Cent. Eur. J. Biol. • 3(4) • 2008 • 359–370
DOI: 10.2478/s11535-008-0045-2

VERSITA

Central European Journal of **Biology**

Novel post-translational modifications of Smad2 identified by mass spectrometry

Research Article

Nimesh Bhaskaran[1], Hiroyuki Iwahana[1,2], Jonas Bergquist[3], Ulf Hellman[4], Serhiy Souchelnytskyi[1*]

[1] *Karolinska Biomics Center, Department of Oncology-Pathology, Karolinska Institutet, Karolinska University Hospital, Solna, SE-17176, Stockholm, Sweden*

[2] *Division of Support in Community Medicine, Center for Community Medicine, Jichi Medical University, 329-0498 Shimotsuke, Tochigi, Japan*

[3] *Department of Analytical Chemistry, Uppsala University, SE-751 24, Uppsala, Sweden*

[4] *Ludwig Institute for Cancer Research, Uppsala University, SE-751 24, Uppsala, Sweden*

Received 19 May 2008; Accepted 1 September 2008

Abstract: Smad2 is a crucial component of transforming growth factor-β (TGFβ) signaling, and is involved in the regulation of cell proliferation, death and differentiation. Phosphorylation, ubiquitylation and acetylation of Smad2 have been found to regulate its activity. We used mass spectrometry to search for novel post-translational modifications (PTMs) of Smad2. Peptide mass fingerprinting (PMF) indicated that Smad2 can be acetylated, methylated, citrullinated, phosphorylated and palmitoylated. Sequencing of selected peptides validated methylation at Gly122 and hydroxylation at Trp18 of Smad2. We also observed a novel, so far unidentified modification at Tyr128 and Tyr151. Our observations open for further exploration of biological importance of the detected PTMs.

Keywords: *Smad2 • Mass spectrometry • Post-translational modifications*

Abbreviations

TGFβ - transforming growth factor-β;
PTM - post-translational modification;
S/TKR - serine/threonine kinase receptor;
CAF - chemically assisted fragmentation;
MH1 and MH2 - mad-homology 1 and 2 domains.

1. Introduction

TGFβ is a polypeptide growth factor which signals by binding to specific transmembrane serine/threonine kinase receptors (S/TKR) [1,2]. Type I and type II TGFβ receptors form a heterotetrameric complex upon the binding of the ligand, TBFβ. It is believed that the type I TGFβ receptor (TβR-I) is the kinase which phosphorylates Smad2 and Smad3 at their C-terminus. This phosphorylation of Smad2 and Smad3 triggers Smad-dependent signaling which is crucial for most activities of TGFβ [1,2]. Smad2 is thought to be vital in development as a Smad2 knock-out results in embryonic lethality [3].

Protein functions are regulated by post-translational modifications (PTMs). More than 400 PTMs of proteins have been described, including phosphorylation, acetylation, methylation, ubiquitylation and proteolytic cleavage [4]. Smad2 has previously been found to be post-translationally modified in several ways. Firstly serine and threonine phosphorylation of Smad2 was

* *E-mail: serhiy.souchelnytskyi@ki.se*

found to be involved in the regulation of Smad2-protein interactions and nuclear-cytoplasm translocation [1,2]. Acetylation of Smad2 was found to interfere with its transcriptional activity [5,6], and ubiquitylation of Smad2 also occurs and controls its protein levels in the cell [7,8].

Mass spectrometry is the best available technology for identifying many PTMs in a given protein simultaneously, due to its high sensitivity, resolution and structural characterization ability [9,10]. PTMs can be detected by a shift in the mass of a peptide with a modification, as compared to the native peptide, for example, an increase in mass of 79.9 Da indicates that a peptide may contain a phosphorylated residue [9,10]. Mass spectrometry can also provide information about location of a modified residue by fragmenting the peptide and observing the fragmentation spectrum for a PTM-dependent mass shift caused by a modified residue [9-11]. Here we report mass spectrometry-based identification of novel PTMs of Smad2 expressed in MCF7 human breast epithelial cells.

2. Experimental Procedures

2.1. Cells and reagents

MCF7 human breast adenocarcinoma cells were obtained from American Type Culture Collection (Manassas, VA) and were cultured in a DMEM medium (Sigma), 10% fetal bovine serum (Gibco). Human Smad2 with a Flag tag at the N-terminus was cloned in a pMEP4 vector under metal-inducible promoter, and was stably transfected in MCF7 cells, followed by selection with hygromycin. Expression of Flag-Smad2 was induced by incubation of MCF7 transfected cells with 5 μM $CdCl_2$ for 4 h, and was monitored by immunodetection with anti-Flag (Sigma) and anti-Smad2 (#543; described in [5,12]) antibodies. As controls, parental and empty pMEP4 vector-transfected MCF7 cells were used, as well as Flag-Smad2-transfected cells without induction of Flag-Smad2 expression.

2.2. Preparation of Flag-Smad2

Proteins were extracted from cells with an extraction buffer (0.5% TritonX-100, 50 mM Tris-HCl, pH 7.2, including a protease inhibitors cocktail (Roche)). Protein extracts were made from MCF7 with or without induced Flag-Smad2 and MCF7 containing empty pMEP4 vector. Flag-Smad2 was immunoprecipitated with anti-Flag antibodies. Immunocomplexes were collected on Sepharose-A beads, extensively washed, and immunoprecipitates were solubilized in SDS sample buffer (25 mM Tris-HCl, pH 6.7, 2.5% SDS, 10% glycerol). Immunoprecipitates were separated by SDS-PAGE, and gels were stained with Coomassie Brilliant Blue or silver, as previously described [12]. Two-dimensional gel electrophoresis was performed as described earlier [12].

2.3. Sample preparation for mass spectrometry

Flag-Smad2 containing protein bands were excised from 1D SDS-PAGE and subjected to in-gel digestion with trypsin, as previously described [12]. In brief, excised gel pieces were de-stained, extensively washed with 0.1 M ammonium bicarbonate, dehydrated by treatment with 100% acetonitrile, and then the gels pieces were air-dried after the removal of acetonitrile. Trypsin (Promega) (10-20 μl; 1 μg/10 μl in 0.1 M ammonium bicarbonate) was added to the dried gel piece, and a second addition of trypsin was made to keep the gel soaked in trypsin solution, digestions were performed for 18 h at 37^{0}C. After digestion the peptides were extracted two times in 60% acetonitrile, 0.1% TFA solution, with extraction volumes of 20 to 50 μl. Combined eluates were subjected to partial drying in a speed-vac to decrease volume of a sample to 1/8 – 1/10 of the initial volume, which reduces the concentration of acetonitrile to less than 10%. Peptides extracted from gels were concentrated using ZipTips (μ-C18; Millipore), eluted in 65% acetonitrile containing the matrix α-cyano-4-hydroxycinnamic acid, and applied directly on a metal MALDI target. For FT-ICR and ESI Q-TOF mass spectrometry, peptides were eluted from ZipTips in 1% acetic acid, 65% acetonitrile, but without the matrix.

2.4. Peptide mass fingerprinting (PMF) and search for PTMs

PMF was performed using three different instruments, MALDI TOF/TOF, FT-ICR and ESI-QStar. The first set of data was generated using a Bruker Ultraflex MALDI TOF/TOF instrument (Bruker Daltonics, Bremen, Germany), as previously described [12,13]. In brief, spectra were collected in a positive mode after 200-300 shots, and were processed using FlexAnalysis software (Bruker Daltonics, Bremen, Germany). Mass range for collection of ions was set as 600 Da to 3,000 Da. Mass spectra were internally calibrated using tryptic peptides (842.51, 1045.56, and 2211.10 Da). The second set of PMF was performed using FT-ICR instrument (Bruker). In the FT-ICR MS experiments the peptides were electrosprayed on-line using a homebuilt instrument that controlled the direct infusion where helium gas at a pressure of 1.3 bars was used to push the

sample through a 30-cm fused silica capillary with an inner diameter of 25 μm [14]. One end of the capillary was lowered into the sample, and the other end, coated by a conductive graphite/polymer layer, was connected to ground, functioning as an electrospray needle. No sheath flow or nebulizing gas was used. The flow rate was 40 nl/min. The ion source was coupled to an Analytica atmosphere-vacuum interface (Analytica, Branford, CT). A potential difference of 2.0-3.0 kV was applied across a distance of 3-5 mm between the spraying needle and the inlet capillary. All mass spectra were acquired using a Bruker Daltonics (Billerica, MA) BioAPEX-94e superconducting, 9.4 T FT-ICR mass spectrometer in broadband mode. All FT-ICR spectra were externally calibrated using seven abundant peaks (m/z 438.26, 613.81, 820.47, 1013.60, 1342.64, 1623.79, and 1811.95) in a bovine serum albumin (BSA) tryptic digest spectrum. The reference peaks fit the calibration curve within 1 ppm error. Typically, 524,288 data points were acquired, adding a minimum of 128 spectra (3 min of acquisition time). Spectra were added as transients and data was collected on a computer running Xmass (Bruker Daltonics). The third data set was generated with Q-TOF (QStar, Applied Biosystems) instrument. The samples were electrosprayed on-line using a single-use capilars, and with the flow rate 100 nl/min. The MS spectra (m/z 300 to 1600) were acquired in the positive ion mode. Q-TOF instrument was externally calibrated with alcohol dehydrogenase (ADH) standard (693.36, 968.49, 1136.57, 1251.67, 16.81, 2019.07 and 2312.15). For internal calibration masses of tryptic peptides 842.51, 1045.56 and 2211.10 were used. A full list of observed peptides is presented in Supplementary Table S1. To identify a protein, we performed searches in the NCBInr sequence database using the ProFound search engine (http://65.219.84.5/service/prowl/profound.html). One miscut, alkylation and partial oxidation of methionine were allowed. Search parameters were set on mass tolerance less than 0.1 Da, no limitations of pI, limitation of Mr 0.0 - 100.0 kDa, and "mammalian" was selected for species search. Significance of the identification was evaluated according to the probability value, "Z" value and sequence coverage. Notably, probability value of 1.0e+00 (indicating 100% confidence that identification is correct) and Z value ensuring confidence of higher than 99.0% (Z values higher than 2.3). Use of both parameters ensured high confidence of identification.

FindMod tool was used to search for post-translationally modified peptides of Smad2. The following settings were used in FindMod searches with collected mass lists (Supplementary Table S1) – monoisotopic masses, tolerance +/- 0.1 Da, one missed cleavage by trypsin, oxidation of methionine and iodoacetamide treatment. The search was to identify Smad2 peptides with PTMs-characteristic mass shifts, and the 32 most common PTMs were considered; these PTMs were defined in FindMode as the most common based on results of mass spectrometry studies (Supplementary Table S2).

2.5. Fragmentation of peptides and detection of PTMs

Fragmentation of selected peptides was performed by post-source decay, *i.e.* "LIFT" in Bruker nomenclature on the Ultraflex MALDI TOF/TOF instrument. To improve fragmentation, chemically-assisted fragmentation and Lys-tagging methods were used [13,15,16]. In brief, Smad2 peptide obtained from the tryptic digestion was labeled with sulphonating CAF reagent. Prior to this reaction blocking of the ε-amino group of lysine residue is required (since the sulfonation reagent reacts with all primary amino groups). This blocking reaction was achieved in two different ways. A solution of O–methylisourea hydrogen sulphate (17.2 mg/ml in 0.25 M $NaHCO_3$, pH 10) which was allowed to react with tryptic peptides overnight at room temperature was used. This lysine blocking reaction adds 40.012 Da per Lys residue. Alternatively 2-methoxy-4, 5-dihydro-1H-imidazole was used (Lys Tag 4H). This reagent adds 68.037 Da specifically to the ε-amino group of lysine residue. The reaction was proceeded at 55^0C for 150 minutes, whereafter the sample was acidified and desalted by ZipTip C18. Both reagents are specific for the ε-amino group of lysine and thus leave the N-terminal amino group intact.

Sulphonation of the peptides was performed by pipeting up and down freshly prepared CAF labeling solution (1mg/1ml in 0.25 M $NaHCO_3$, pH 9.4) through a ZipTip and allowed to react with adsorbed sample at room temperature. The reagent washed out with 2 x 10 μl of 0.1% TFA. A preparation of 5% hydroxylamine in the same labeling solution was drawn through a ZipTip a few times in order to hydrolyze unwanted uptake of the CAF reagent. The sulphonated peptides were eluted by drawing 1-3 μl of 0.1 TFA/ 60% acetonitrile up and down the ZipTip. The sulphonation adds 135.983 Da to primary amines in the peptide. A peptide with C-terminal Arg will increase its mass by 135.983 Da. A peptide with a C-terminal Lys will increase by a total of 178.005 Da or 204.020 Da with 40.012 Da or 68.037 Da from the Lysine blocking and 135.983 Da from the sulphonation. Peptides with additional Lys will take up 40.012 Da or 68.037 Da/Lys each depending on the blocking method used. As MALDI TOF/TOF fragmentation delivered confident results, this sequencing approach was used.

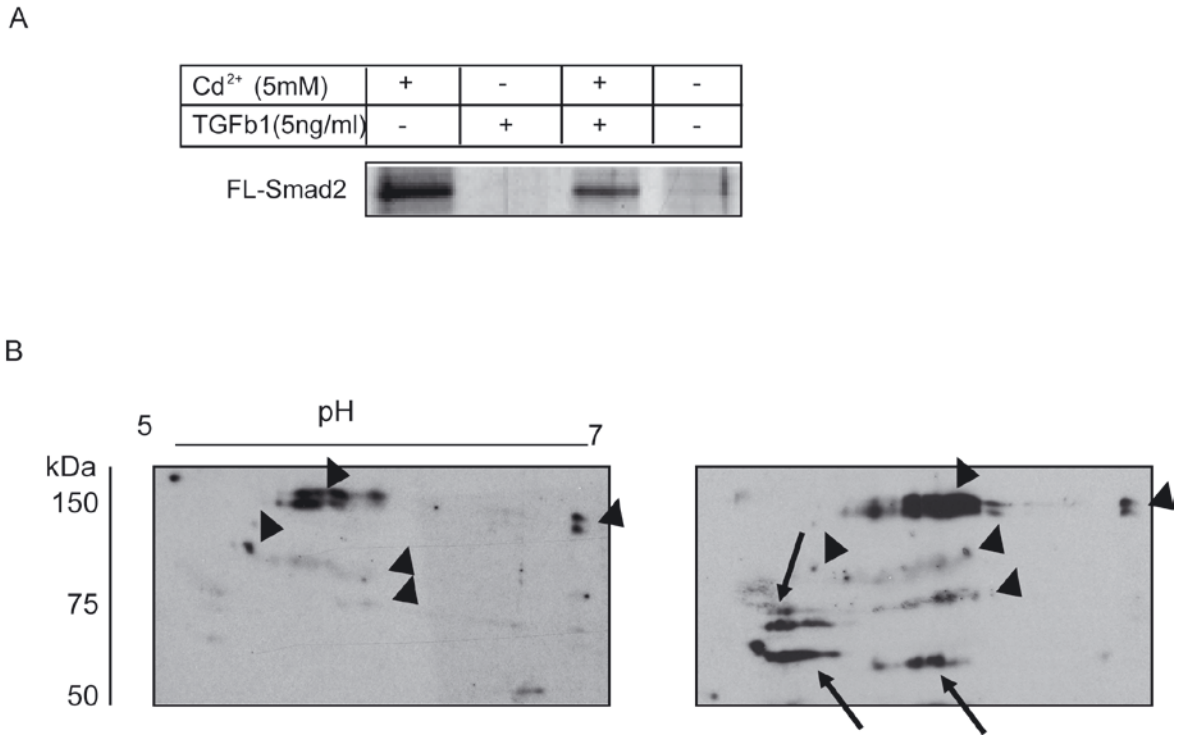


Figure 1. Expression of Flag-Smad2 in MCF7 cells.

(A) The image of a 1D SDS-PAGE gel, stained with silver, is shown. Migration position of Flag-Smad2 (FL-Smad2) and treatment of cells with $CdCl_2$ and TGFβ1 are indicated. (B) The images of immunoblots of Flag-Smad2 expressed in MCF7 cells and separated by 2-DE (right panel), as compared to the sample from MCF7 cells transfected with control empty vector (left panel). Arrows indicate migration of Flag-Smad2 forms. Arrowheads indicate migration of non-specific to Flag-Smad2 proteins, i.e. background. Migration positions of molecular mass markers and pH gradient are indicated.

Fragmentation spectra of selected peptides were interpreted manually, as expected sequences of candidate peptides could be predicted from PMF and predicted PTM-dependent mass shift. To confirm manual interpretation, fragmentation spectra were also analysed using Mascot search engine. Mass spectra were uploaded in Mascot, and peptides were identified with confidence of $p< 0.05$, *i.e.* Mowse score 80 for native 1213.84 Da and score 129 for 1613.83 Da peptides (Figure 4). As observed modifications of tyrosine residues (Figure 4) have not been described, MS/MS searches did not detect it, with decrease of overall confidence of identification (data not shown).

3. Results

3.1. Stable expression of Smad2 in human breast epithelial cells

Human Flag-Smad2 was constitutively transfected in MCF7 cells in pMEP4 vector under Cd^{2+}-inducible promoter. Inducible expression allowed maintaining cells without enhanced level of Flag-Smad2, and therefore not affecting the cell physiology. Flag-Smad2 was induced for only 4 h before treatment of cells. Stable transfection also ensured that the level of Flag-Smad2 expression was comparable in all cells, and did not vary, as it would be expected for transiently transfected proteins. Presence of the Flag tag at the N-terminus was reported not to interfere with Smad2 functions [17]. MCF7 cells are often used as a model of human breast epithelial cells, and represent transformed, *i.e.* cancer, cells. MCF7 have also intact TGFβ specific receptors. Expression of human Flag-Smad2 in these cells allowed the study Smad2 PTMs in an environment similar to the human breast cancer situation.

Flag-Smad2 expressed in MCF7 cells was detected as a distinct protein band in the lane of proteins extracted from $CdCl_2$ treated cells (Figure 1A). Cells expressing Flag-Smad2 were treated or not with 10 ng/ml of TGFβ1, and compared to untreated control cells. Migration position of expressed Flag-Smad2 in SDS-PAGE was the same as for endogenous Smad2. The identity of induced Flag-Smad2 was confirmed by immunoblotting of the whole cell extracts prepared as in Figure 1A, with anti-Flag and anti-Smad2 antibodies (data not shown). We observed multiple forms of Flag-Smad2, when the proteins were resolved by two-dimensional gel electrophoresis (2-DE; Figure 1B). The limited number of multiple forms of Flag-Smad2 suggested

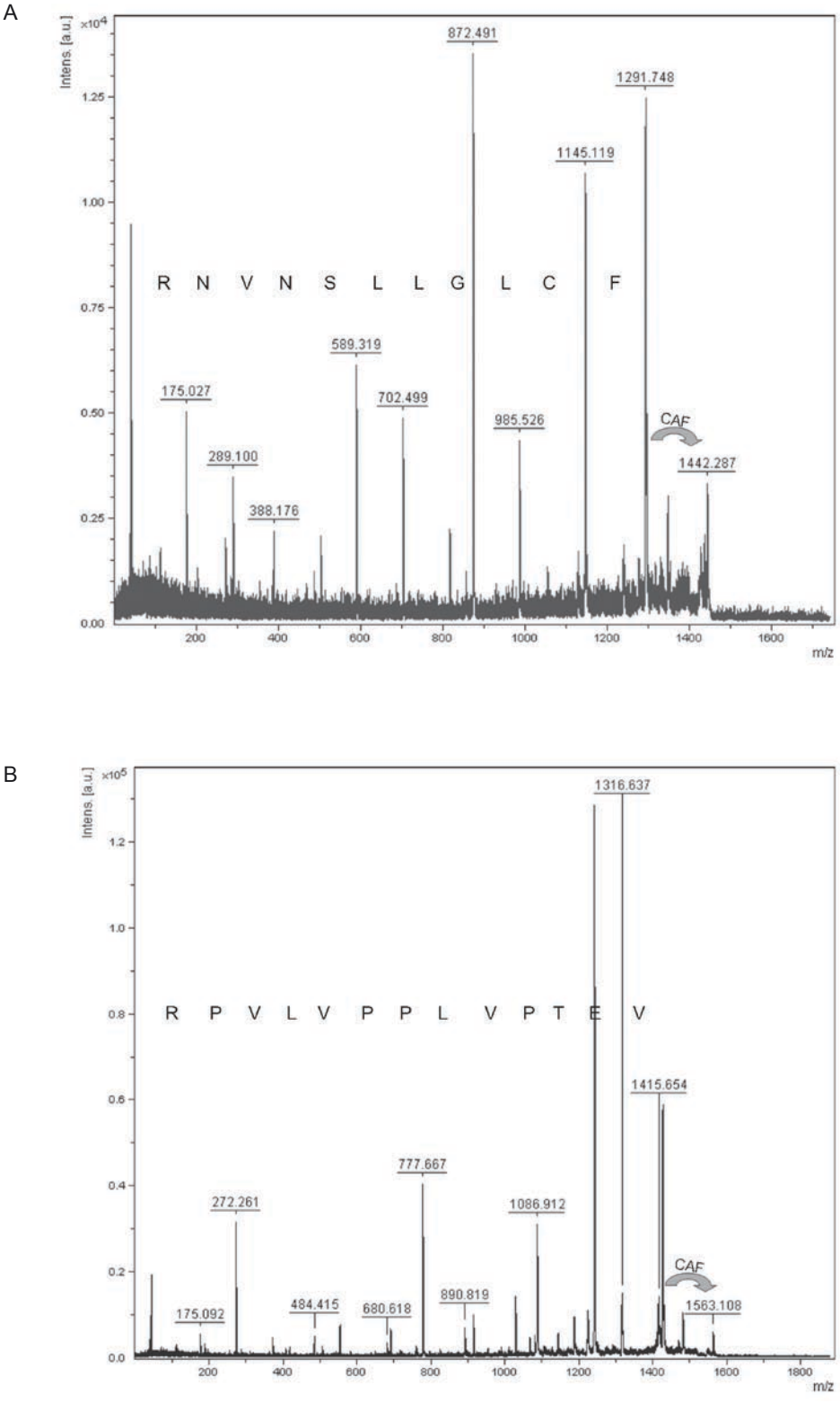


Figure 2. Fragmentation of selected peptides confirmed identification of Smad2.

Fragmentation spectra of 1442.287 Da (A) and 1563.108 Da (B) peptides are shown. Identified amino acid sequences are indicated in the images. Fragments with indicated masses were used for determination of the sequence. Parental ions and CAF-related shift are indicated. MALDI TOF/TOF platform was used.

No:	Peptide a)	Modification b)	Modified Residue	Position c)	Mass Observedd)	Theoretical Mass of Modified Peptide d)	Delta mass observed-theoretical PTM d)	Delta mass specific to PTM d)	Type of Instrument
1	RLLGWK	CITR	R	14	773.48	772.48	0.022	0.98	MALDI TOF, ESI FTICR
2	IPPGCNLK	NTRY	C	381	870.54	841.46	0.049	28.99	MALDI TOF, ESI FTICR
3	IPPGCNLK	TRIMETH	C	381	940.6	898.48	0.071	43.05	ESI FTICR, ESI QTOF
4	WCEKAVK	TRIMETK	K	43	921.48	863.44	0.013	59.04	MALDI TOF, ESI FTICR
5	CSSMS	GERA	C	463	802.40	530.15	0.053	272.25	MALDI TOF,ESI FTICR, ESI QTOF
6	CSSMS	PALM	C	463	825.31	587.18	0.088	239.22	MALDI TOF,ESI FTICR, ESI QTOF
7	MCTIRMSFVK	OCTA	S	417	1398.66	1272.62	0.04	126.10	MALDI TOF, ESI QTOF
8	QTVTSTPCWIELHLNGPLQWLDK	DIMETH	C/Q	429(Q)/436(C)	2707.48	2679.35	0.025	28.03	MALDI TOF, ESI FTICR
9	QTVTSTPCWIELHLNGPLQWLDK	DIMETP	P	435/445	2707.48	2679.35	0.025	29.03	MALDI TOF, ESI FTICR

Table 1. PTMs identified by PMF of Flag-Smad2.

a) Peptide sequence is shown, with modified amino acid residue in bold.
b) For abbreviation of modifications, see Supplementary Table S2
c) Position(s) of modified residues, with alternative PTM sites if applicable, are indicated.
d) Mass is in Da. Due to the lower resolution and accuracy of MALDI TOF MS platform, as compared to ESI FT-ICR, accuracy corresponding to MADLI TOF MS is shown.

that PTMs are present mostly in limited combinations in the same molecule of Flag-Smad2 (Figure 1B). For studies of PTMs, we used Flag-Smad2 which migrated in SDS-PAGE at the similar position as the endogenous protein. PTMs which introduce a significant mass shift, e.g. ubiquitylation, were therefore not considered in this study. For MALDI TOF mass spectrometry, we performed two experiments, that is to say, two preparations of Flag-Smad2, with two repeats for the TGFβ1-treated cells and three repeats for non-treated conditions in each experiment. We collected 70 to 100 mass peaks in each experiment for Flag-Smad2 in TGFβ1-treated and not-treated cells. Mass lists are shown in Supplementary Table S1. We also used mass spectrometers with ESI technique for ionization of peptides. With ESI FT-ICR mass spectrometry we detected 77 and 92 peptides from TGFβ1-treated and 93 and 100 peptides from non-treated cells, in two experiments, respectively. Mass spectrometry analysis with a Q-TOF instrument (QStar) provided a list of 34 peptides from TGFβ1-treated and 158 peptides from non-treated cells (Supplementary Table S1). PMF with the detected peptide masses identified Smad2 with high confidence (Z>2.4, coverage >25 %, up to 11 Smad2 peptides detected, search tolerance 0.05 Da, no restriction on species, pI and molecular mass). Fragmentation of selected peptides confirmed identification of Smad2 (Figures 2, 3 and 4).

3.2. Identification of PTMs of Flag-Smad2 with peptide mass fingerprinting data

We noted that a significant number of detected Flag-Smad2 peptides were not matched to non-modified Smad2 peptides. One of the reasons maybe that the presence of modifying groups have changed the molecular mass of a peptide, and therefore excluded this peptide from assigning it to the Smad2. These mass shifts are specific for specific modifications, and a modified peptide may be identified by addition or deduction of a given PTM mass value. To identify potentially modified peptides of Smad2, we performed searches with the FindMod tool (http://expasy.org/tools/findmod/). We used the mass lists generated by us, and performed search for Smad2 peptides which have mass shifts specific for particular modifying groups at a particular amino acid residue in the peptide. For example, observation of an increase in 42.037 Da mass shift of a lysine-containing peptide may indicate acetylation of that peptide. We searched for the 32 most common PTMs (Supplementary Table S2) with 3 sets of mass lists generated by MALDI TOF (two experiments), FT-ICR (two experiments) and Q-TOF (one experiment) mass spectrometry. We considered only those peptides with PTMs which were observed in at least two out of

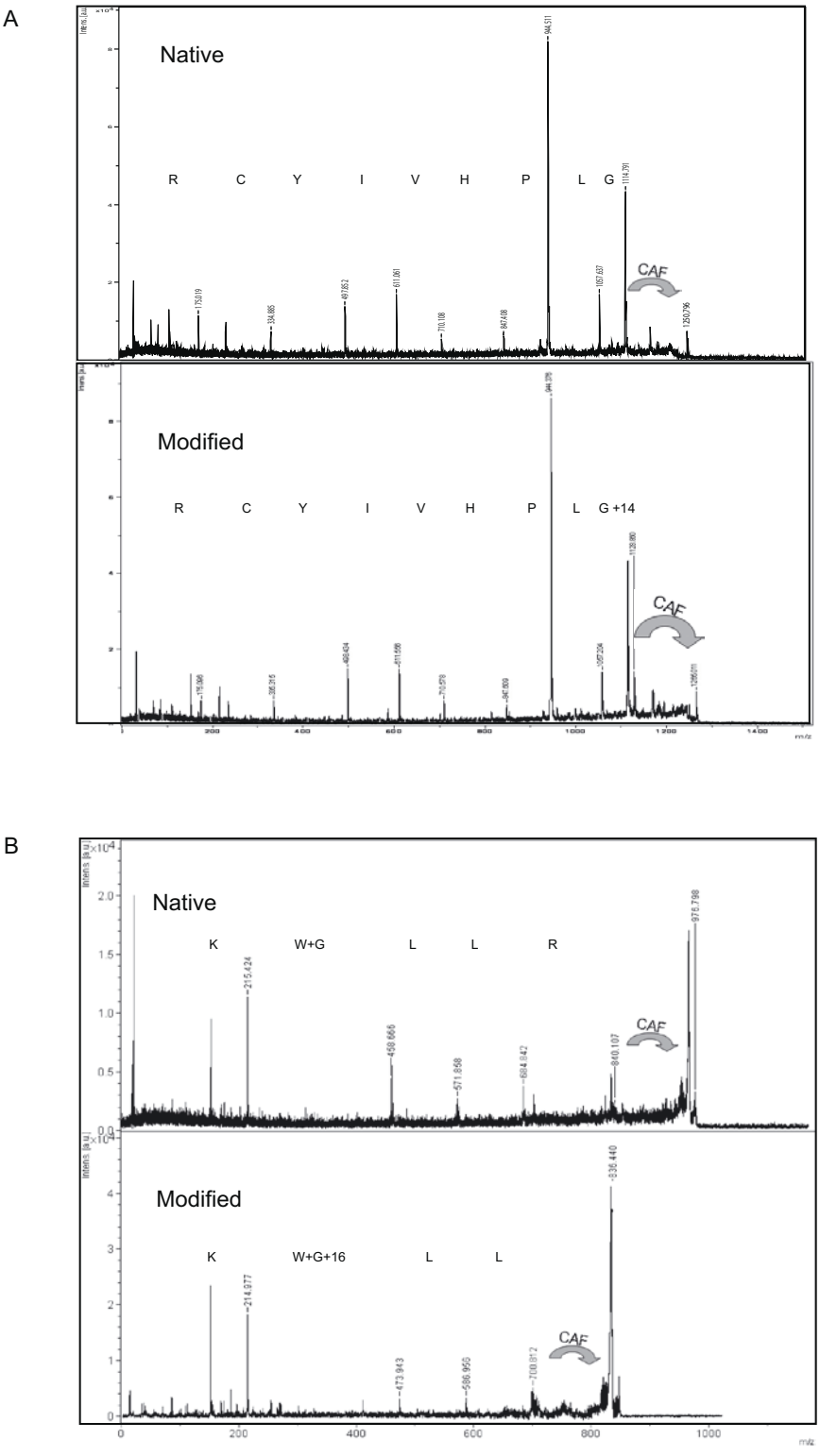
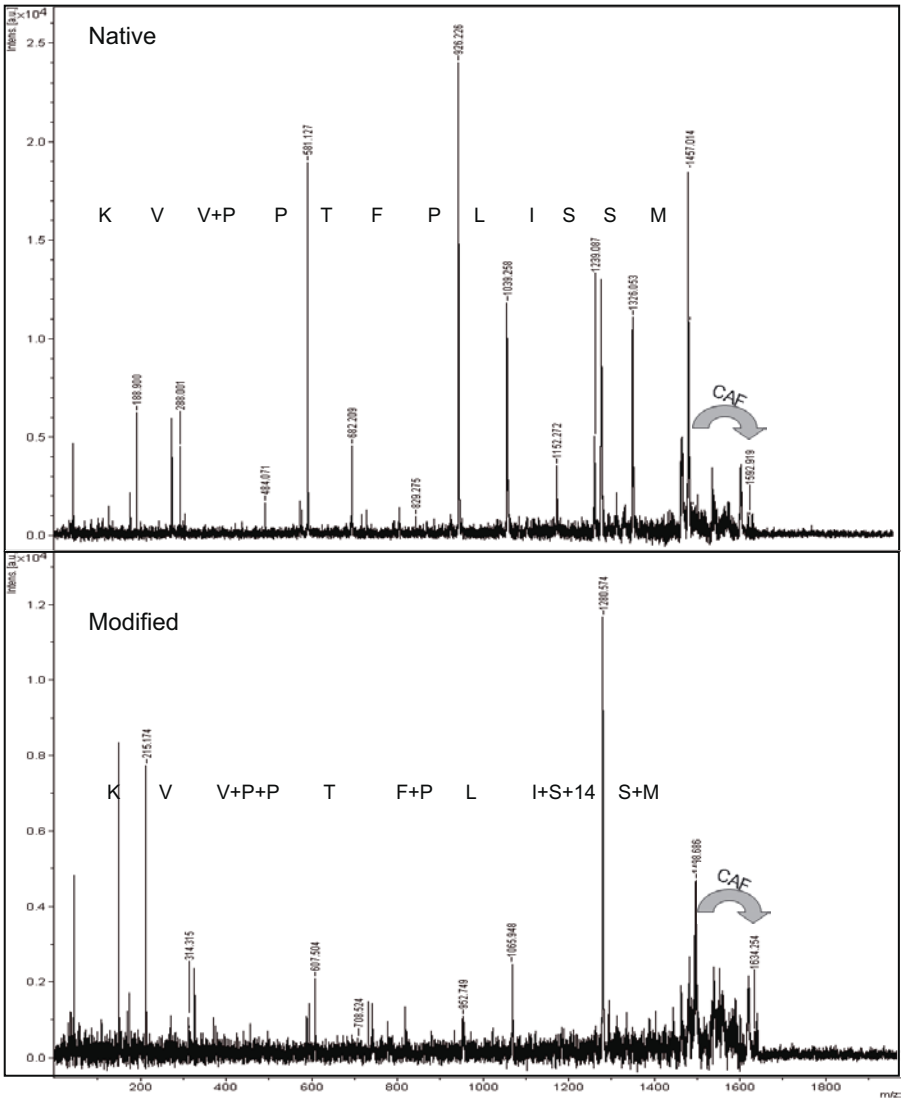


Figure 3. Flag-Smad2 was methylated at glycine Gly122, serine/isoleucine Ser3/Ile4 and hydroxylated at tryptophan Trp18.

Fragmentation spectra of 1114.79 Da and 1128.88 Da (A), 840.10 and 700.81 Da (B), 1457.01 and 1498.68 Da (C) peptides are shown. Sequencing of 1114.79 Da peptide showed that it was a Gly122-Arg130 peptide of Smad2, and the same peptide was observed with methylated Gly122 (A), as molecular mass corresponding to Gly position was increased with 14 Da. Hydroxylation of tryptophan residue Trp18 was observed by a mass shift of 16 Da at the Trp position in Leu15-Lys19 peptide (B) Methylation at Ser3-Ile4 position was observed in the N-terminal peptide of Smad2 (C). MALDI TOF/TOF platform was used.

C



continued **Figure 3.** Flag-Smad2 was methylated at glycine Gly122, serine/isoleucine Ser3/Ile4 and hydroxylated at tryptophan Trp18.

Fragmentation spectra of 1114.79 Da and 1128.88 Da (A), 840.10 and 700.81 Da (B), 1457.01 and 1498.68 Da (C) peptides are shown. Sequencing of 1114.79 Da peptide showed that it was a Gly122-Arg130 peptide of Smad2, and the same peptide was observed with methylated Gly122 (A), as molecular mass corresponding to Gly position was increased with 14 Da. Hydroxylation of tryptophan residue Trp18 was observed by a mass shift of 16 Da at the Trp position in Leu15-Lys19 peptide (B) Methylation at Ser3-Ile4 position was observed in the N-terminal peptide of Smad2 (C). MALDI TOF/TOF platform was used.

three datasets. Due to differences in resolution and accuracy of the employed mass spectrometry platforms, with FT-ICR MS having the highest resolution and accuracy, and referring to peptide identification with at least two platforms, we show results from the lower-resolution/accuracy instrument out of the employed platforms (Table 1).

The FindMod tool identified 12 peptides in Flag-Smad2 mass list, which corresponded to peptides with PTMs. Nine of PTMs were observed in two MS datasets and are shown in Table 1. We identified acetylation at Lys19, thus confirming previous observation by Simonsson *et al.* [5]; as acetylation at Lys19 was observed with MALDI TOF MS only, it is not included in the Table 1. We also observed phosphorylation at the C-terminal peptide which contains receptor-phosphorylatable serine residues, and at the peptide containing PKC-phosphorylated Ser110 (by MALDI TOF MS, data not shown), which were reported earlier [17,18]. Of the novel PTMs seen, we observed palmitoylation at Cys412 and Cys463, and geranylation at Cys463. Methylation was observed in three peptides

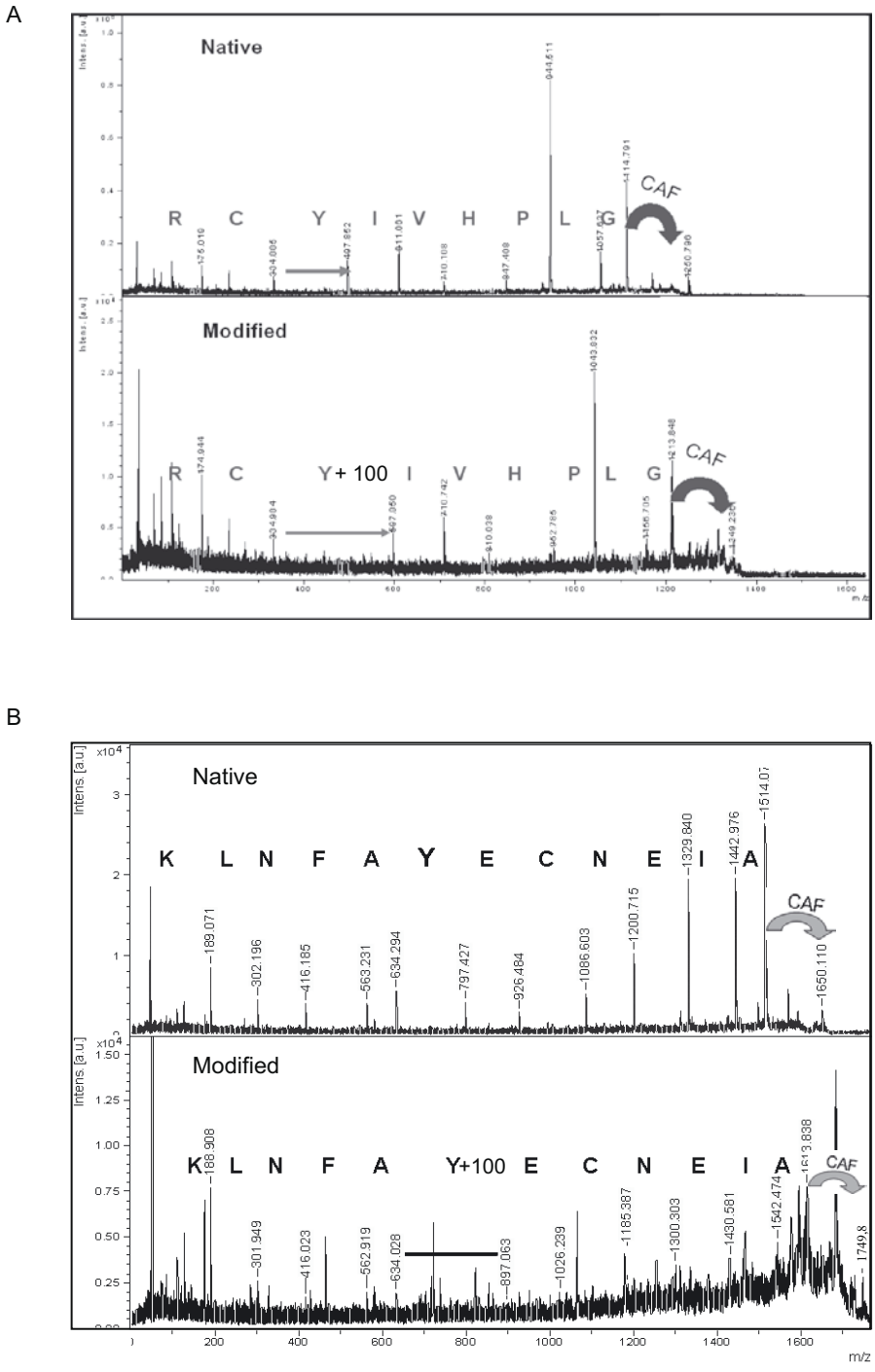


Figure 4. Tyrosine modifications of Flag-Smad2.

Mass spectra of non-modified and tyrosine-modified peptide Ala145-Lys156 (A) and Gly122-Arg130 (B) are shown. Fragmentations at tyrosine positions indicated mass shifts of 263.0+/-0.8 Da, which indicate Tyr+100.0 Da. MALDI TOF/TOF platform was used.

with possible sites of modification at Cys381 or Lys383 for the peptide Ile377-Lys383, Lys43 or Lys46 for the peptide Trp40-Lys46, and Trp429 or Cys436 or Lys451 for the peptide His429-Lys451. As methylation of Lys residues should interfere with the efficiency of tryptic digestion, Lys 383, Lys 46 and Lys 451 are unlikely to be methylated. Therefore, methylation is potentially assigned to Cys381, Lys43, Trp429 and Cys436. We also observed citrullination at Arg14, O-octanoyl Ser417, and S-nitrosylation at Cys381 (Table 1). Thus, PMF confirmed previously reported modifications and detected novel PTMs for Smad2.

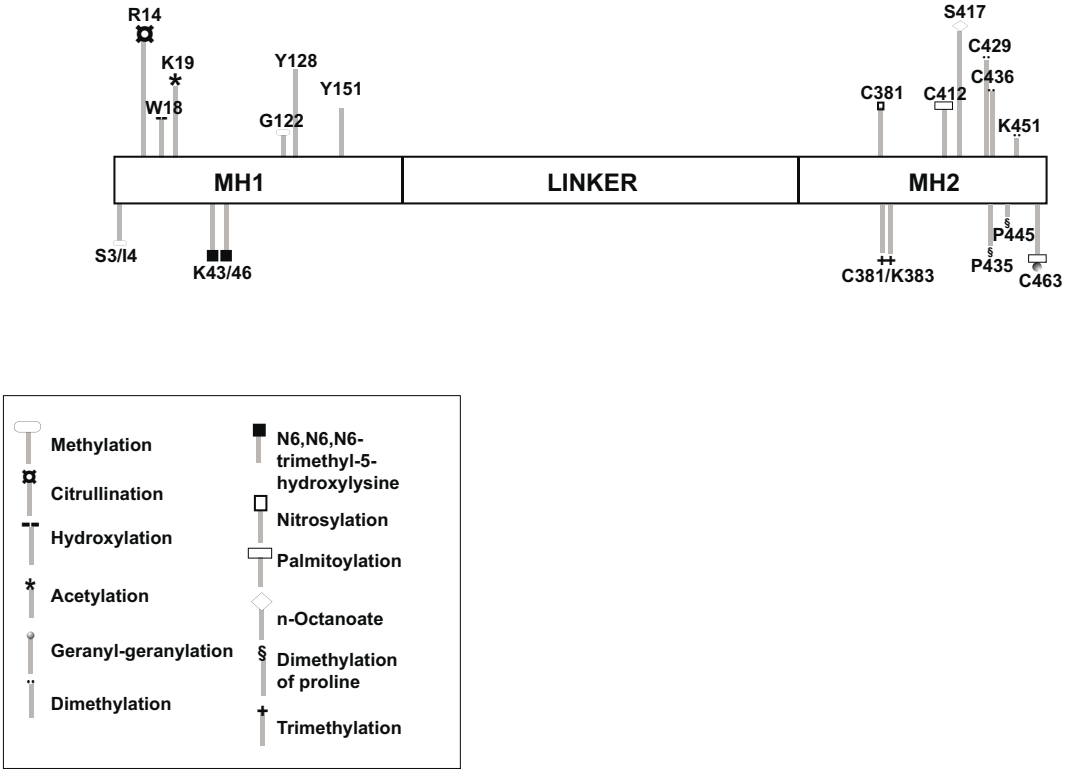


Figure 5. Presentation of Smad2 PTMs identified by PMF and sequencing.

Location of reported in this article PTMs in human Smad2 is shown. MH1, linker and MH2 domains are indicated. For description of PTMs, see Table 1 and the text.

3.3. Identification of PTMs of Flag-Smad2 with peptide sequencing by fragmentation

To validate that potentially modified peptides originate from Smad2, we performed fragmentation of selected peptides with molecular masses corresponding to modified peptides. We applied Post-Source Decay (PSD) using the MALDI TOF/TOF instrument. Chemically assisted fragmentation (CAF) and Lys-tagging were used to facilitate fragmentation and detection of ions [13].

Fragmentation of a peptide of Mr 1114.79 Da confirmed identification of Smad2 (Gly122-Arg130 peptide; Figure 3A). Fragmentation of the 1128.88 Da peptide showed a mass shift corresponding to methylation of glycine residue Gly122 (Figure 3A). Fragmentation of the 700.81 Da peptide indicated hydroxylation at the tryptophan residue Trp18 in the Leu15-Lys19 peptide; a non-modified peptide with a miscut at an arginine residue was also identified (Figure 3B). We observed methylation at the N-terminus of Smad2 at Ile4 or Ser3 residues; but due to the low fragmentation efficiency between Ser3 and Ile4, both residues are candidates for methylation (Figure 3C).

Fragmentation of the 1613.83 Da and 1213.84 Da CAF modified peptides indicated a molecular mass shift of 99.866 Da for tyrosine residues Tyr128 and 100.035 for Tyr151, as compared to Tyr 163.063 Da (Figure 4). The mass of 263 Da does not correspond to a mutation at these positions. PTMs which would render Tyr to 263 Da are not in the list of the reported PTMs (http://www.abrf.org/index.cfm/dm.home, www.expasy.org). The mass of 263 Da for tyrosine residue may be due to formation of a covalent link with S-cysteinyl (3'-(S-cysteinyl)-tyrosine). However, this possibility has to be explored further.

TGFβ1 treatment of Smad2-expressing cells resulted in no detection of the peptide with palmitoylation of Cys463 at the C-terminus of Smad2. We did not observe the peptide with acetylated Lys43 or Lys46 upon treatment with TGFβ1, while other acetylated peptides were observed. All other reported PTMs here were not affected by the TGFβ1 treatment of cells. The lack of pronounced TGFβ-dependencies for most of the detected PTMs may be explained by the presence of PTMs in both TGFβ1-treated and non-treated cells. The lack of TGFβ-dependency suggests that these PTMs may be convergence points for other non-TGFβ signaling mechanisms.

4. Discussion

Smad2 is a transcription factor crucial for TGFβ signaling, and its phosphorylation, acetylation and ubiquitylation have earlier been described as potent mechanisms modulating Smad2 activity [1-3,5-8,17,18]. Phosphorylation of Smad2 is crucial for its activation, interaction with other proteins and DNA, and localization in cells [1-3,5-8]. Ubiquitylation was found to be the main mechanism regulating not only degradation of Smad2, but also affecting its transcriptional activity [1-3,7]. Using mass spectrometry, we identified PTMs previously not observed in Smad2. We performed analysis of Smad2 peptides using three different mass spectrometry approaches, *i.e.* MALDI TOF/TOF, FT-ICR and Q-TOF. These platforms have different resolution and accuracy, with FT-ICR having the highest. To eliminate false identifications, we used stringent conditions for peptide analysis in each approach separately, as well as considered PTMs which were observed in at least two approaches. This analysis of mass spectrometry data eliminated more than 85% of the detected peptides, as not conforming requirements for significant and confident identification of PTMs.

Our observation of a number of PTMs of Smad2 is not an exception, but a confirmation of the high frequency of various PTMs in proteins [4,9,10]. We observed previously described PTMs, such as acetylation at Lys19 [5], phosphorylation at the C-terminal serine residues [1,17] and Ser110 [18]. We did not observe PTMs which would change significantly the molecular mass of Smad2, e.g. more than 10,000 Da, due to the analysis of the Smad2 migrating in SDS-PAGE at the position of a non-modified protein. It also has to be noted that PTMs observed by us do not necessarily occur at the same Smad2 molecule, and the studied sample may be a combination of Smad2 with various patterns of PTMs.

PTMs observed by us are located in areas of Smad2 which are involved in regulation of transcriptional activity and protein-protein interactions of Smad2 (Figure 5). As an example, palmitoylation at the C-terminus may enhance membrane tethering of Smad2 in the proximity to TGFβ receptors, and facilitate Smad2 activation. Acetylation of Smad2 has been shown to be important for regulation of Smad2 degradation and transcriptional activity [1,5-7]. Our report of novel sites of acetylation, methylation, hydroxylation, palmitoylation, geranylation and citrunillation open for further studies of their role in Smad2-dependent TGFβ signaling.

Acknowledgments

We are grateful to Lars Rönnstrand and Kristina Masson for help with QStar MS. This work was supported in part by Karolinska Biomics Center, Dep. of Oncology-Pathology, Ludwig Institute for Cancer Research, grants from the Swedish Cancer Foundation, the Swedish Research Council, INTAS, RTN EpiPlast Carcinoma, and Hiroshima University to S.S, and of the Swedish Research Council grants #629-2002-6821, #342-2004-3944 and #621-2005-5379 to J.B.

References

[1] Feng X.-H., Derynck R., Specificity and versatility in TGF-b signaling through Smads, Annu. Rev. Cell Dev. Biol., 2005, 21, 659-693

[2] Siegel P.M., Massague J., Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer, Nat. Rev. Cancer, 2003, 3, 807-820

[3] Goumans M.J., Mummery C., Functional analysis of the TGFbeta receptor/Smad pathway through gene ablation in mice, Int. J. Dev. Biol., 2000, 44, 253-265

[4] Witze E.S., Old W.M., Resing K.A., Ahn N.G., Mapping protein post-translational modifications with mass spectrometry, Nat. Methods, 2007, 4, 798-806

[5] Simonsson M., Kanduri M., Grönroos E., Heldin C.H., Ericsson J., The DNA binding activities of Smad2 and Smad3 are regulated by coactivator-mediated acetylation, J. Biol. Chem., 2006, 281, 39870-39880

[6] Tu A.W., Luo K., Acetylation of Smad2 by the co-activator p300 regulates activin and transforming growth factor bets response, J. Biol. Chem, 2007, 282, 1187-1196

[7] Lo R.S., Massagué J., Ubiquitin-dependent degradation of TGF-beta-activated Smad2, Nat. Cell Biol., 1999, 1, 472-478

[8] Moustakas A., Souchelnytskyi S., Heldin C.H., Smad regulation in TGF-beta signal transduction, J. Cell Sci., 2001, 114, 4359-4369

[9] de Hoog C.L., Mann M., Proteomics, Annu. Rev. Genomics Hum. Genet., 2004, 5, 267-293

[10] Jensen O.N., Interpreting the protein language using proteomics, Nat. Rev. Mol. Cell Biol., 2006, 7, 391-403

[11] Wilkins M.R., Gasteiger E., Goole, A., Herbert B., Molloy M.P., Binz P.A., et al., High-throughput mass spectrometric discovery of protein post-translation-

al modifications, J. Mol. Biol., 1999, 289, 645-657
[12] Dubrovska A., Souchelnytskyi S., Efficient enrichment of intact phosphorylated proteins by modified immobilized metal-affinity chromatography, Proteomics, 2005, 5, 4678-4683
[13] Conrotto P., Hellman U., Lys Tag: an easy and robust chemical modification for improved de novo sequencing with a matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometer, Rapid Commun. Mass Spectrom., 2008, 22, 1823-1833
[14] Nilsson S., Wetterhall M., Bergquist J., Nyholm L., Markides K.E., A simple and robust conductive graphite coating for sheathless electrospray emitters used in capillary electrophoresis/mass spectrometry, Rapid Commun. Mass Spectrom., 2001, 15, 1997-2000
[15] Keough T., Youngquist R.S., Lacey M.P., A method for high-sensitivity peptide sequencing using post-source decay matrix-assisted laser desorption ionization mass spectrometry, Proc. Natl. Acad. Sci. USA, 1999, 96, 7131-7136
[16] Keough T., Lacey M.P., Youngquist R.S., Derivatization procedures to facilitate de novo sequencing of lysine-terminated tryptic peptides using post-source decay matrix-assisted laser desorption/ionization mass spectrometry, Rapid Commun. Mass Spectrom., 2000, 14, 2348-2356
[17] Souchelnytskyi S., Tamaki K., Engström U., Wernstedt C., ten Dijke P., Heldin C.H., Phosphorylation of Ser465 and Ser467 in the C terminus of Smad2 mediates interaction with Smad4 and is required for transforming growth factor-beta signalling, J. Biol. Chem., 1997, 272, 28107-28115
[18] Yakymovych I., ten Dijke P., Heldin C.H., Souchelnytskyi S., Regulation of Smad signaling by protein kinase C, FASEB J., 2001, 15, 553-555