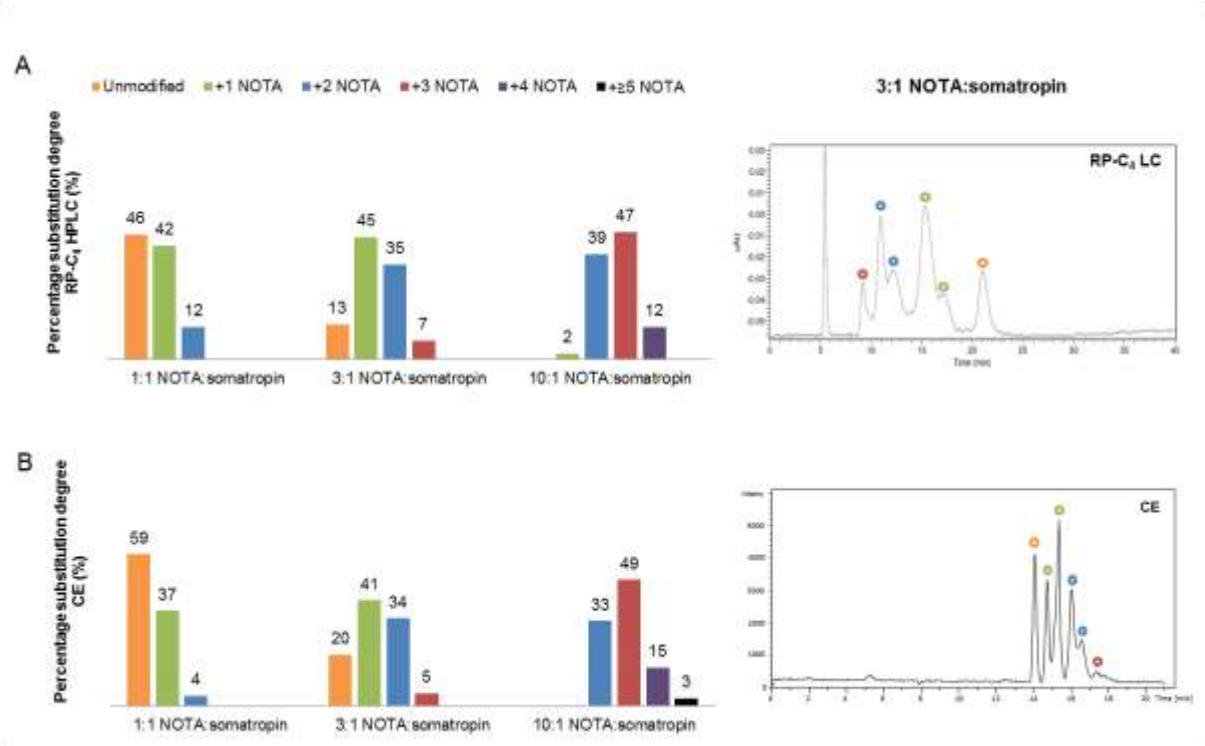

Fig. 3: Analytical results of the substitution degree of NOTA-modified somatropin obtained with RP-C_4 (A) and CE (B). A typical LC chromatogram (A: 50 mM ammonium bicarbonate pH 7.5 and B: n-propanol: a 60 min linear gradient from 70% A (v/v) + 30% B (v/v) to 50% A (v/v) 50% B (v/v)) and CE electropherogram of the 3:1 NOTA-somatropin sample are given.

Our findings were confirmed by CE-MS (Fig. 3B), a technique also included in the somatropin monograph of the European Pharmacopoeia (Ph. Eur. 8.0: 01/2008:0951) [58]. In CE, the separation is based on charge differences [59], in our case a loss of one positive charge of lysine with the

simultaneous addition of negative charges of NOTA, whereas in RP HPLC mainly the decrease in hydrophobicity due to the attachment of NOTA leads to the separation [56]. Overall, the 1:1 and 3:1 samples show similar degrees of somatropin modification as measured with CE-MS compared to LC-MS (Fig. 3). In the 10:1 sample, CE-MS revealed somatropin with NOTA substitution degrees ≥5 (amounted for 3%). As these highly polar compounds end up in the LC dead time they were not detected, whereas in CE they migrate later and, therefore, could be detected. In general, we can conclude that both CE-MS and LC-MS techniques lead to similar conclusions regarding the extent of modification. However, higher order substitution degrees are detectable using CE-MS.

3.3 Peptide mapping of the NOTA-modified somatropins

The current gold standard in protein characterization is the "bottom-up approach". This method relies on the digestion of a mixture of proteins of interest and subsequent analysis of the digested peptides by LC-MS. All peptides were identified based on their peptide mass fingerprint (m/z value) and CID fragmentation pattern, thereby establishing the validity of these patterns for peptide identification and structural elucidation of the protein modification (Fig. 4). Peptides containing the NOTA-label are characterized by a mass increase of 449.52 Da and specific NOTA-losses in their CID spectra (Fig. 4A-B). Another observation made was the lower residual protein quantity upon the different modification and chelation steps, which are each followed by a desalting step and a lyophilisation step in case of NOTA-modification. Therefore, the UV absorption upon NOTA-modification (going from A to B) is lower.

The monograph of somatropin in the European Pharmacopoeia (Ph. Eur.) includes a peptide mapping method using trypsin (Ph. Eur. 8.0: 01/2008:0951) [58]. It is commonly known that trypsin is a specific protease that cleaves at the *c*-terminal of arginine and lysine residues [60]. This yields theoretically 21 peptide fragments for somatropin (PeptideCutter [61]). However, trypsin skips lysine cleavage sites when NOTA-modifications are present, *e.g.* for the NOTA-labeled Lys-158 protein fragment (fragment 146-164) one trypsin cleavage site was missed. The individual peptides with NOTA-modification are given in Table 1. They represent modifications of lysine residues at Lys-70, Lys-158, Lys-140 and Lys-172, with Lys-70 found in all samples. The sequence coverage of the

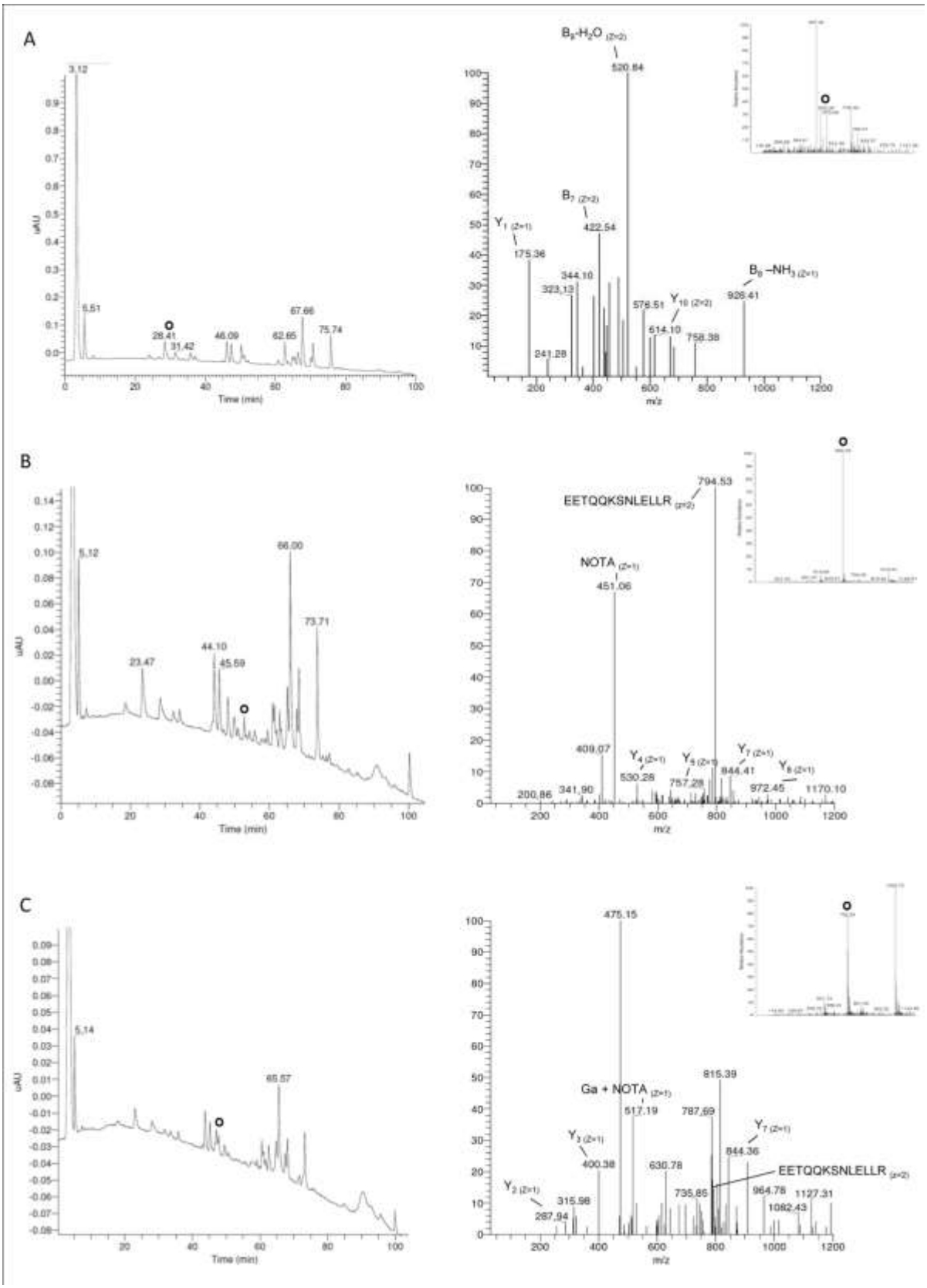

Fig. 4: LC chromatogram of the tryptic digest and MS^2 spectra of the peptide EETQQKSNLELLR (A) with NOTA-modification (B) and gallium labeling (C), indicated by a black circle. Inset: MS spectra, with the selected precursor ion indicated by a black circle.

trypsin digest was more than 95% for the control (*i.e.* unmodified somatropin sample) and more than 90% for the NOTA-labeled somatropin samples; all lysine residues were recovered (100% lysine coverage) in the control and NOTA:somatropin samples. The NOTA-Lys-70 modification yield was very high in all samples (94-100%). The Lys-70 coverage amounted for 6% (*i.e.* the ratio between number of peptides containing Lys-70 and number of peptides containing a Lys residue), which is half of the theoretical specific Lys coverage of 11%. This means that relatively less fragments with Lys-70 were recovered compared to the other Lys-peptides. Indeed, for Lys-140, Lys-158 and Lys-172, the specific lysine (LysX) coverage was more than 11%.

Table 1: Peptide mapping results (tryptic digest) of the individual peptides with NOTA modification

Product	Sequence	Th. mass (Da) (z)	Exp. mass (Da)	RT (min)	NOTA-LysX distribution yield (%)	NOTA-LysX modification yield (%)
1:1	EETQQ**K**^{70}SNLELLR + NOTA	680.07 (3)	680.13	56.16	100	100 (n=1)
3:1	EETQQ**K**^{70}SNLELLR + NOTA	680.07 (3)	680.28	52.95	64	94 (n=4)
	FDTNSHNDDALL**K**^{158}NYGLLY + NOTA	888.61 (3)	888.68	72.37	30	24 (n=4)
	DTNSHNDDALL**K**^{158}NY + NOTA	690.71 (3)	691.03	77.37		
	DMD**K**^{172}VETFLR + NOTA	568.64 (3)	568.71	67.53	4	7 (n=4)
	TGQIF**K**^{140}QTY + NOTA	768.36 (2)	768.45	59.69	2	4 (n=4)
10:1	EETQQ**K**^{70}SNLELLR + NOTA	680.07 (3)	680.25	55.61	80	100 (n=1)
	DTNSHNDDALL**K**^{158}NY + NOTA	690.71 (3)	691.1	80.02	13	100 (n=1)
	TGQIF**K**^{140}QTY + NOTA	768.36 (2)	768.52	56.75	7	8 (n=1)

Product: molar ratio of NOTA over somatropin. NOTA-LysX distribution yield: percentage NOTA of LysX among all NOTA modified lysine residues. NOTA-LysX modification yield: percentage of LysX that is NOTA modified (i.e. lysine site depicted in bold in the sequence).

Not only in HPLC but also in the hydrolysis step of the peptide mapping, orthogonal systems can be used to further analyze the modification yield from the different molar ratios. Chymotrypsin mainly cleaves peptide bonds in which the carboxyl group is contributed by phenylalanine, tryptophan and tyrosine, theoretically also yielding 21 fragments for somatropin [61]. In addition, leucine and methionine may be cleaved as well, although at a much lower rate. This would theoretically yield 54 somatropin fragments [61]. Results are given in S.I.1. A relatively low overall sequence coverage was

obtained after chymotryptic peptide mapping of control and NOTA-modified samples (below 80%). The Lys-70 residue was found in 37% of the Lys containing peptide fragments, *i.e.* a much higher Lys-70 coverage compared to trypsin-cleaved peptide mapping, but at the expense of Lys-158 with a LysX coverage of 0%. The NOTA-LysX distribution yield in sample 3:1 NOTA:somatropin was also found mostly on Lys-70 (86%) followed by Lys-140 (14%), with NOTA-LysX modification yields of 50% and 67%, respectively. The use of chymotrypsine therefore gives another perspective on the modification yield compared to trypsine (94% vs 50% for Lys-70), confirming the structural information. Similar conclusions were also obtained with an enzyme combination of trypsin and chymotrypsin (see S.I.2.). The use of *S. aureus* V8 protease was not suited for peptide mapping: the sequence coverage was below 40% and the lysine coverage amounted 22% for NOTA-labeled somatropin and control (see S.I.3.).

Our data indicate that Lys-70 is a hotspot for NOTA-modification (Table 1), which was also suggested by Sakal *et al.* after modification of somatropin with fluorescein isothiocyanate (FITC) [62]. Our *in silico ab initio* pKa calculations revealed that the Lys-70 residue has a lower pKa value (pKa = 8.3) than the other eight Lys-residues and is therefore more reactive under our modification conditions (Fig. 5). In addition, Lys-70 is positioned outside the binding pocket of the somatropin:hGHR interaction: modification of this lysine residue will therefore not directly interfere with receptor binding (Fig. 5). According to our data, *i.e*. the NOTA-LysX distribution yield and modification yield, we conclude that Lys-70 followed by Lys-158 are most reactive towards *p*-SCN-Bn-NOTA and Lys-140 and Lys-172 are the less reactive lysine residues. Except for Lys-172, all found lysine residues lie outside the binding pocket with the receptor.

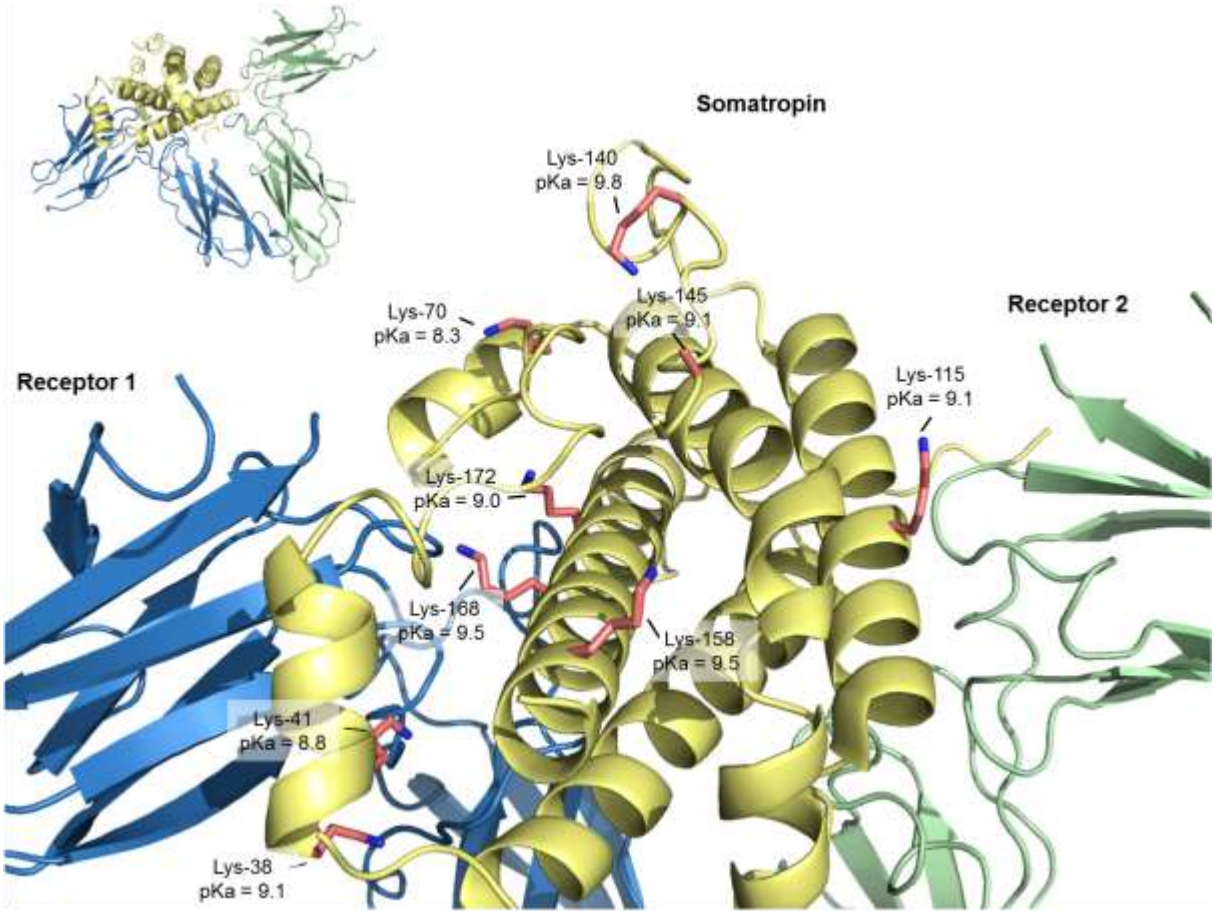

Fig. 5: Position of lysine residues on somatropin in complex with hGHR (PDB: 3HHR) [26]. Yellow: somatropin; blue: hGHR (site I); green: hGHR (site II). *In silico* calculated pKa values for each lysine residue are indicated.

3.4 Analysis of gallium labeled NOTA-somatropin

For the complexation of gallium in NOTA, two times molar excess of gallium was used compared to the used amount of p-SCN-Bn-NOTA during NOTA:somatropin synthesis. Similar results as for NOTA-modified somatropin (section 3.2) were obtained: the overall sequence coverage after tryptic peptide mapping was 96%, with 100% lysine coverage. Tryptic peptide mapping of gallium labeled NOTA-somatropin demonstrated that all NOTA molecules were complexed with gallium (Fig. 4C), resulting in labeling efficiencies of 100%.

4. CONCLUSION

The presence of multiple reactive sites on biomolecules towards chemical modifications during conjugation reactions can have a great impact on the product composition, and hence, the biological

activity. Therefore, it is important to analytically characterize the products originating from different synthesis procedures. Direct analytical and bottom-up approaches were used to profile the different modified somatropin proteins and demonstrated that higher amounts of *p*-SCN-Bn-NOTA during synthesis led to a heterogeneous product with higher order substitution degrees, as well as different position isomers. The 1:1 NOTA:somatropin synthesis procedure yielded the highest mono-NOTA-somatropin fraction (42%) with less higher order substitution degrees (≥2 NOTA, 12%); Lys-70 was found to be the modification hotspot towards *p*-SCN-Bn-NOTA. We conclude that Lys-70 followed by Lys-158 are most reactive towards p-SCN-Bn-NOTA and Lys-140 and Lys-172 are the less reactive lysine residues. Except for Lys-172, all found lysine residues lie outside the binding pocket with the receptor. The synthesis of NOTA-somatropin is a suited synthesis procedure for the production of target-specific radiopharmaceuticals for further investigation of the treatment and visualization of growth hormone receptor overexpressing cancers.

5. ACKNOWLEDGEMENTS

This research project was supported by grants from Ghent University (BOF special research fund 01J22510) to BDS and EW and the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) to MD (101529). We thank Dr. Valentijn Vergote for his operational help in the initial experiments. The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by Ghent University, the Hercules Foundation and the Flemish Government – department EWI.

5. REFERENCES

[1] A. Mullard, 2012 FDA drug approvals, Nature Reviews Drug Discovery. 12 (2013) 87-90.

[2] S. Lawrence, Billion dollar babies - biotech drugs as blockbusters (vol 25, pg 380, 2007), Nat. Biotechnol. 25 (2007) 380-382.

[3] C.K. Schneider, C. Vleminckx, I. Gravanis, F. Ehmann, J.H. Trouvin, M. Weise, S. Thirstrup, Setting the stage for biosimilar monoclonal antibodies, Nat. Biotechnol. 30 (2012) 1179-1185.

[4] S.A. Berkowitz, J.R. Engen, J.R. Mazzeo, G.B. Jones, Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars, Nature Reviews Drug Discovery. 11 (2012) 527-540.

[5] J. Arrowsmith, A decade of change, Nature Reviews Drug Discovery. 11 (2012) 17-18.

[6] A. Signore, A. Annovazzi, M. Chianelli, F. Corsetti, C. Van de Wiele, R.N. Watherhouse, Peptide radiopharmaceuticals for diagnosis and therapy, Eur. J. Nucl. Med. 28 (2001) 1555-1565.

[7] I.U. Khan, A.G. Beck-Sickinger, Targeted tumor diagnosis and therapy with peptide hormones as radiopharmaceuticals, Anticancer Agents Med Chem. 8 (2008) 186-199.

[8] S.M. Okarvi, Peptide-based radiopharmaceuticals and cytotoxic conjugates: potential tools against cancer, Cancer Treat. Rev. 34 (2008) 13-26.

[9] E. Wynendaele, N. Bracke, S. Stalmans, B.D. Spiegeleer, Development of Peptide and Protein Based Radiopharmaceuticals, Curr. Pharm. Des. (2013). *In press.*

[10] C.S. Cutler, H.M. Hennkens, N. Sisay, S. Huclier-Markai, S.S. Jurisson, Radiometals for combined imaging and therapy, Chem. Rev. 113 (2013) 858-883.

[11] J. Buscombe, Radiolabeled Probes Targeting G-Protein-Coupled Receptors For Personalized Medicine, Curr. Pharm. Des. (2013). *In press.*

[12] T.K. Koch, B.O. Berg, S.J. Dearmond, R.F. Gravina, Creutzfeldt-Jakob Disease in a Young-Adult with Idiopathic Hypopituitarism - Possible Relation to the Administration of Cadaveric Human Growth-Hormone, N. Engl. J. Med. 313 (1985) 731-733.

[13] G. Walsh, Biopharmaceutical benchmarks 2010, Nat. Biotechnol. 28 (2010) 917-924.

[14] A. Dubois, S. Gsteiger, S. Balser, E. Pigeolet, J.L. Steimer, G. Pillai, F. Mentre, Pharmacokinetic similarity of biologics: analysis using nonlinear mixed-effects modeling, Clin. Pharmacol. Ther. 91 (2012) 234-242.

[15] H.K. Heim, Similar biological medicinal products currently licensed in the European union--overview of non-clinical study programs, Biologicals. 39 (2011) 284-288.

[16] P. Saenger, Current status of biosimilar growth hormone, Int J Pediatr Endocrinol. 2009 (2009) 370329.

[17] H.T. Jiang, S.L. Wu, B.L. Karger, W.S. Hancock, Mass Spectrometric Analysis of Innovator, Counterfeit, and Follow-On Recombinant Human Growth Hormone, Biotechnol. Prog. 25 (2009) 207-218.

[18] M. Thevis, M. Bredehoft, M. Kohler, W. Schanzer, Mass spectrometry-based analysis of IGF-1 and hGH, Handb Exp Pharmacol. (2010) 201-207.

[19] M. Bidlingmaier, C.J. Strasburger, Growth hormone, Handb Exp Pharmacol. (2010) 187-200.

[20] C. Ehrnborg, B.A. Bengtsson, T. Rosen, Growth hormone abuse, Baillieres Best Pract Res Clin Endocrinol Metab. 14 (2000) 71-77.

[21] R.I.G. Holt, I. Erotokritou-Mulligan, P.H. Sonksen, The history of doping and growth hormone abuse in sport, Growth Horm. IGF Res. 19 (2009) 320-326.

[22] D. Courtheyn, B. Le Bizec, G. Brambilla, H.F. De Brabander, E. Cobbaert, A.V. de Wiele, J. Vercammen, K. De Wasch, Recent developments in the use and abuse of growth promoters, Anal. Chim. Acta. 473 (2002) 71-82.

[23] S.J. Olshansky, T.T. Perls, New developments in the illegal provision of growth hormone for "anti-aging" and bodybuilding, Jama-Journal of the American Medical Association. 299 (2008) 2792-2794.

[24] B.C. Cunningham, J.A. Wells, Comparison of a structural and a functional epitope, J. Mol. Biol. 234 (1993) 554-563.

[25] T. Amit, H. Hacham, O. Daily, P. Hertz, R.J. Barkey, Z. Hochberg, The Hep G2 cell line in the study of growth hormone receptor/binding protein, Mol. Cell. Endocrinol. 101 (1994) 29-36.

[26] A.M. de Vos, M. Ultsch, A.A. Kossiakoff, Human growth hormone and extracellular domain of its receptor: crystal structure of the complex, Science. 255 (1992) 306-312.

[27] D. Poger, A.E. Mark, Turning the growth hormone receptor on: evidence that hormone binding induces subunit rotation, Proteins. 78 (2010) 1163-1174.

[28] M. Bidosee, R. Karry, E. Weiss-Messer, R.J. Barkey, Regulation of growth hormone receptors in human prostate cancer cell lines, Mol. Cell. Endocrinol. 309 (2009) 82-92.

[29] E. Weiss-Messer, O. Merom, A. Adi, R. Karry, M. Bidosee, R. Ber, A. Kaploun, A. Stein, R.J. Barkey, Growth hormone (GH) receptors in prostate cancer: gene expression in human tissues and cell

lines and characterization, GH signaling and androgen receptor regulation in LNCaP cells, Mol. Cell. Endocrinol. 220 (2004) 109-123.

[30] M. Gebre-Medhin, L.G. Kindblom, H. Wennbo, J. Tornell, J.M. Meis-Kindblom, Growth hormone receptor is expressed in human breast cancer, Am. J. Pathol. 158 (2001) 1217-1222.

[31] X.D. Yang, F.K. Liu, Z. Xu, C. Chen, G. Li, X.Y. Wu, J.S. Li, Growth hormone receptor expression in human colorectal cancer, Dig. Dis. Sci. 49 (2004) 1493-1498.

[32] P.E. Clayton, I. Banerjee, P.G. Murray, A.G. Renehan, Growth hormone, the insulin-like growth factor axis, insulin and cancer risk, Nat Rev Endocrinol. 7 (2011) 11-24.

[33] J.K. Perry, D.X. Liu, Z.S. Wu, T. Zhu, P.E. Lobie, Growth hormone and cancer: an update on progress, Curr Opin Endocrinol Diabetes Obes. 20 (2013) 307-313.

[34] M. Maamra, J. Finidori, S. Von Laue, S. Simon, S. Justice, J. Webster, S. Dower, R. Ross, Studies with a growth hormone antagonist and dual-fluorescent confocal microscopy demonstrate that the full-length human growth hormone receptor, but not the truncated isoform, is very rapidly internalized independent of Jak2-Stat5 signaling, J. Biol. Chem. 274 (1999) 14791-14798.

[35] M.W. Brechbiel, Bifunctional chelates for metal nuclides, Q J Nucl Med Mol Imaging. 52 (2008) 166-173.

[36] E.W. Price, C. Orvig, Matching chelators to radiometals for radiopharmaceuticals, Chem Soc Rev. 43 (2014) 260-290.

[37] R. Haselberg, V. Brinks, A. Hawe, G.J. de Jong, G.W. Somsen, Capillary electrophoresis-mass spectrometry using noncovalently coated capillaries for the analysis of biopharmaceuticals, Analytical and Bioanalytical Chemistry. 400 (2011) 295-303.

[38] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10037-10041.

[39] J.E. Nielsen, G. Vriend, Optimizing the hydrogen-bond network in Poisson-Boltzmann equation-based pK(a) calculations, Proteins-Structure Function and Genetics. 43 (2001) 403-412.

[40] R.E. Bank, M. Holst, A new paradigm for parallel adaptive meshing algorithms, Siam Journal on Scientific Computing. 22 (2000) 1411-1443.

[41] T.J. Dolinsky, J.E. Nielsen, J.A. McCammon, N.A. Baker, PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations, Nucleic Acids Res. 32 (2004) W665-W667.

[42] T.J. Dolinsky, P. Czodrowski, H. Li, J.E. Nielsen, J.H. Jensen, G. Klebe, N.A. Baker, PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, Nucleic Acids Res. 35 (2007) W522-W525.

[43] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, Journal of Physical Chemistry B. 102 (1998) 3586-3616.

[44] A.D. Mackerell, M. Feig, C.L. Brooks, Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, Journal of Computational Chemistry. 25 (2004) 1400-1415.

[45] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, Journal of Computational Chemistry. 26 (2005) 1781-1802.

[46] L. Chantalat, N.D. Jones, F. Korber, J. Navaza, A.G. Pavlovsky, The Crystal-Structure of Wild-Type Growth-Hormone at 2.5 Angstrom Resolution, Protein and Peptide Letters. 2 (1995) 333-340.

[47] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A Smooth Particle Mesh Ewald Method, J. Chem. Phys. 103 (1995) 8577-8593.

[48] E.T. Clarke, A.E. Martell, Stabilities of the Fe(Iii), Ga(Iii) and in(Iii) Chelates of N,N',N''-Triazacyclononanetriacetic Acid, Inorganica Chimica Acta. 181 (1991) 273-280.

[49] M.C. Manning, K. Patel, R.T. Borchardt, Stability of protein pharmaceuticals, Pharmaceutical research. 6 (1989) 903-918.

[50] W.J. Zhang, X. Luo, Y.L. Liu, X.X. Shao, J.D. Wade, R.A.D. Bathgate, Z.Y. Guo, Site-specific DOTA/europium-labeling of recombinant human relaxin-3 for receptor-ligand interaction studies, Amino Acids. 43 (2012) 983-992.

[51] A.E. Loot, A. van Buiten, A.J. Roks, R.H. Henning, The suitability of iodinated Angiotensin-(1-7) peptides as pharmacological tools, J Pharmacol Toxicol Methods. 51 (2005) 51-55.

[52] J.M. Jeong, M.K. Hong, Y.S. Chang, Y.S. Lee, Y.J. Kim, G.J. Cheon, D.S. Lee, J.K. Chung, M.C. Lee, Preparation of a promising angiogenesis PET imaging agent: Ga-68-labeled c(RGDyK)-isothiocyanatobenzyl-1,4,7-triazacyclononane-1,4,7-triacetic acid and feasibility studies in mice, J. Nucl. Med. 49 (2008) 830-836.

[53] Z.B. Li, K. Chen, X. Chen, Ga-68-labeled multimeric RGD peptides for microPET imaging of integrin alpha(v)beta(3) expression, European Journal of Nuclear Medicine and Molecular Imaging. 35 (2008) 1100-1108.

[54] A. Staub, D. Guillarme, J. Schappler, J.L. Veuthey, S. Rudaz, Intact protein analysis in the biopharmaceutical field, J. Pharm. Biomed. Anal. 55 (2011) 810-822.

[55] R. Haselberg, G.J. de Jong, G.W. Somsen, Capillary electrophoresis-mass spectrometry for the analysis of intact proteins 2007-2010, Electrophoresis. 32 (2011) 66-82.

[56] S. Fekete, J.L. Veuthey, D. Guillarme, New trends in reversed-phase liquid chromatographic separations of therapeutic peptides and proteins: theory and applications, J. Pharm. Biomed. Anal. 69 (2012) 9-27.

[57] B.A. Garcia, What does the future hold for Top Down mass spectrometry?, J. Am. Soc. Mass Spectrom. 21 (2010) 193-202.

[58] European Directorate for the Quality of Medicines and HealthCare, European Pharmacopoeia 8.0., Strassbourg: Council of Europe 2014.

[59] R. Haselberg, G.J. de Jong, G.W. Somsen, Capillary electrophoresis-mass spectrometry for the analysis of intact proteins, J. Chromatogr. A. 1159 (2007) 81-109.

[60] J.V. Olsen, S.E. Ong, M. Mann, Trypsin cleaves exclusively C-terminal to arginine and lysine residues, Mol Cell Proteomics. 3 (2004) 608-614.

[61] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein identification and analysis tools in the ExPASy server, Methods Mol. Biol. 112 (1999) 531-552.

[62] E. Sakal, A. Gertler, Y. Shechter, Biological-Activity of a Fluorescein Human Growth-Hormone Derivative Prepared by Specific Covalent Labeling of Lysine-70, Biochemistry (Mosc). 30 (1991) 8899-8904.

Supplementary data

S.I.1.: Peptide mapping results (chymotryptic digest) of the individual peptides with NOTA modification

Product	Sequence	Th. mass (Da) (z)	Exp. mass (Da)	RT (min)	NOTA-LysX distribution yield (%)	NOTA-LysX modification yield (%)
1:1	SESIPTPSNREETQQ**K**^{70}S + NOTA	790.17 (3)	789.59	39.12	100 (n=1)	100 (n=1)
3:1	NREETQQ**K**^{70}SNL + NOTA	599.64 (3)	599.75	31.88	86 (n=3)	50 (n=3)
	SESIPTPSNREETQQ**K**^{70}S + NOTA	790.17 (3)	790.25	36.30		
	SESIPTPSNREETQQ**K**^{70}SNL + NOTA	865.91 (3)	865.99	45.47		
	K^{140}QTY + NOTA	495.02 (3)	495.23	34.85	14 (n=3)	67 (n=3)
10:1	NREETQQ**K**^{70}SNL + NOTA	599.64 (3)	599.86	36.06	100 (n=1)	20 (n=1)
	SESIPTPSNREETQQ**K**^{70}S + NOTA	790.17 (3)	790.22	39.25		

Product: molar ratio of NOTA over somatropin. NOTA-LysX distribution yield: percentage NOTA of LysX among all NOTA modified lysine residues. NOTA-LysX modification yield: percentage of LysX that is NOTA modified (*i.e.* lysine site depicted in bold in the sequence).

S.I.2.: Enzyme combination (trypsine and chymotrypsin) used in peptide mapping

The 3:1 NOTA-modified somatropin (45 nmol protein) was dissolved in 1000 µL 6 M guanidine HCl, 35 mM Tris, 20 mM DL-dithiothreitol at pH 7.5 and incubated for 30 min at 37°C, while shaking at 300 rpm. S-carboxymethylation of cysteine residues was performed by addition of 20 µL of iodoacetate (pH 7.2; 58 mM) and subsequent incubation for 30 min at 37°C (at 300 rpm). 100 µL of DL-dithiothreitol (1 M) was added and mixed. The sample was loaded onto a PD-10 column sephadex G-25M (equilibrated with digestion buffer consisting of 10 mM ACES, 20 mM $CaCl_2$, pH 7.0). After collection of the protein fraction in 2.0 mL, 1.0 mL was transferred into 300 µL of immobilized trypsin and chymotrypsin solution (20 TAME and 8.3 ATEE units resp., Pierce, Erembodegem, Belgium), mixed, and incubated for 24 hours at 37°C (300 rpm). After incubation, the solutions were centrifuged at 100 g for 10 sec. 1.0 mL of the supernatant was transferred to 10 µL of formic acid (10% v/v), mixed and centrifuged at 20 000 g for 10 min. The supernatant was analyzed by LC-MS.

Peptide mapping results (enzymatic combination of trypsin and chymotrypsin) of the individual peptides with NOTA modification

Product	Sequence	Th. mass (Da) (z)	Exp. mass (Da)	RT (min)	NOTA-LysX distribution yield (%)	NOTA-LysX modification yield (%)
3:1	EETQQ**K**^{70}S + NOTA	433.77 (3)	433.89	21.89	33 (n=1)	5 (n=1)
	DMD**K**^{172}VE + NOTA	593.65 (2)	593.65	48.76	67 (n=1)	6 (n=1)
	DMD**K**^{172}VETF + NOTA	478.84 (3)	479.01	0.38		

Product: molar ratio of NOTA over somatropin. NOTA distribution: percentage NOTA of LysX among all NOTA modified lysine residues. Modification yield: percentage of LysX that is NOTA modified (*i.e.* lysine site depicted in bold in the sequence).

S.I.3.: *S. aureus* V8 protease used in peptide mapping

The 3:1 NOTA-modified somatropin (45 nmol protein) and control (45 nmol protein) were separately dissolved and treated as described in S.I.2. to perform S-carboxymethylation and reduction. The samples were loaded onto a PD-10 column sephadex G-25M (equilibrated with digestion buffer consisting of 10 mM ACES, 20 mM $CaCl_2$, pH 7.0) and eluted in 2.0 mL with V-8 digestion buffer pH 7.8 (Pierce, Erembodegem, Belgium). 1.0 mL was transferred into 125 µL of V-8 protease solution (6 ± 1.6 units, Pierce, Erembodegem, Belgium), mixed, and incubated for 18 hours at 37°C (300 rpm). After incubation, the solutions were centrifuged at 100 g for 10 sec. 1.0 mL of the supernatant was transferred to 10 µL of formic acid (10% v/v), mixed and centrifuged at 20 000 g for 10 min. The supernatant was analyzed by LC-MS.