UPPSALA UNIVERSITET

2021-05-21

# *Discovery of peptides in Chinese medicinal plants*

Mostafa Afshar

**Degree Project in Pharmacognosy, 30 hp, Spring 2021**

Supervisor: Erik Jacobsson

Examinator: Ulf Göransson

Pharmacognosy

Department of Pharmaceutical Biosciences

Faculty of Pharmacy

Uppsala University

# Contents

## Abstract

In this study, 20 well-known Chinese medicinal plants are included. Traditionally, these medicinal herbs were consumed for different diseases but have one application in common, treatment of rheumatism arthritis. Peptides are pharmacologically attractive substances and the reason behind this work as there is only one study on peptide content of included herbs. Commonly, plant-derived peptides are cysteine-rich peptides.

The plants were extracted in series by 60%, 30% and 10% AcN in $H_2O$ and FA. Then, the combined extracts were purified by SPE and SEC. Additionally, reduction by DDT, alkylation by IAM, and digestion by trypsin performed. UPLC-QToF was used for separation and identification of potential peptides.

The investigation resulted in finding peptides in *Coix lacryma-jobi L*, *Atractylodes lancea*, *Astragalus membranaceus.* Peptides found in these species are mainly in the range 2,5-5 kDa. This finding raises new questions about the function of peptides in these species and the possibility of having pharmacological activity.

## Background

This work is part of a big project, a collaboration between Uppsala University, Karolinska Institute, and Guangdong Provincial hospital of Chinese Medicine. In this study, 20 well-known Chinese Medicinal Plants (CMPs) are included. All these herbs have been previously investigated phytochemically and pharmacologically. Traditionally, these herbal medicines were consumed for different diseases but have one application in common, treatment of rheumatoid arthritis (RA). Traditionally, an amount of 2-15 g of these herbs are decocted with water as an oral dose. In some cases, they are made into pills or powder for oral usage. Exceptionally, fresh parts of *Artemisia Annua* are wringed to extract for taking orally. *Kadsura Heteroclita* and *Rehmannia glutinosa* are even made into wine. Furthermore, these plants are included in different Chinese medicinal formulas (1).

Except for one species, there are almost no studies on peptide contents of these herbs. Peptides are attractive substances pharmacologically and the reason behind this investigation. The specific objective of this study was to investigate Chinese medicinal herbs in search for peptides. Additionally, found peptides identified and characterised.

### Peptides

Peptides are defined as polymers of up to 51 amino acids (AAs) . Peptide in plants functions as a defence mechanism like Antimicrobial peptides (AMPs), cell division control, reproductive mechanism etc (2). Commonly, plant-derived peptides are cysteine-rich peptides (CRPs). Beside being nutritious, peptides can show bioactivity e.g., anti-inflammatory, anti-fungal, anti-microbial, anti-cancer, etc (3,4). Most peptides originate from inactive precursor proteins. Peptides are classified into CRPs and post-translationally modified peptides (PTM -peptides). PTM-peptides can have modifications such as tyrosine sulfation, glycosylation, and proline hydroxylation (2). CRPs typically consist of 2-16 cysteine (Cys) residues with different characteristics and arrangement (5).

Peptides have commonly good solubility in H2O and a little acid can increase dissolution. For more hydrophobic peptides, acetonitrile (AcN) can be added to H2O/acid. Peptides containing many lipophilic residues can aggregate in H2O. In this case, H2O/AcN is not enough and higher concentration of acid, or Dimethylformamide (DMF) or Dimethyl Sulfoxide (DMSO) as solvent is needed for dissolution. DMF and DMSO are high-boiling solvents which are difficult to evaporate and can get early elution in HPLC. Usually, laboratories use H2O/AcN/0,1 % Formic Acid (FA) as eluent. Trifluoroacetic Acid (TFA) is a better counterion and gives higher column performance than FA, while FA gives lower suppression of the MS-ionization in contrast to TFA. Normally, C18-columns are used in HPLC for small- to medium-sized peptides and C4/C8-columns for medium- to large-sized (6).

Mass spectrometry (MS) is the most important technique for identification and quantification of proteins/peptides. Disulfide bonds between Cys give the peptide stability and functional structure but makes the identification difficult due to irregular fragmentation in MSMS. Reduction of peptides results in free sulfhydryl groups of Cys. Sulfhydryl groups are reactive and tend to oxidize spontaneously with each other. To prevent this issue, alkylation is needed. The most common alkylating compounds are Iodoacetic Acid (IAA), Iodoacetamide (IAM), N-Ethylmaleimide (N-EM) and acrylamide (7).

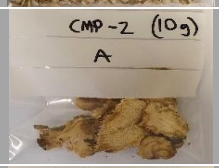
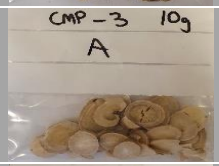
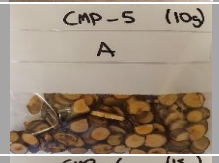
Trypsin is the gold standard for digestion of protein/peptide due to its high cleavage specificity and stability as well as being beneficial to have a basic AA in the C-terminal to facilitate the sequencing. Trypsin is a serine protease that cleaves the peptide after arginine and lysine at C-terminal. The tryptic cleavage leads to peptide sequence in an ideal range for MS-fragmentation. For identification, the spectra are compared to the plant's transcriptome. In a process called autolysis, trypsin can hydrolyse itself, resulting in additional fragments that interfere with peptide identification. However, $Ca^{2+}$ ion present in solution can bind to the enzyme and prevent autolysis (8,9). Modified trypsin does not interfere as much and is used in this work.

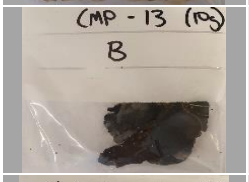
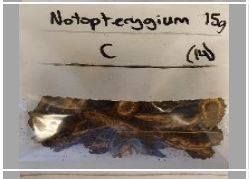
## Rheumatoid arthritis

RA is a chronic autoimmune and inflammatory disease. In RA, the immune system attacks healthy cells of joints and causes pain, swelling, and in later stages damage and cause deformity in joints as well as in bones. The cause of the disease is still unknown, but studies show genetic factors and lifestyle (e.g., smoking, alcohol consumption) can increase risk for RA. This disease affects not only an individual but also is a social problem by its cost for the society (10). According to World Health Organisation (WHO) the prevalence is 0.3-1% and is higher in women than men and higher in developed countries (11). The aim of RA treatment is to prevent joint and bone damages, pain, and disability. The treatment divides in physiotherapy and medication. RA patients usually are medicated by anti-inflammatory, analgesics, glucocorticoids, and immunosuppressive drugs e.g., Non-Steroidal Anti-Inflammatory Drugs (NSAID), Disease Modifying Anti-Rheumatic Drugs (DMARD), Sulfasalazine, TNF-α inhibitors (12).

Plant species included in this study is presented in table 1. For more information about plant species see Appendix 1.

Tabell 1: List and info about included species.

| CMP number | Species | Family | Organ | |
|---|---|---|---|---|
| | | | Traditionally used | Used in project |
|   | *Coix lacryma-jobi* L.var.ma-yuen (Roman.) Stapf | Poaceae | Seed | Seed |
|   | *Angelica sinensis* (Oliv.) Diels | Apiaceae | Root | Root |
|   | *Paeonia lactiflora* PalL. | Ranunculaceae | Root | Root |
|   | *Ephedra sinica* Stapf | Ephedraceae | Stem | Stem |
|   | *Cinnamomum cassia* Presl | Lauraceae | Bark | young stem |
|   | *Glycyrrhiza uralensis* Fisch. | Fabaceae | Root, Rhizome | Root |
|   | *Atractylodes lancea* (Thunb.) DC. | Asteraceae | Rhizome | Root |
|   | *Astragalus membranaceus* (Fisch.) *Bge.var.mongholicus* (Bge.) | Fabaceae | Root | Root |
|   | Yeast cultured on *Polygonum hydropiper L.*, *Artemisia annua L.*, *Xanthium Sibiricum* Patr. *Vigna umbellata* Ohwi et Ohashi, *Prunus armeniaca* L.var.ansu Maxim., wheat | Polygonaceae Asteraceae Asteraceae Fabaceae Rosaceae | Root seed | Root seed |
|   | *Morinda officinalis* How | Rubiaceae | Root | Root |

| CMP number | Species | Family | Organ | |
|---|---|---|---|---|
| | | | Traditionally used | Used in project |
|  | *Salvia miltiorrhiza* Bge. | Lamiaceae | Root, Rhizome | Root |
|  | *Dioscorea nipponica* Makino | Dioscoreaceae | Rhizome | Root |
|  | *Rehmannia glutinosa* Libosch. | Scrophulariaceae | Root | Root |
|  | *Notopterygium incisum* Ting ex H.T.Chang | Apiaceae | Root, Rhizome | Root |
|  | *Gentiana macrophylla* Pall. | Gentianaceae | Root | Root |
|  | *Kadsura heteroclita* (Roxb) Craib | Schisandraceae | stem | Root |
|  | *Morus alba L.* | Moraceae | All parts | Stem |
|  | *Ligusticum chuanxiong* Hort. | Apiaceae | Rhizome | young stem |
|  | *Aucklandia lappa* Decne. | Asteraceae | Root | Root |
|  | *Aconitum carmichaelii* Debx. | Ranunculaceae | Root | Root |

# Method and Material

## Screening

### Work-up

Extracts were previously prepared by two different solvent, A and B. Solvent A consists of 50:50 DCM and MeOH and solvent B, 60:40 AcN and H2O. A small amount of extract was transferred to a 1,5 mL Eppendorf tube (see Appendix 2) and dissolved in methanol (Honeywell, > 99,9%) to a concentration of 10 mg/mL. To dissolve properly, all samples vortexed and in some cases sonicated. To sediment eventual particles samples were centrifuged by 16060 RCF in 10 min. Samples prepared for injection to HPLC-LTQ by transferring 150 µL of samples to glass vials.

### HPLC-MS

In HPLC system used gradient mode and the solvent used as mobile phase were A: 5% AcN, 0,1% FA in H2O and B:95% AcN, 0,1% FA in H2O, gradient: 0-3 min A, 3-18 min reducing A to 0% and increasing B to 100%, 19-28 min A. The HPLC column was a Kinetex 2.6 µm, XB-C18 100 Å, 100*3 mm. Mass spectrometer was a Finnigan LTQ from Thermo Fisher. The system was set on an analysis time of 28 min, flow rate 0,3 mL/min and injection volume 1 µL. Methanol was used as blank. Spectra have been analysed and those samples with multi- charge signals, samples with many signals in the area above 1000 m/z and samples belonging to families with known peptide picked for further investigation.

### UPLC-QToF

Samples CMP-4,5,6,15,19 B and CMP-6,8,12,13, 16 A with promising LTQ-spectra have been chosen first to inject to the UPLC-QToF system and the remaining samples after them. The mobile phase was gradient of solvent A: H2O, 0,5 % MeOH, 0,1% FA and solvent B: AcN, 0,1% FA,

gradient mode: 1-90% B in 50 min, 100% A in 50-75 min. Acquity Peptide BEH, C18, 130Å, 1,7 µm, 75 µm* 150 mm. Masswindow (m/z) set to 300 - 2000 and 75 min. The flow rate was 0,3 µL/min. The nanoACQUITY UPLC System coupled to Waters® Micromass® Q-Tof micro™ has been used.

## Series 1 and 2

In Series 1, 10 mL and in Series 2, 60 mL of the extract processed. Only on Series 2, tryptic cleavage was performed, and the intension was to produce more of the sample for being able to perform sequencing.

### Extraction

From each CMPs 10 g weighed up in a mortar and grinded if possible. The grinded mass transferred to a 200 mL flask and weighed up again. In the first step of extraction, 100 mL solvent A (60% ACN, 0,1% Formic acid in H2O) was added and the sample was gently shaken by shaker apparatus in 2 hours. The solvent poured out and collected in another flask. This step repeated 2 more times by adding 100 mL solvent B (30% ACN, 0,1% FA in H2O) and 100 mL solvent C (10% ACN, 0,1% FA in H2O) (see Appendix 3). To separate particles the collected samples were filtered through filter paper.

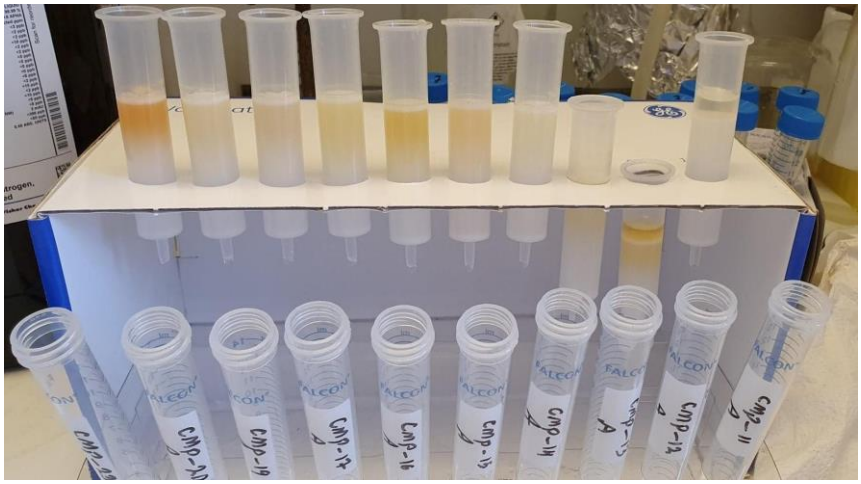
### Solid Phase Extraction (SPE)

To 50 mL falcon tubes, 10 mL of extract transferred. These samples diluted by 23 mL 0,1% FA in H2O to gain 10% AcN in each of them. In Serie 2, 60 mL of the extract transferred and diluted in the same manner (to 10% AcN). Samples were centrifuged in 5 min by 5560 RCF to sediment large particles. SPE column Isolute® C18 (EC), 500 mg was used. Firstly, columns activated by MeOH for at least 1 hour. Columns were calibrated by 2 volumes of 10% ACN, 0,1% FA in H2O. Extracts were loaded and columns were washed by 1 volume 10% ACN, 0,1% FA in H2O. For elution, 5

mL 60% ACN, 0,1% Formic acid in H2O was used and the eluates were collected in 15 mL falcon tubes. After freeze drying, samples were weighed to calculate the yield (See Appendix 4 for series 1, Appendix 6 for series 2).

**Size Exclusion Chromatography (SEC)**

Freeze-dried SPE-fractions dissolved in 3 mL 30% ACN, 0,1% FA in H2O. As reference, 1 mL of each sample saved in a 1,5 mL Eppendorf tube. PD-10 Desalting Columns from GE Healthcare were used. Columns were used according to instructions. In short, the columns were equilibrated using 25 mL 30% ACN, 0,1% FA in H2O. The 2 mL remaining samples were added to columns, let to enter the bed completely and an additional 0,5 mL 30% ACN, 0,1% FA in H2O were added to set the loading volume to 2,5 mL. The flow-through was discarded. Samples eluted by 3,5 mL solvent (30% ACN, 0,1% FA in H2O) and collected in preweighed 15 mL falcon tubes. These samples are High Molecular Weight (HMW) substances and marked as HMW. Columns were eluted again by 4 mL solvent (30% ACN, 0,1% FA in H2O) to catch Low Molecular Weight (LMW) substances which were marked as LMW and collected if needed. After freeze-drying samples were weighed (see Appendix 5 for series 1 and Appendix 7 for series 2).



**Reduction and Alkylation**

Buffer 1: 1M Tris-HCl, pH≈ 8,5, 4mM EDTA (Ethylenediaminetetraacetic)

Buffer 2: 8M Guanidine-HCl, pH≈ 8,5

Buffer 3: 0,5M Citric Acid

Buffer A: 1 mL buffer 1 mixed with 3 mL buffer 2

Buffer B: 3 mg DTT (Dithiothreitol) dissolved in 0,5 mL buffer A

To dried samples have been added 180 µL Buffer A and 20 µL Buffer B. After vortexing, the air inside Eppendorf tubes was removed by flushing nitrogen gas into the tubes. Samples incubated in 2,5 h in dark at 37 °C.

In a falcon tube, 1 g IAM weighed up and 5 mL diluted Buffer 1 (2,5 ml Buffer 1 diluted by 2,5 mL $H_2O$) added. Heating the mixture to 65 °C dissolved IAM in the dark. After cooling down the mixture, 250 µL of mixture is added to each sample. The reaction stopped after 3 min by adding 200 µL Buffer 3. The alkylated samples were compared to the native peptides by ULPC-QToF to identify peptides that harbour disulphides.

**Digestion (Only Serie 2)**

To one aliquot of trypsin (1 µg) 0,5 mL 50 mM NH4HCO3 was added and vortexed. From the trypsin solution, 100 µL pipetted to dried samples and incubated in water bath at 37 °C. After 3 h incubation, 100 µL trypsin solution added to the samples ones again and incubated overnight at 37 °C. Samples were dried in SpeedVac for 3 h. Dry samples were dissolved in 100 µL solvent C (10% ACN, 0,1% FA in H20). After centrifugation, 50 µL transferred to glass vials for analysis by UPLC-QToF.

**Identification (UPLC-QToF)**

A nanoACQUITY UPLC System coupled to Xevo® G2-XS QToF has been used. Run time set to 60 min, flow rate on 0,3 µL and injection volume on 1 µL. Solvent used were A: MQ (Milli-Q

Water), 0,1% FA, B: AcN, 0,1 FA in gradient mode, 0-0,5 min 99% A, 0,5-45 min linear increasing B to 60%, 45-47 min increasing to 85% B, 47-49 min 85% B, 49-50 min re-equilibration at 99% A. The columns used was a peptide BEH C18 (Acquity UPLC® M-Class, Peptide BEH C18 column, 130Å, 1,7 µm, 75 µm* 150 mm).

For identification of peptides MSMS-spectra is used. Immonium ion masses of amino acids are tabulated in Appendix 8.

**Cys-content**

Before and after reduction and alkylation, samples were injected to the UPLC-QToF. The data before and after were compared to each other. Alkylation of Cys by IAM gave a mass change by 58,00548 Da (monoisotopic). One disulphide bond gives a mass change by $2 \times 58{,}00548$ and so on. For the calculation, Excel was used.

# Results

Screening part shows the attempt to sort out and put the focus on species with higher possibility of containing peptides. The criteria for screening part are known peptides in their family/genus, multi-charged signals, and signals above 1000 m/z. In part 2 (Series 1 and 2), results for discovered peptides in plants is presented. Analysis shows that CMP-1, -7, and -8 contain peptides (in agreement with Appendix 8).

The attempt to decide Cys-content did not give any result. Additionally, peptide sequencing approach did not accomplish due to time constraints, but some hints was seen in CMP-1.

## Screening

LTQ-spectra shows signals matching the criteria described in Experimental for most of the samples. In some cases, signals just appear in one extract or signals are much stronger in one extract in contrast to the other extract. Samples CMP-1,4,6,7,8,9,11,12,15,16,17,19, and 20 have been chosen from table 2 for investigation by UPLC-QToF.

Table 2: Analysis of the LTQ spectra.

| CMP | Extract A | | Extract B | | Pep in family (Reference) | Pep in Species (Reference) |
|---|---|---|---|---|---|---|
| | signals | signal above 1000 m/z | signals | signal above 1000 m/z | | |
| **CMP-1** | ✓ | - | - | - | ✓ (13) | Linear cyclotides (13) |
| **CMP-2** | ✓ | - | ✓ | + | | |
| **CMP-3** | - | - | ✓ | + | | |
| **CMP-4** | ✓ | + | ✓ | + | | |
| **CMP-5** | - | - | ✓ | + | | |
| **CMP-6** | ✓ | + | ✓ | ++ | ✓ (14) | |
| **CMP-7** | - | - | ✓ | - | ✓ (15) | |
| **CMP-8** | ✓ | + | ✓ | ++ | ✓(14) | |
| **CMP-9** | ✓ | - | ✓ | + | ✓(14–16) | |
| **CMP-10** | - | - | - | - | ✓(17) | |
| **CMP-11** | ✓ | + | ✓ | + | ✓(18) | |
| **CMP-12** | ✓ | ++ | ✓ | + | ✓(19) | |
| **CMP-13** | ✓ | + | - | - | | |
| **CMP-14** | ✓ | - | ✓ | - | | |
| **CMP-15** | ✓ | ++ | ✓ | ++ | | |
| **CMP-16** | ✓ | ++ | ✓ | ++ | | |
| **CMP-17** | ✓ | + | ✓ | + | ✓(20) | |
| **CMP-19** | ✓ | + | ✓ | + | | |
| **CMP-20** | ✓ | - | ✓ | - | ✓(15) | |
| **CMP-23** | ✓ | - | ✓ | - | | |

++ Many strong signals
+ Some signals
- No signals

The QToF-spectra were analysed, and the result is presented in table 3. As a result, CMP-1, -6, -7, -8 and -12 have been chosen for further analysis.

Table 3: Analysis of the QToF-spectra.

| CMP | A (multi-charged peak) | B (multi-charged peak) | m/z found in both A and B |
|---|---|---|---|
| **CMP-1** | - | * | - |
| **CMP-4** | - | - | - |
| **CMP-6** | ✓ | ✓ | 1235, 1646, 1259 |
| **CMP-7** | ✓ | ✓ | 1686, 1840, 1698, 920 |
| **CMP-8** | ✓ | ✓ | 1241, 827, 1304, 869 |
| **CMP-9** | - | - | - |
| **CMP-11** | - | - | - |
| **CMP-12** | ✓ | ✓ | 1547, 1032, 1482, 1571 |
| **CMP-15** | - | ✓ | 1017, 982, 1235** |
| **CMP-16** | - | - | - |
| **CMP-17** | - | - | - |
| **CMP-19** | - | - | - |
| **CMP-20** | - | - | - |

* No sample available

** Signals appeared just in extract B.

## Series 1 and 2

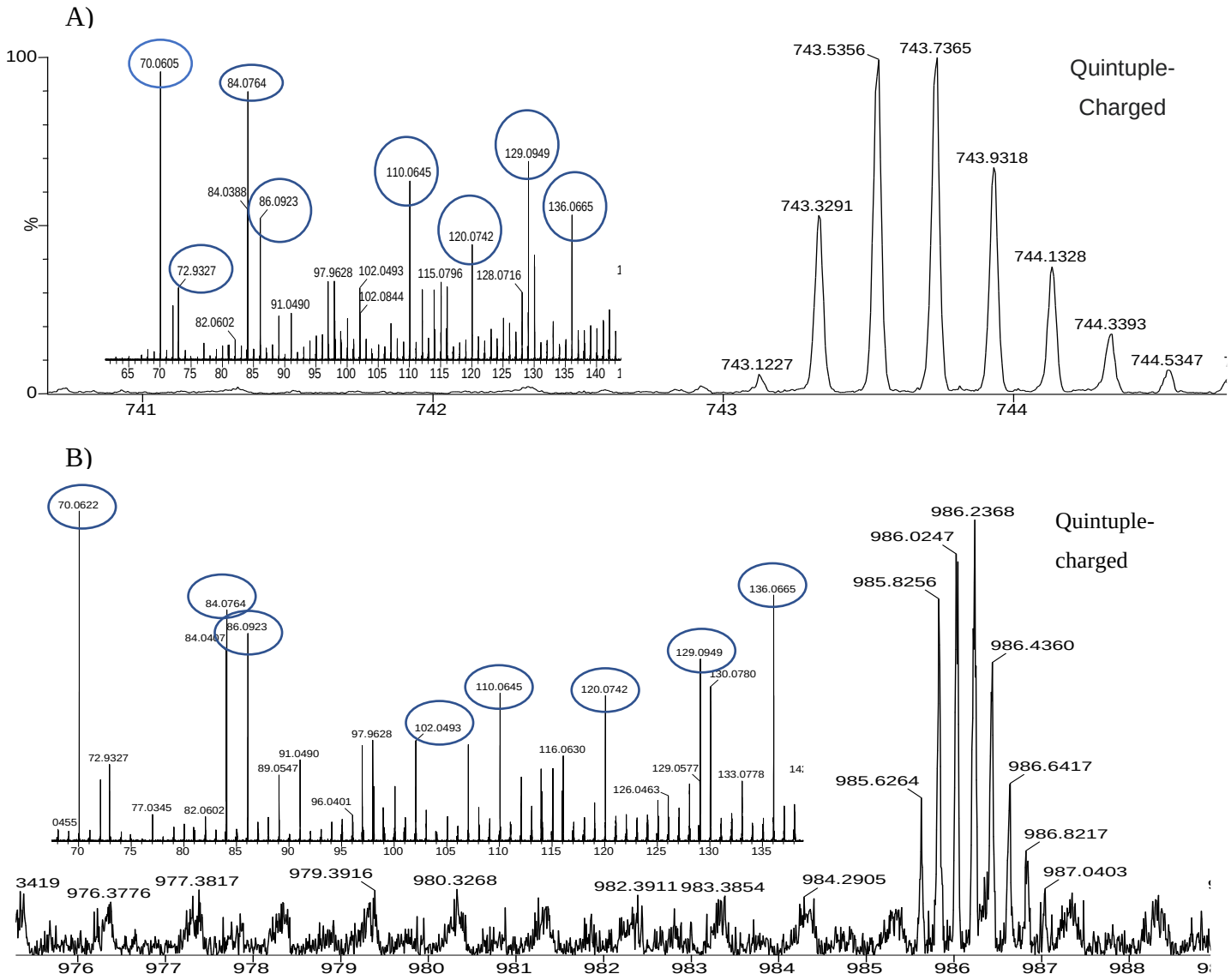



Figure 1: presents the spectra of CMP-1. Circled signals show immonium ions of amino acids. A and B show spectra for two different peptides.

Spectrum A belongs to a peptide with a mass of 3710 Da and spectrum B belongs to a peptide weighing 3936 Da, according to Figure 1. CMP-1 showed signals for peptides in the range of 2,5-4 kDa. These peptides appeared mostly in the beginning of the chromatogram between 19-25 m.

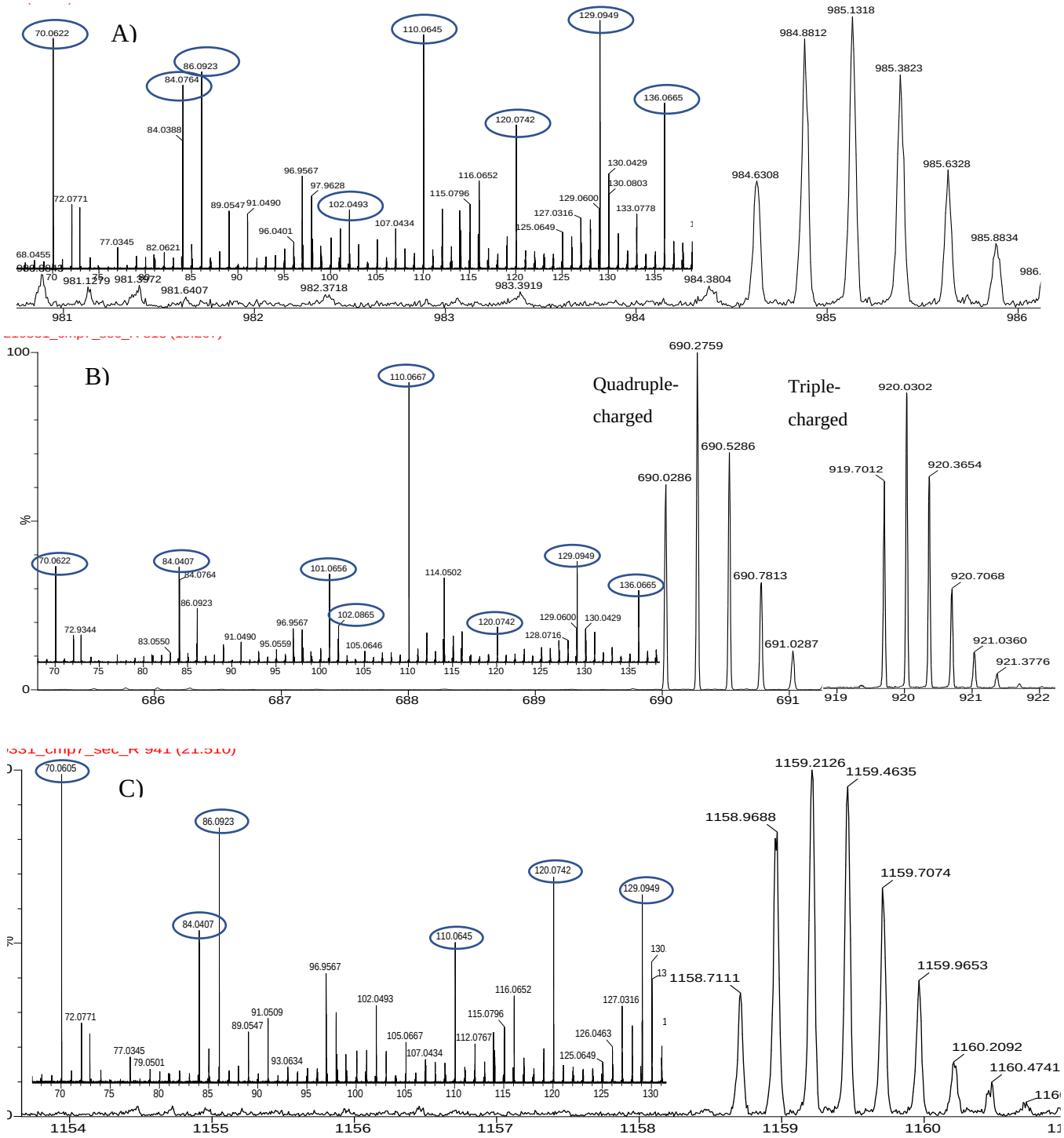



Figure 2: Presents some of CMP-7 data. Spectra show multi-charged signals and its fragmentation spectra with immonium ions of amino acids marked. A, B and C are three different spectra, each belongs to one peptide.

CMP-7 shows peptides appearing mainly in 19-25 min in the chromatogram. Spectrum A belongs to 3932 Da peptide, B shows a 2754 Da peptide and C has 4628 Da mass, according to Figure 2.

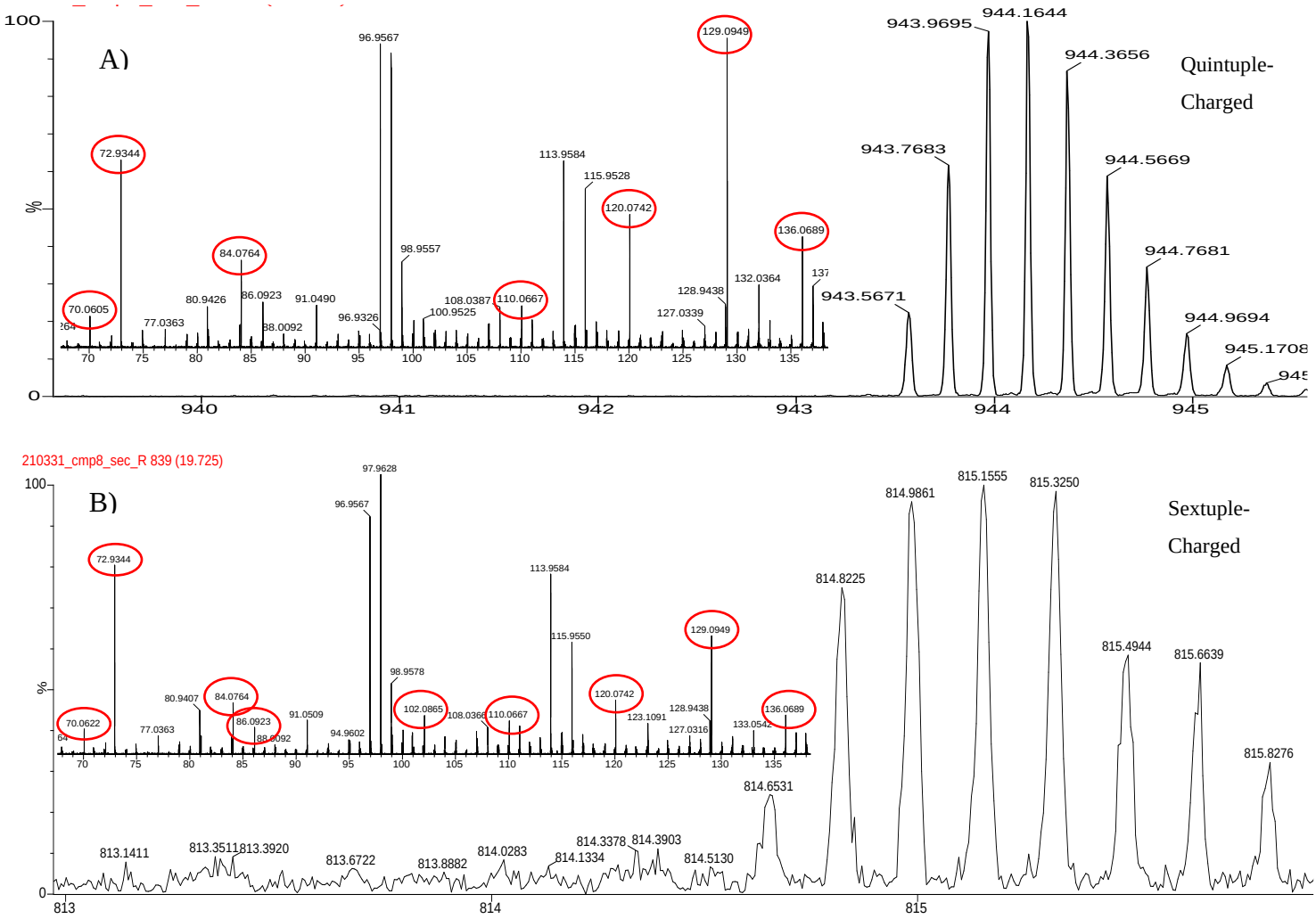



Figure 3: Presents some of CMP-8 data. Spectra show multi-charged signals and its fragmentation spectra with immonium ions of amino acids marked. A and B are two different spectra, each belongs to one peptide.

As shown in Figure 3, CMP-8 showed peaks for even bigger peptides that mostly appeared in 3,5-5 kDa. Spectrum A shows a 4710 Da peptide and B a 4878 Da peptide,

## Discussion

Analysis of these 20 herbal medicines resulted in finding peptides in CMP-1, -7 and -8. Peptides appeared mostly in the beginning of the chromatogram and were mainly big peptides 2,5-5 kDa. According to table 2 in the screening part of the result, all these 3 species belong to the plant family with already known peptides. Immonium ions of amino acids in fragmentation spectra are the same as what table 4 suggests would show strong signals. This evidence proves the existence of peptides in these plants.

CMP-1 was already studied by Nguyen GKT et al (13) and they discovered linear cyclotides in the mass range 2,5-4 kDa. Peptides in the same mass range have been found in this project as well. In screening part, it was hard to found multi-charged signals in CMP-1. Firstly, there was not extract B available. Secondly, it can depend on the solubility of these peptides in MeOH: AcN as extract A was based on. In addition, different part of the plant has been used in the project. Series 1 and 2 were based on same solvent as extract B (AcN, FA and $H_2O$), where a series of extractions were performed to get a broad range extract. The result of Series 1 and 2 indicates that peptides found in CMP-1 have good solubility in this solvent.

CMP-7 and -8 belong to the plant families Asteraceae respective Fabaceae. Studies show the existence of peptides in these families although, there is no previous study on peptide content of these 2 species. According to Franke et al (15), Paws-derived peptides (PDPs) are discovered in Zinnia haageana and also mentioned that PDPs are abundantly found in the Millereae and Heliantheae tribes in Asteraceae family. There is a possibility that peptides found in CMP-7 are PDPs. Snakin-like peptide was discovered in *Peltophorum dubium* from the family Fabaceae, by Rodríguez-Decuadro et al (14). In this family, AMPs e.g., Snakins are widespread, and peptides found in CMP-8 can belong to this class of peptides.

The total yield in both series (1 and 2) shows almost under 1% for most of the plants, according to Appendix 5 and 7. Total Yields are estimated because just 10 mL in Series 1 and 60 mL in Series 2 from the total 300 mL gained from almost 10 g material were used which decrease the liability of these values. Also, this extraction process needs to be done at least in triplet for more accurate total yield. Some negative values and values above 100% are also highlighted in Appendix 5. These values can not exist and can depend on low amount (10 mL of the total 300 mL) extract used in Series 1 which the analytical scale used was not suitable for.

The result of Cys-content attempt was questionable. It was a time-consuming process to determine the Cys-content and no match was found. The question raising is, are these cyclotides and the method for determination of Disulfide bonds lack effectiveness? To answer to these questions further examination is required. Nguyen GKT et al (13) used another approach that previously was used by Göransson and Craik (21) where partial reduction and stepwise alkylation was performed. Can these peptides be the reason for the anti-inflammatory effect and prescription for treatment of RA? As mentioned in the background, traditionally medicinal plants included in this work are decocted with water for orally taking. Peptides are generally sensitive to high temperature and peptides do not have a good stability in gastrointestinal tract. This knowledge implies that found peptides probably are not available with same structure for absorption by taking orally and the bioactivity cannot be predicted.

This finding shows that there is a lot that we do not know about these medicinal herbs. Understanding the benefits and functions of found peptides in the plants can be useful in treatment of diseases. The immediate next step should be to sequence these peptides either by isolating or using peptidomics, follow by determination of the possible bioactivity of discovered peptide by performing different bioassays. I would suggest firstly an antimicrobial and then an anti-inflammatory bioassay.

## Acknowledgments

I wish to express my gratitude to Erik Jakobsson my supervisor, Ulf Göransson, Luke Robertson and all members of pharmacognosy department for the support and mentorship.

## References

1. Xi S, Gong Y. Essentials of Chinese Materia Medica and Medical Formulas: New Century Traditional Chinese Medicine [Internet]. San Diego, UNITED STATES: Elsevier Science & Technology; 2017 [cited 2021 Feb 28]. Available from: http://ebookcentral.proquest.com/lib/uu/detail.action?docID=4853140

2. De Coninck B, De Smet I. Plant peptides – taking them to the next level. J Exp Bot. 2016 Aug 1;67(16):4791–5.

3. Hayes M, Bleakley S. 21 - Peptides from plants and their applications. In: Koutsopoulos S, editor. Peptide Applications in Biomedicine, Biotechnology and Bioengineering [Internet]. Woodhead Publishing; 2018 [cited 2021 Mar 6]. p. 603–22. Available from: https://www.sciencedirect.com/science/article/pii/B9780081007365000259

4. Hemu X, Serra A, Darwis DA, Cornvik T, Sze SK, Tam JP. Peptidomic Identification of Cysteine-Rich Peptides from Plants. In: Schrader M, Fricker L, editors. Peptidomics: Methods and Strategies [Internet]. New York, NY: Springer; 2018 [cited 2021 Mar 6]. p. 379–93. (Methods in Molecular Biology). Available from: https://doi.org/10.1007/978-1-4939-7537-2_26

5. Marshall E, Costa LM, Gutierrez-Marcos J. Cysteine-Rich Peptides (CRPs) mediate diverse aspects of cell–cell communication in plant reproduction and development. J Exp Bot. 2011 Mar 1;62(5):1677–86.

6. Jensen KJ, Tofteng Shelton P, Pedersen SL, editors. Peptide Synthesis and Applications [Internet]. Totowa, NJ: Humana Press; 2013 [cited 2021 Mar 6]. (Methods in Molecular Biology; vol. 1047). Available from: http://link.springer.com/10.1007/978-1-62703-544-6

7. Suttapitugsakul S, Xiao H, Smeekens J, Wu R. Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics. Mol Biosyst. 2017 Nov 21;13(12):2574–82.

8. Hustoft H, Malerod H, Wilson S, Reubsaet L, Lundanes E, Greibrokk T. A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. In 2012.

9. Papaleo E, Fantucci P, De Gioia L. Effects of Calcium Binding on Structure and Autolysis Regulation in Trypsins. A Molecular Dynamics Investigation. J Chem Theory Comput. 2005 Nov;1(6):1286–97.

10. Ebringer A. Rheumatoid Arthritis and Proteus [Internet]. London: Springer London; 2012 [cited 2021 Mar 5]. Available from: http://link.springer.com/10.1007/978-0-85729-950-5

11. WHO | Chronic rheumatic conditions [Internet]. WHO. World Health Organization; [cited 2021 Mar 5]. Available from: http://www.who.int/chp/topics/rheumatic/en/

12. Sturfelt ON Gunnar. Reumatiska sjukdomar | Läkemedelsboken [Internet]. [cited 2021 Mar 5]. Available from: https://lakemedelsboken.se/kapitel/rorelseapparaten/reumatiska_sjukdomar.html?search=reumatism&id=p2_14#p2_14

13. Nguyen GKT, Lian Y, Pang EWH, Nguyen PQT, Tran TD, Tam JP. Discovery of Linear Cyclotides in Monocot Plant Panicum laxum of Poaceae Family Provides New Insights into Evolution and Distribution of Cyclotides in Plants *. J Biol Chem. 2013 Feb 1;288(5):3370–80.

14. Rodríguez-Decuadro S, Barraco-Vega M, Dans PD, Pandolfi V, Benko-Iseppon AM, Cecchetto G. Antimicrobial and structural insights of a new snakin-like peptide isolated from *Peltophorum dubium* (Fabaceae). Amino Acids. 2018 Sep;50(9):1245–59.

15. Franke B, Jayasena AS, Fisher MF, Swedberg JE, Taylor NL, Mylne JS, et al. Diverse cyclic seed peptides in the Mexican zinnia (*Zinnia haageana*). Biopolymers. 2016 Nov;106(6):806–17.

16. Ma M-S, Bae IY, Lee HG, Yang C-B. Purification and identification of angiotensin I-converting enzyme inhibitory peptide from buckwheat (*Fagopyrum esculentum* Moench). Food Chem. 2006 May 1;96(1):36–42.

17. Figueiredo PO, Matos M de FC, Perdomo RT, Kato WH, Barros MVGO, Garcez FR, et al. Rubiaceae-Type Cyclopeptides from *Galianthe thalictroides*. J Nat Prod. 2016 Apr 22;79(4):1165–9.

18. Khani S, Seyedjavadi SS, Zare-Zardini H, Hosseini HM, Goudarzi M, Khatami S, et al. Isolation and functional characterization of an antifungal hydrophilic peptide, Skh-AMP1, derived from *Satureja khuzistanica* leaves. Phytochemistry. 2019 Aug 1;164:136–43.

19. Zhang L, Ng TB, Lam JKW, Wang SW, Lao L, Zhang KY, et al. Research and Development of Proteins and Peptides with Therapeutic Potential from Yam Tubers. Curr Protein Pept Sci. 2019;20(3):277–84.

20. Bokesch HR, Charan RD, Meragelman KM, Beutler JA, Gardella R, O'Keefe BR, et al. Isolation and characterization of anti-HIV peptides from *Dorstenia contrajerva* and *Treculia obovoidea*. FEBS Lett. 2004 Jun 4;567(2–3):287–90.

21. Göransson U, Craik DJ. Disulfide Mapping of the Cyclotide Kalata B1: CHEMICAL PROOF OF THE CYCLIC CYSTINE KNOT MOTIF*. J Biol Chem. 2003 Nov 28;278(48):48188–96.

22. Zhang J, Wang Z, Liu Y. A Brief History of Dietary Chinese Herbs. In: Liu Y, Wang Z, Zhang J, editors. Dietary Chinese Herbs: Chemistry, Pharmacology and Clinical Evidence [Internet]. Vienna: Springer; 2015 [cited 2021 Jan 22]. Available from: https://doi.org/10.1007/978-3-211-99448-1_1

23. Kuang H, Xia Y, Yang B, Wang Q, Wang Y. Screening and comparison of the immunosuppressive activities of polysaccharides from the stems of *Ephedra sinica* Stapf. Carbohydr Polym. 2011 Jan 10;83(2):787–95.

24. World Health Organization, editor. WHO monographs on selected medicinal plants. Vol. 1. Geneva: World Health Organization; 1999.

25. Jun X, Fu P, Lei Y, Cheng P. Pharmacological effects of medicinal components of *Atractylodes lancea* (Thunb.) DC. Chin Med. 2018 Nov 27;13(1):59.

26. Carbonara T, Pascale R, Argentieri MP, Papadia P, Fanizzi FP, Villanova L, et al. Phytochemical analysis of a herbal tea from *Artemisia annua* L. J Pharm Biomed Anal. 2012 Mar 25;62:79–86.

27. World Health Organization, editor. WHO monograph on good agricultural and collection practices (GACP) for *Artemisia annua* L. Geneva: WHO; 2006. 49 p.

28. Ayaz M, Junaid M, Ahmed J, Ullah F, Sadiq A, Ahmad S, et al. Phenolic contents, antioxidant and anticholinesterase potentials of crude extract, subsequent fractions and crude saponins from *Polygonum hydropiper* L. BMC Complement Altern Med. 2014 May 3;14(1):145.

29. Peng ZF, Strack D, Baumert A, Subramaniam R, Goh NK, Chia TF, et al. Antioxidant flavonoids from leaves of *Polygonum hydropiper* L. Phytochemistry. 2003 Jan 1;62(2):219–28.

30. Jiang H, Wang X, Yang L, Zhang J-X, Hou A-J, Man W-J, et al. The fruits of *Xanthium sibiricum* Patr: A review on phytochemistry, pharmacological activities, and toxicity. World J Tradit Chin Med. 2020 Jan 10;6(4):408.

31. World Health Organization, editor. WHO monographs on selected medicinal plants. Vol. 3. Geneva: World Health Organization; 1999.

32. Ong SP, Cai B, Zhang P, Liu X, Ng WW, Li HH, et al. High Performance Liquid Chromatography Fingerprinting Technology Of The Commonly-used Traditional Chinese Medicine Herbs [Internet]. Singapore, SINGAPORE: World Scientific Publishing Company; 2010 [cited 2021 Feb 20]. Available from: http://ebookcentral.proquest.com/lib/uu/detail.action?docID=919079

33. Yan X, editor. Dan Shen (*Salvia miltiorrhiza*) in Medicine, Biology and chemistry. Dordrecht Heidelberg New York London: Springer; 2015. 19–25 p. (Dan Shen (Salvia miltiorrhiza) in Medicine; vol. 1).

34. Ou-yang S, Jiang T, Zhu L, Yi T. *Dioscorea nipponica* Makino: a systematic review on its ethnobotany, phytochemical and pharmacological profiles. Chem Cent J [Internet]. 2018 May 11 [cited 2021 Jan 22];12. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5945570/

35. Identification of Autotoxic Compounds in Fibrous Roots of Rehmannia (*Rehmannia glutinosa* Libosch.) [Internet]. [cited 2021 Jan 22]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0028806

36. Jiang Z-B, Liu H-L, Liu X-Q, Shang J-N, Zhao J-R, Yuan C-S. Chemical constituents of *Gentiana macrophylla* Pall. Nat Prod Res. 2010 Sep 10;24(14):1365–9.

37. Jia N, Li Y, Wu Y, Xi M, Hur G, Zhang X, et al. Comparison of the anti-inflammatory and analgesic effects of *Gentiana macrophylla* Pall. and *Gentiana straminea* Maxim., and identification of their active constituents. J Ethnopharmacol. 2012 Dec 18;144(3):638–45.

38. Cao L, Li B, Shehla N, Gong L, Jian Y-Q, Peng C-Y, et al. Triterpenoids from stems of *Kadsura heteroclita*. Fitoterapia. 2020 Jan 1;140:104441.

39. Yu H, Zeng R, Lin Y, Li X, Tasneem S, Yang Z, et al. *Kadsura heteroclita* stem suppresses the onset and progression of adjuvant-induced arthritis in rats. Phytomedicine. 2019 May 1;58:152876.

40. Gundogdu M, Muradoglu F, Sensoy RIG, Yilmaz H. Determination of fruit chemical properties of *Morus nigra* L., *Morus alba* L. and *Morus rubra* L. by HPLC. Sci Hortic. 2011 Dec 5;132:37–41.

41. Yuan Q, Zhao L. The Mulberry (Morus alba L.) Fruit—A Review of Characteristic Components and Health Benefits. J Agric Food Chem. 2017 Dec 6;65(48):10383–94.

42. Chen Z, Zhang C, Gao F, Fu Q, Fu C, He Y, et al. A systematic review on the rhizome of *Ligusticum chuanxiong* Hort. (Chuanxiong). Food Chem Toxicol. 2018 Sep 1;119:309–25.

43. Seo C-S, Lim H-S, Jeong S-J, Shin H-K. Anti-allergic effects of sesquiterpene lactones from the root of *Aucklandia lappa* Decne. Mol Med Rep. 2015 Nov 1;12(5):7789–95.

44. Yang X, Zhang X, Yang S-P, Le T, Chen B. Evaluation of *Aucklandia lappa* Decne extracts as antiulcer activity in animals. Pak J Pharm Sci. 2016 Sep 1;29:1695–701.

45. He Y-N, Ou S-P, Xiong X, Pan Y, Pei J, Xu R-C, et al. Stems and leaves of *Aconitum carmichaelii* Debx. as potential herbal resources for treating rheumatoid arthritis: Chemical analysis, toxicity and activity evaluation. Chin J Nat Med. 2018 Sep 1;16(9):644–52.

46. Azietaku JT, Ma H, Yu X, Li J, Oppong MB, Cao J, et al. A review of the ethnopharmacology, phytochemistry and pharmacology of *Notopterygium incisum*. J Ethnopharmacol. 2017 Apr 18;202:241–55.

# Supplementary Information

## Appendix 1: Plant species

CMP-1, *Coix Lacryma-jobi* or job´s tears can grow up to 200 cm. The flowers have white, yellow, or red colour. The white oval seed has a sweet taste and uses traditionally as medicine. To date 41 different chemical components of the seed has been isolated and identified which includes coixenolides, triglycerides, fatty acids, lactams, coixol, saccharides, sterols, and triterpenes. Additionally, pharmacological studies proved the antitumor, anti-inflammatory, anti-proliferative and antiallergic effect of Job´s seed(22).

CMP-2, *Angelica Sinensis* or Danggui in Chinese is a 15–25 cm long plant with light brown to brown colour. Over 70 small molecules have been identified and studied which can be divided into volatile and water-soluble substances. Above all, there are organic acids, vitamins, amino acids and certain trace elements(22).

CMP-3, *Paeonia lactiflora Pall.* even called Shaoyao in Chinese is a well-known herb which originally used for treatment of typhoid. Furthermore, Shaoyao prescribed for rheumatoid arthritis, spasms, dysmenorrhea, and muscle cramp. Chemical compounds isolated from Shaoyao are mainly monoterpene glycosides, flavonoids, polyphenolic compounds, and daucosterols(22).

CMP-4, *Ephedra Sinica* is well-known for its active component, phenylpropylamino alkaloids. Ephedrine alkaloids are adrenergic agonist. Chinese Ephedra or Caomahuang is included in various herbal preparations and mostly uses as anti-inflammatory, antipyretic, respiratory, and central nervous system stimulant medicine. Besides, studies shows some polysaccharides produced in Caomahuang have anti suppressive effects and can be a potential treatment for autoimmune diseases(23,24).

CMP-5, *Cinnamomum Cassia* is an evergreen tree with a high up to 17 m. Rougui harvests at autumn and the dark brown bark dries in shade. The bark of Rougui is rich in oils and various types

of polysaccharides, sesquiterpenes, glycosides, flavanols, and flavanol polymers. Rougui has anti-inflammatory, antidiabetic, anti-ulcer, and anti-microbial properties(22,24).

CMP-6, *Glycyrrhiza Uralensis*, licorice (Gan cao) is a perennial herb with umami taste and used traditionally in most chinese herbal medicines to reduce the bitter taste. Licorice has purple and white coloured flower. Root and underground stem of the plant utilizes the traditional medicine. Licorice consist of different class of triterpenoid saponins, phenolic compounds such as coumarins, flavanone, chalcones, and alkaloids. Pharmacological studies states anti-inflammatory, antioxidant, hepatoprotection, and antitumor activity of licorice(22,24).

CMP-7, *Atractylodes lancea* (AL) or Cangzhu is a perennial herb that can grow up to 1 m tall. The medicinal part is the rhizome. AL prescribed traditionally for treatment of rheumatic diseases, digestive disorders, and influenza. The chemical components isolated from Cangzhu are mostly volatile oil, sesquiterpene glycosides, tryptpphane, atraclyloin, furaldehyde, atractylodin. Pharmacological studies confirm anti-inflammatory, inticancer, and anti.obesity activity of AL(1,25).

CMP-8, *Astragalus membranaceus*, also called huangqi in Chinese is a perennial plant with up to 80 cm tall and hairy stem and leaves. The white or yellow root of the 4-year-old plant is used for medicinal purposes in Chinese traditional medicine. Huangqi contains flavonoids, saponins and polysaccharides(22,24).

CMP-9, *Artemisia Annua L*, Huanghuahao is an annual plant. Artemisia can grow to 150 cm tall. The entire plant can be used for medical purposes. Phytochemical studies demonstrate sesquiterpenoids, artemisinin, flavonoids, coumarins and volatile oils. The herb is known for its unique component artemisinin which is the most potent antimalaria drug(26,27).

CMP-9, *Polygonum hydropiper L*, Shuiliao is an annual herb. Shuiliao is 20-60 cm tall and usually aerial part, root, and fruit of the herb uses in traditional Chinese medicines. The chemical components are flavones, flavonoid glycosides, sesquiterpene acid, and polygonic acid(28,29).

CMP-9, *Xanthium sibiricum*, Cang'er is a 20-90 cm tall annual plant. All part of the plant can be consumed as traditional medicine. Commonly, the herb harvests, cuts in pieces, sun-dries or uses fresh. The herb was traditionally used for common cold, itching skin, haemorrhoid, spasms, etc. Pharmacologically, the herb has anti-inflammatory, hypoglycaemic, diuretic, respiratory, cardiovascular activity. The main chemical constituents are, phenylpropanoids, sesquiterpene, glycosides, thiazids, etc(1,30).

CMP-9, *Vigna Umbellata* (Chixiaodou, Rice Bean) is a nutritious, protein-rich small grain that grows 1 m tall. The red seed of rice bean is recorded as medicine in Shennong Bencao Jing, first Chinese pharmacy monograph written for 2000 years ago. Three major classes of compound were found in rice bean, triterpenoid saponins, polyphenols and flavonoids. In traditional Chinese medication red bean was prescribed as diuretic, against edema, rheumatic arthritis, etc. Pharmacological studies proved anti-inflammatory, antioxidant, hepatoprotective and hypotonic activity of the herb(22).

CMP-9, *Prunus armeniaca L* (Kuxingren, bitter Apricot) is a fruit tree. Normally apricot grows on hillslopes and up to 9 m tall. Kuxingren is the seed of the plant and traditionally is consumed mainly for respiratory problem e.g., cough and asthma. Kuxingren contains cyanohydrin glycoside amygdalin, fatty acids, and volatile components such as linalool, γ-decanolactone, etc. Studies are currently focused on amygdalin the cyanohydrin component of the herb due to its reducing effect on development of atherosclerosis disease by attenuating the inflammatory response(22,31)

CMP-10, *Morinda Officinalis* (Bajitian) is one of the most famous Chinese traditional medicines. The bark of the root is light purple in color and is used as medicine. The plant consists mainly of anthraquinones, iridoids and oligosaccharides. In addition, steroids, amino acids, and certain volatile substances have been isolated. Studies suggest that the herb exhibits anti-inflammatory, antinociceptive, anti-osteoporotic, and antihepatotoxic activity(22).

CMP-11, *Salvia miltiorrhiza* (Danschen) grows to a maximum of 60 cm. The flowers can be blue or purple. The chemical components isolated and identified are phenolic acids and

phenanthraquinones. The root of Danschen is traditionally used for treatment of irregular menstrual period, abdominal pain, etc. The herb exhibits anti-inflammatory, intioxidant, and anti-coagulative activity(32,33).

CMP-12, *D. nipponica Makino* (Chuanlongshuyu) has been traditionally used to treat rheumatoid arthritis, pain, cough, and asthma. Pharmacological studies have shown the herb's anti-inflammatory, antitumor, antidiuretic, and analgesic effects. The herb mainly contains steroids and saponins(34).

CMP-13, *Rehmannia glutinosa Libosch* (Dihuang) is a perennial that has been used in traditional Chinese medicine. Traditionally, the root, either fresh or dried, has been used for treatment of diabetes, insomnia, rheumatism arthritis, and haematological conditions. Bioactive substances identified in the herb include carbohydrates, iridoid glycosides, rehmannin, aucubin, catalpol, and campesterol(31,32,35).

CMP-14, *Gentiana macrophylla Pall.* (Qinjiao) shows significant anti-inflammatory and antinociceptive activity. These pharmacological effects are thought to be linked to substances such as flavonoids, secoiroid glycosides and triterpenes detected in the extract. Qinjiao used traditionally for treatment of pain, constipation and rheumatism arthritis(36,37).

CMP-15, *Kadsura heteroclita (Roxb.) Craib* (Xuetong) has long been used as a prophylaxis and treatment for rheumatoid arthritis. Studies show that the extract of the herb consists of lignans and triterpenes(38). Studies confirm the anti-inflammatory activity of the KHC(39).

CMP-16, *Morus Alba L*. (Sangzhi) is registered in the Chinese Pharmacopoeia. All parts of the plant are used for medicinal purposes. Above all, the extract consists of flavonoids, alkaloids, polysaccharides, anthocyanins, and various Diel-Alder adducts(40,41).

CMP-17, *Ligosticum chuanxiong* Hort. has long been used for headaches, rheumatism, menstrual disorders and swelling in traditional Chinese medical treatment. Studies indicate that the biological components discovered include phthalides, alkaloids and volatile oils. Pharmacologically, the herb has anti-inflammatory, intioxidant, anti-atherosclerosis, anti-cerebral ischemia, etc(32,42).

CMP-19, *Aucklandia Lappa* Dence is a well-known herbal medicine used for treatment of rheumatism arthritis, coughs, asthma, tuberculosis, etc. Bioactive constituents of A.Lappa are mostly sesquiterpene lactones, dehydrocostus lactone, alantolactone and costunolide. Several pharmacological activity of A.Lappa has been reported including anti-inflammatory, anti-ulcer(1,43,44).

CMP-20, *Aconitum carmichaelii Debx* (Chuanwu) is a perennial herb recorded in Chinese Materia Medica. The herb has shown significant effect against rheumatoid arthritis mainly by reducing pain and inflammation. Besides, studies showed pharmacological activity in treating diarrhoea, edema, gastroenteritis, bronchial asthma, etc. Chuanwu consists mostly of alkaloids as bioactive compounds. This sample were preboiled before extraction(1,45).

CMP-23, *Notopterygium incisum* (qiang hua) is an herbal medicine collected and sun-dried root in autumn. The herb has been used for treatment of common cold, rheumatoid arthritis, headache, etc. The main bioactive component of the root are coumarins, phenoloids, sesquiterpenes, and alkaloid(1,46).

## Appendix 2: Amount extracts, screening

Screening: amounts of extracts A and B used for screening by HPLC-MS

| Extracts A | (mg) | Extracts B | (mg) |
|---|---|---|---|
| **CMP-1** | 115,7 | **CMP-1** | - |
| **CMP-2** | 2,0 | **CMP-2** | 15,4 |
| **CMP-3** | 5,6 | **CMP-3** | 4,7 |
| **CMP-4** | 10,8 | **CMP-4** | 5,6 |
| **CMP-5** | 2,8 | **CMP-5** | 29,4 |
| **CMP-6** | 9,6 | **CMP-6** | 3,1 |
| **CMP-7** | 3,2 | **CMP-7** | 34,3 |
| **CMP-8** | 1,8 | **CMP-8** | 19,1 |
| **CMP-9** | 5,4 | **CMP-9** | 6,4 |
| **CMP-10** | 20,1 | **CMP-10** | 8,0 |
| **CMP-11** | 1,5 | **CMP-11** | 7,6 |
| **CMP-12** | 1,2 | **CMP-12** | 5,5 |
| **CMP-13** | 3,2 | **CMP-13** | 6,8 |
| **CMP-14** | 14,9 | **CMP-14** | 22,4 |
| **CMP-15** | 7,5 | **CMP-15** | 4,7 |
| **CMP-16** | 1,9 | **CMP-16** | 4,1 |
| **CMP-17** | 6,2 | **CMP-17** | 2,1 |
| **CMP-19** | 10,4 | **CMP-19** | 20,6 |
| **CMP-20** | 11,1 | **CMP-20** | 9,2 |
| **CMP-23** | 5,0 | **CMP-23** | 22,5 |

## Appendix 3: plants mass

Extraction: weighted plants mass and total extract volume.

| Plants | Mass (g) | Total Extract (mL) |
|---|---|---|
| **CMP-1** | 10,04 | 300 |
| **CMP-2** | 10,08 | 300 |
| **CMP-3** | 10,04 | 300 |
| **CMP-4** | 10,01 | 300 |
| **CMP-5** | 10,09 | 300 |
| **CMP-6** | 10,01 | 300 |
| **CMP-7** | 10,02 | 300 |
| **CMP-8** | 10,28 | 300 |
| **CMP-9** | 10,03 | 300 |
| **CMP-10** | 10,2 | 300 |
| **CMP-11** | 10,22 | 300 |
| **CMP-12** | 10,12 | 300 |
| **CMP-13** | 9,99 | 300 |
| **CMP-14** | 10,07 | 300 |
| **CMP-15** | 10,03 | 300 |
| **CMP-16** | 10,05 | 300 |
| **CMP-17** | 10,03 | 300 |
| **CMP-19** | 9,99 | 300 |
| **CMP-20** | 10,15 | 300 |
| **CMP-23** | 9,97 | 300 |

## Appendix 4: yield for Series 1 after SPE

Series 1: The extract yield obtained after SPE in series 1.

| Plants | Extract (mL) | Gained mass (mg) | Yield (%) |
|---|---|---|---|
| **CMP-1** | 10 | 0,2 | 0,060 |
| **CMP-2** | 10 | 0,5 | 0,149 |
| **CMP-3** | 10 | 3,2 | 0,956 |
| **CMP-4** | 10 | 3,9 | 1,169 |
| **CMP-5** | 10 | 6,3 | 1,873 |
| **CMP-6** | 10 | 22,6 | 6,773 |
| **CMP-7** | 10 | 1,6 | 0,479 |
| **CMP-8** | 10 | 2,0 | 0,584 |
| **CMP-9** | 10 | 2,8 | 0,837 |
| **CMP-10** | 10 | 0,3 | 0,088 |
| **CMP-11** | 10 | 4,7 | 1,380 |
| **CMP-12** | 10 | 13,5 | 4,002 |
| **CMP-13** | 10 | 0,2 | 0,060 |
| **CMP-14** | 10 | 8,2 | 2,443 |
| **CMP-15** | 10 | 8,4 | 2,512 |
| **CMP-16** | 10 | 0,9 | 0,269 |
| **CMP-17** | 10 | 2,2 | 0,658 |
| **CMP-19** | 10 | 5,3 | 1,592 |
| **CMP-20** | 10 | 0,7 | 0,207 |
| **CMP-23** | 10 | 0,2 | 0,060 |

## Appendix 5: Yield after SEC, Series 1.

Series 1: Yield after SEC and total yield.

| species | Actual Yield SPE (mg) | Actual Yield SEC (mg) | Yield SEC (%) | Total Yield (%) |
|---|---|---|---|---|
| **CMP-1** | 0,2 | 0,1 | 50 | 0,03 |
| **CMP-2** | 0,5 | 0,3 | 60 | 0,09 |
| **CMP-3** | 3,2 | 0 | 0,0 | 0,00 |
| **CMP-4** | 3,9 | 0,3 | 7,7 | 0,09 |
| **CMP-5** | 6,3 | 4,8 | 76,2 | 1,43 |
| **CMP-6** | 22,6 | 4,2 | 18,6 | 1,26 |
| **CMP-7** | 1,6 | 0,1 | 6,2 | 0,03 |
| **CMP-8** | 2,0 | 0 | 0,0 | 0,00 |
| **CMP-9** | 2,8 | 3,3 | 117,9 | 0,99 |
| **CMP-10** | 0,3 | -0,2 | -66,7 | -0,06 |
| **CMP-11** | 4,7 | 0,1 | 2,1 | 0,03 |
| **CMP-12** | 13,5 | 3,9 | 28,9 | 1,16 |
| **CMP-13** | 0,2 | 3,1 | 1550 | 0,93 |
| **CMP-14** | 8,2 | 0,2 | 2,4 | 0,06 |
| **CMP-15** | 8,4 | 0,6 | 7,1 | 0,18 |
| **CMP-16** | 0,9 | 1,8 | 200 | 0,54 |
| **CMP-17** | 2,2 | 0,2 | 9,1 | 0,06 |
| **CMP-19** | 5,3 | -0,1 | -1,9 | -0,03 |
| **CMP-20** | 0,7 | 0,1 | 14,3 | 0,03 |
| **CMP-23** | 0,2 | -0,2 | -100 | -0,06 |

## Appendix 6: Yield after SPE, Series 2.

Series 2: amount extract used and yield after SPE.

| Species | Extract (mL) | Actual Yield | Yield % |
|---|---|---|---|
| **CMP-1** | 60 | 7,3 | 0,4 |
| **CMP-6** | 60 | 134,5 | 6,7 |
| **CMP-7** | 60 | 23,4 | 1,2 |
| **CMP-8** | 60 | 11,6 | 0,6 |
| **CMP-12** | 60 | 77,9 | 3,8 |

## Appendix 7: Yield after SEC, Series 2.

Series 2: Yield after SEC and the total yield

| species | Actual Yield SPE (mg) | Actual Yield (mg) | Yield SEC (%) | Total Yield (%) |
|---|---|---|---|---|
| **CMP-1** | 7,3 | 1,9 | 26,0 | 0,09 |
| **CMP-6** | 134,5 | 120,5 | 89,6 | 6,02 |
| **CMP-7** | 23,4 | 5,4 | 23,1 | 0,27 |
| **CMP-8** | 11,6 | 4 | 34,5 | 0,19 |
| **CMP-12** | 77,9 | 33,3 | 42,7 | 1,65 |

## Appendix 8: Immonium ions

Immonium ion: masses needed for elucidation of the spectra.

| Residue | Immonium Ion | Related ion |
|---|---|---|
| **Alanine** | 44 | |
| **Arginine** | *129* | 59,70,*73*,87,100,112 |
| **Asparagine** | *87* | *70* |
| **Aspartic acid** | *88* | *70* |
| **Cysteine** | *76* | |
| **Glutamic acid** | *102* | |
| **Glutamine** | 101 | 56,84,*129* |
| **Glycine** | 30 | |
| **Histidine** | **110** | *82*,*121*,*123*,*138*,166 |
| **Isoleucine** | **86** | 44,*72* |
| **Leucine** | **86** | 44,*72* |
| **Lysine** | *101* | *70*,84,112,129 |
| **Methionine** | *104* | 61 |
| **Phenylalanine** | **120** | *91* |
| **Proline** | **70** | |
| **Serine** | 60 | |
| **Threonine** | 74 | |
| **Tryptophan** | **159** | 77,117,**130**,132,**170,171** |
| **Tyrosine** | **136** | *91*,*107* |
| **Valine** | 72 | 41,55,69 |

Bold face indicates strong signal and italics indicates weak. Taken from:

https://proteomicsresource.washington.edu/mascot/help/fragmentation_help.html