RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Structure-based prediction of CAD response factors of dammarane-type tetracyclic triterpenoid saponins and its application to the analysis of saponin contents in raw and processed *Panax Notoginseng*

Ming Peng [a,b], Tong Zhang [*c], Yue Ding [c], Yaxiong Yi [a], Yongjian Yang [b], Jian Le [b]

a. School of Pharmacy, Shanghai University of Traditional Chinese Medicine, Shanghai, 201203, China;

b. Department of Chemistry, Shanghai Institute for Food and Drug Control, Shanghai, 201203, China;

c. Experiment Center for Teaching and Learning, Shanghai University of Traditional Chinese Medicine, Shanghai,201203, China

* Corresponding author: Tel.: +86 21 51322318; Fax: 86 21 51322337

E-mail address: zhangtdmj@hotmail.com

Abstract: The analysis of saponin contents in *Panax Notoginseng* (Sanqi) is challenged by the lacking of authentic reference standards. In this study, a gradient eluted HPLC method coupled with charged aerosol detector (CAD) has been established to solve this problem. The impact of structural features, including the type of aglycon, the optical rotations at C-20, the glycosyl substituent and the glycosyl linkage of dammarane-type tetracyclic triterpenoid saponins on their CAD response factors has been discovered. The rules of the impact have been utilized to predict CAD response factors of saponins in raw and processed notoginseng based on their structures elucidated by LC-QTOFMS. An intensive investigation of the saponin contents in raw (different cultivate places, sizes, and medicinal parts) and processed (steaming, baking, autoclaving, stewing and frying) *Panax Notoginseng* were implemented. This method was successfully applied to distinguishing the quality of raw and processed *Panax Notoginseng*, finding out biomarkers in processed notoginseng, and screening the best processing technique for this herb.

Key words: *Panax Notoginseng*; charged aerosol detector (CAD); dammarane-type tetracyclic triterpenoid saponin; response factor; raw; processed

Abbreviations: Ara: arabinose; CAD: Charge aerosol detector; Chp: Chinese Pharmacopoeia; ESI: electrospray ion; Glc: glucose; Man: mannose; *P. notoginseng*: *Panax notoginseng;* PPD: protopanaxadiol; PPT: protopanaxatriol; QAMS: multi-component with single marker; QTOFMS: quadrupole time-of-flight mass spectrometry; RCF: relative correction factor; RF: response factor; Rha: rhamnose; Xyl: xylose

1. Introduction

Notoginseng, the dry root or rhizome of *Panax notoginseng* (Burk.) F. H. Chen (*P. notoginseng*), also called 'Sanqi' or 'Sanchi', is a precious traditional Chinese medicine with a long history of medical use. The saponin components have been discovered to contribute to the main pharmacological functions of this herb, such as the treatment of cardiovascular diseases,[1] the biological activities of anti-cancer,[2] anti-hyperlipoidemia,[3] and anti-hyperglycemia,[4] *etc*. In traditional Chinese medical applications, processed notoginseng is distinguished from the raw herb by the claim of its capability to "nourish" blood.[5] Furthermore, contemporary researches have reported that processed notoginseng exhibit more potent pharmaceutical activities than raw notoginseng, such as anticancer,[6-8] antiplatelet, anticoagulant, and platelet aggregation inhibition effects,[9] *etc*. Apparently, different compound basis of raw and processed notoginseng directly influences their pharmacological activities.

Dammarane-type tetracyclic triterpenoid saponins have been found to be the major active components in *P. notoginseng*,[10] and can essentially be classified into two types: protopanaxadiol (PPD) and protopanaxatriol (PPT) type. The lacking of authentic reference standards of rare saponins, especially those secondary saponins only existed in processed notoginsengs has impeded the quality control of this herb. Recently, some strategies of quantitative analysis of multi-component with single marker (QAMS) have been developed for the determination of saponin content in *P. notoginseng* mostly based on HPLC-UV and LC-MS platforms. A QAMS method focused on 11 saponins in *P. notoginseng* has been established and validated at UV 200 nm.[11] The slopes of the equations of linear regressions for each saponin were used to calculate the relative correction factor (RCF). Although this method is simple and accurate, the RCFs of each

saponin need to be calculated before the testing on real samples. Moreover, the RCFs of those saponins without authentic reference available are still not achievable and predicted, and the intensive analysis of the whole saponin contents in notoginsengs, especially those in processed herbs could not easily be accomplished. Lai *et al.*[12] developed a green protocol for the utilizing of specific enzymatic hydrolyzing process to calculate relative response factor of specific PPD type saponins, with less consumption of solvent and authentic reference standards. However, this protocol has only focused on 4 PPD saponins so far. Further researches are needed to find the specific enzymes for the hydrolyzing of other types of saponins. Moreover, a HPLC-ESI-MS coupled with mobile-phase compensation method has been investigated for the determination of saponins in *P. notoginseng* calculated based on normalized data of saponin peaks.[13] However, the variations of MS responses of different saponins owing to their structural types and molecular weights could still not be neglected, which limits the extensive application of this method.

Charged aerosol detector (CAD) was firstly introduced in 2002.[14] CAD is a mass sensitive and universal detector for the routine determination of any non-volatile and many semi-volatile chemical species. The liquid mobile phase is nebulized in CAD chamber by N_2 to become aerosol droplets. Then the small droplets containing analytes enter the drying tube, and the big droplets which are composed of the majority of mobile phase enter the wasting tube. After that, the dry particles are mixed with a charged N_2 gas flow which has just passed through the corona discharge needle, and at the meantime the charges are transferred to the dry particles. The charged analyte particles are then collected and the electrical charges are measured with an electrometer. CAD has extensively been applied for the analysis of impurities in pharmaceuticals,[15,16] food products and herbal dietary supplements,[17-19] pharmaceutical formulations,[20,21] and environmental pollutants,[22] *etc*. Moreover, HPLC-CAD has been performed on the analysis of major saponins in raw notogingseng by external standard method using commercially available reference standards.[23,24] However, the content of those minor saponins were not mentioned due to the absence of authentic reference standards. CAD was claimed to be generating identical peak response for all non-volatile substances, however, quite a few studies have also investigated that CAD responses of the analytes are not always the same.[25-27] The variations of the responses may be due to the particle density, hygroscopicity, and volatility, etc., of the analytes in the particle

phase during nebulization.[28] This means that it is inappropriate to arbitrarily assume an identical CAD response for all the saponins which embrace close but different structures without figuring out the their relationships. However, once the relation between the saponin structure and its CAD response is elucidated, this detector is still a convenient and stable detector for the determination of saponins which are short of chromophores in their structure.

In this article, a gradient eluted HPLC-CAD method with post-column mobile phase compensation has been developed to determine the saponin contents in raw and processed *P. notoginseng*. The impact of the structural features, including types of aglycon, optical rotation, glysosyl substituent and glycosyl linkage of dammarane-type tetracyclic triterpenoid saponins on their CAD response factors (RFs) has been discovered. Moreover, the rules have been successfully utilized to predict CAD RFs of the saponins based on their structures, which were identified by LC-QTOFMS in our study, and the prediction has also been validated. An in-depth investigation on saponin contents in raw *P. notoginseng* of different sizes, cultivated places and medicinal parts, as well as the secondary saponins and biomarkers in processed notoginseng of different processing procedures was then implemented.

2. Experimental

2.1. Chemicals and reagents

Reference standards of notoginsenoside R_1, ginsenoside Rg_1, Re, Rf, Rb_1, 20(S)-Rg_2, 20(S)-Rh_1, 20(R)-Rg_2, 20(R)-Rh_1, Rb_2, Rb_3, F_1, Rd, F_2, 20(S)-Rg_3, 20(R)-Rg_3, 20(S)-PPT, CK, 20(R)-Rh_2 and 20(R)-PPD were purchased from Chengdu Must Bio-technology Co. LTD. Ginsenoside 20(S)-Rh_2 and 20(S)-PPD were kindly supplied by Shanghai Pharm Valley Co. LTD. Except for 20(R)-Rh_1 (purity 97.65%), 20(R)-Rh_2 (purity 92.33%) and 20(S)-PPD (purity 95.99%), the purities of all the above reference standards were labeled above 98% by the manufacturers. Reference standard of gypenoside XVII was obtained from Shanghai Winherb Medical Science Co., Ltd (purity 98.93%). Tinidazole (purity 100%) was provided by Zhejiang Supor Pharmaceuticals Co., Ltd. HPLC grade acetonitrile and methanol were obtained from Merck (Darmstadt, Germany). Deionized water was purified using a Milli-Q system (Millipore, Bedford, MA).

2.2. Raw herb of *Panax notoginseng*

Main roots of raw *P. notoginseng* were purchased from two cultivated places in China. Seven different sizes of main root (20, 30, 40, 60, 80, 120 and countless heads) were purchased from Wenshan county, Yunnan province. And five different sizes of main root (30, 40, 60, 120 and countless heads) were purchased from Bobai county, Guangxi province. The term "head" in Panax notoginseng refers to the size of this herb, which has been used in China for a long history. It refers to the number of pieces of notoginseng main roots per 0.5 kg. For example, "40 heads" in this article means that each 0.5 kg of notoginseng herb contains 40 pieces. Apparently, the greater the number of head is, the smaller is the size of main root. In Chinese market, the smaller the number of head is, the more expensive is the notoginseng, because people believe that notoginseng of bigger sizes contain greater amount of total saponins, thus have more potent pharmacology effects. In addition, other medicinal parts of *P. notoginseng* (rhizome, branch root and root hair) were all obtained from Wenshan county. All the raw *P. notoginsengs* were pulverized and screened through an 80 mesh sieve. Water contents of the notoginseng powders were determined by Karl Fischer method.

2.3. Processing procedures of notoginseng

Processed notoginsengs were prepared by different procedures, i.e., steaming, autoclaving, baking, stewing and frying. All the processed notoginsengs were prepared using raw 120-head main root of *P. notoginseng* cultivated in Wenshan county. All the processed notoginsengs were dried under vacuum at 80 ℃ for 48 hr before being pulverized. Then the pulverized powder was being screened through an 80 mesh sieve. Karl Fischer titration was performed afterwards to determine the water contents of all processed notoginseng powder.

2.3.1 Steaming

Raw notoginsengs were steamed at 100℃ in a steamer for 1, 2, 3 and 4 hrs. Before being steamed, all the raw notoginsengs were soaked in water for 2 hrs.

2.3.2 Autoclaving

Raw notoginsengs were steamed in an autoclave at 100 or 120 ℃ for 2, 4, 6, 8, 12, 24, 36, 48 and 72 hrs, respectively. Before being autoclaved, all the raw notoginsengs were soaked in water for 2 hrs.

2.3.3. Baking

Raw notoginsengs were baked in an electric blast drying oven at 100 or 120 ℃ for 24, 48 and 72 hrs, respectively.

2.3.4. Stewing

Raw notoginseng powder (80 mesh) were stewed in water. After the water was boiling for 10 minutes, the remaining notoginseng powder and liquid was collected and evaporated to dryness.

2.3.5. Frying

Raw notoginsengs were steamed until tender before being cut into slices. Then the notoginseng slices were fried in tea-seed oil over moderate heat till both sides of the slices appeared to be golden in color.

2.4 Preparation of reference standard stock solution and internal standard solution

The stock reference solution was prepared using 70% methanol aqueous solution. Since the contents of different saponins vary greatly, the concentrations of 22 reference standards in stock solution ranged from 0.05~1.5 mg/mL according to the saponin contents in notoginseng. Tinidazole was dissolved in 70% methanol aqueous solution to prepare an internal standard (IS) solution at the concentration of 2 mg/mL.

2.5. Preparation of sample solution

About 0.5 g of notoginseng powder was accurately weighed and transferred into a 20 mL volumetric flask. Then 5 mL of IS solution and 10mL of 70% methanol aqueous solution was added into the volumetric flask. The flask was ultra-sonicated (500 W, 50 Hz) for 60 min, and 70% methanol aqueous solution was added to volume afterwards. After the sample solution was shaken well and standing for a while, the supernatant was withdrawn and filtered through a 0.22 μm polytetrafluoroethylene (PTEE) filter.

2.6. HPLC Chromatographic conditions

HPLC analysis was performed on a Dionex Ultimate 3000 series HPLC system, equipped with vacuum degasser, dual gradient pump, autosampler, ultraviolet detector, and Corona Ultra charged aerosol detector (Munich, Bavaria, Germany). A Waters HSS C18 column (25 cm×4.6 mm, i.d., 3.5 μm, Ireland) was used for chromatographic separation at 30 ℃. The separation was achieved using a binary gradient elution system consisted of water and acetonitrile (ACN) as mobile phases. The gradient program was as follows: 0-31 min (20.5% ACN), 31-32 min (20.5%→30% ACN),

32-50.5 min (30%→35% ACN), 50.5-61 min (35%→50% ACN), 61-81 min (50%→90% ACN), 81-91 min (90% ACN). The flow rate was set at 1.0 mL/min and the sample injection volume was 10 μL. The sample elution was eluted to UV and CAD detector successively. UV detector wavelength was set at 203 nm. CAD data collection frequency was 2 Hz, and nebulizer temperature was 35 ℃.

In order to keep the organic modifier content to be constant when the mobile phases reached CAD detector, post column compensation of mobile phases was introduced. Since dual gradient pumps of this HPLC system had a slight difference in the dead volume, the post column counter gradient program for CAD detector was set for a 0.3min's delay, which was as follows: 0-31.3 min (79.5% ACN), 31.3-32.3 min (79.5%→70% ACN), 32.3-50.8 min (70%→65% ACN), 50.8-61.3 min (65%→50% ACN), 61.3-81.3 min (50%→10% ACN), 81.3-91.0 min (10% ACN), with the total flow rate of 1.0 mL/min.

2.7 Validation of HPLC method

2.7.1 Calibration curves

The stock reference solution was serially diluted by 70% methanol aqueous solution to prepare 7 levels of calibration standard solutions for 22 saponins with authentic reference standards available. For instance, the concentrations of ginsenoside R_1 in calibration standard solutions were 0.0048, 0.120, 0.240, 0.480, 0.1201, 0.2402 and 0.3603 mg/mL (L1~L7), respectively. In each level of calibration standard solution, IS maintained a constant concentration of 0.5 mg/mL. Each calibration standard solution was injected in triplicate. And the calibration curves were established by plotting the peak area ratio of each analyte versus IS against the concentration of each analyte.

2.7.2 Limits of detection (LOD) and limits of quantification (LOQ)

LOD and LOQ of each analyte were calculated on the peak response at signal-to-noise (S/N) of 3 and 10, respectively.

2.7.3 Accuracy

The accuracy of this HPLC method was evaluated by the recovery test. For the preparation of each recovery sample solution, 0.25 g of raw notoginseng powder (120 heads, Yunnan) was transferred to a 20 mL volumetric flask. And then different quantities of individual stock reference

199 solution were added in to prepare 3 spiked concentration levels. In each recovery sample solution,
200 IS concentration was maintained to be 0.5 mg/mL constantly. For each spiked level, recovery
201 sample solution was prepared in triplicate.

202 2.7.4. Precision

203 Precision of this method were evaluated by intra-day precision, injection precision, and
204 sample repeatability, respectively. The intra-day precisions were calculated based on the variations
205 of the accuracies in recovery tests. For the validation of injection precision, 10 μL of reference
206 standard solution was injected in triplicate, and the variations of the area ratios of each analyte
207 versus IS were calculated.

208 2.8. LC-MS conditions

209 LC-MS data were acquired on an Agilent 1290 Infinity UPLC coupled with Agilent 6538
210 UHD Accurate-Mass QTOF LC/MS system and an ESI source (Agilent Technologies, Santa Clara,
211 USA). The chromatographic conditions, including type of chromatographic column, column
212 temperature, gradient elution, flow rate and injection volume, were exactly the same as those of
213 HPLC-CAD system, except that the mobile phase A was 0.01% formic acid instead of water. A
214 post-column tee joint was used to split the flow rate, and the actual flow rate passed through ESI
215 source kept 0.2 mL/min constantly. The optimized mass parameters were as follows: electrospray
216 ion (ESI) source; gas temperature, 350 ℃; drying gas (N_2), 10 L/min; nebulizing gas pressure, 40
217 psig; capillary voltage, 3500 V; capillary current, 0.032 μA; chamber current, 2.20 μA. ESI
218 negative and positive modes were performed on both MS and tandem MS in the m/z range of
219 100~1400. Besides, MS/MS analysis was achieved using collision energies of 10 V, 20 V, 30 V
220 and 40 V, respectively. Prior to mass data acquisition, the mass spectrometry was tuned and
221 optimized using Agilent ESI-L low concentration Tunning mix (lot：LB95102). The accurate mass
222 was measured to identify the structures of saponins in raw and processed notoginsengs using
223 Agilent MassHunter Workstation software (Version B.04.00).

224 2.9. Prediction of CAD response

225 In this article, 22 dammarane-type tetracyclic triterpenoid saponins or aglycons with
226 authentic reference standards were utilized to set up the HPLC-CAD method. And these above 22
227 saponins have covered the scope of different structure features, including type of aglycon, optical

rotation at C-20, glycosyl substituent and glycosyl linkage, of the saponins basically existed in notoginsengs. In this experiment, the slope of calibration curve for each known saponin was regarded as the RF of its chromatographic peak. According to our results, the differences of CAD RFs of saponins with different structures were much smaller than UV RFs. The impact of structural features of PPD and PPT saponins on the CAD RFs has been investigated. Since the structures of all the saponins in notoginsengs have already been identified by LC-MS, the CAD RFs of saponins without authentic reference standards could then be predicted.

2.10. Validation of the prediction of CAD RFs

The prediction of CAD RFs was validated by some of the saponins with authentic reference standards. The differences between their CAD slopes obtained from the linear regressions and their predicted CAD RFs are calculated as: $\frac{ABS(\mathrm{Predicted\ CAD\ RF} - \mathrm{CAD\ slope})}{\mathrm{CAD\ slope}} \times 100\%$.

Ten PPT type saponins with authentic reference standards in this article were regarded as sample saponins so that their CAD RFs were going to be assigned based on the rule of impact of their structural features. Furthermore, another authentic reference standard of PPD type saponin, namely gypenoside XVII, was chosen to perform the validation. The stock validation solution was prepared using 70% methanol aqueous solution. A series of validation solutions were prepared by diluting the stock validation solution for 6 concentration levels, and the concentrations were between 0.0041 ~ 0.1220 mg/mL. In each level of validation solutions, IS maintained a constant concentration of 0.5 mg/mL. The validation solutions were then injected into HPLC under the developed chromatographic condition. Then the CAD linear regression was set up by plotting the peak area ratio versus IS against the concentration. The difference between CAD slope of linear regression and the predicted RF value was calculated.

2.11. Saponin content determination of notoginseng

The saponin contents of real notoginseng samples were calculated using internal standard method based on the CAD peak area ratio versus IS for each analyte. For the saponins with authentic reference standard available, an internal standard curve method was applied based on their validated calibration curve equations. Nevertheless, for those saponins without reference standards available in this experiment, their CAD RFs were predicted based on the structures

identified by LC-MS. Thus, the content of each saponin was calculated using a simple internal standard method according to the predicted CAD RFs.

3. Results and discussion

3.1 Optimisation of gradient elution program

Based on the polarities, the 22 saponins with authentic reference standards available in this experiment can be divided into three groups: (1) high polarity saponins, including R_1 and Rg_1; (2) medium polarity saponins, including Re, Rf, Rb_1, 20(S)-/20(R)-Rg_2, 20(S)-/20(R)-Rh_1, Rb_2, Rb_3, F_1 and Rd; (3) low polarity saponins, including F_2, 20(S)-/20(R)-Rg_3, CK, 20(S)-/20(R)-Rh_2, 20(S)-PPT and 20(S)-/20(R)-PPD. Since there is one more hydroxyl group in PPT aglycon than in PPD aglycon, the polarities of PPT type saponins are basically higher than those of PPD type saponins, given the same number and type of glycosyl substituents. The most difficult part of establishing this LC method is the separation of optical isomers, e.g., 20(S)-/20(R)-Rg_3 and 20(S)-/20(R)-Rh_2, and geometric isomers at C-20 position, e.g., RK_3/Rh_4, and RK_1/Rg_5, since these isomers bear very similar polarities and close chemical properties. Gradient elution program was set up, and the saponin peaks were successfully separated. In the sample chromatogram of raw notoginsengs, the retention times of all the saponins were within 60 min, which indicated that the saponins were of high and medium polarities. However, upon steaming for 3 hrs, low polar saponins emerged in processed notoginseng. Furthermore, the content of low polar saponins increased dramatically when notoginsengs were autoclaved at 120 ℃ for 18 hrs (Fig. 1).

3.2 Optimisation of post column compensation program of CAD detector

One of the distinguishing features of CAD detector is that different compounds theoretically exhibit similar responses. However, the amount of organic modifier in the mobile phase significantly influences the CAD response. It has been reported that with the increasing of organic modifier from 0% to 100% in the mobile phase, the CAD response may increase dramatically from 5 to 10 times.[29] Unfortunately, the wide range of polarities of saponins in raw and processed notoginsengs hinders the use of isocratic elution of mobile phase. Dual pump HPLC system has been invented to overcome this restraint. A counter gradient program was designed for post column compensation, and a constant quantity of organic modifier in mobile phase reached the CAD detector at any time. Due to a slight difference in the dead volumes of the dual pump system,

a 0.3-minute time lag was considered in the post column counter gradient program.

The total content of ACN in mobile phase in post-column gradient elution was evaluated. It has been found that if ACN content increased from 20% to 50%, the peak height and S/N ratio of R_1 increased by 40% and 100%, respectively. However, if ACN content further increased to 80%, the peak height and S/N ratio of R_1 decreased dramatically by 200% and 400%, respectively. This result indicated that a higher organic modifier content may not always bring better CAD responses to analytes. Thus, the optimal post-column gradient program was established, where a constant content of 50% ACN was eluted to CAD detector, and the baseline drifting caused by pre-column gradient elution was effectively avoided.

3.3 Validation of HPLC-CAD method

Linearity, LOQ, accuracy, and precision were validated by this developed HPLC-CAD method. The linearities and LOQs were also compared with those results simultaneously acquired at UV 203nm. The correlation coefficients (r) of most of the saponins were above 0.999, and those of the rest of the analytes were above 0.994, which basically met the requirements of quantification determination. In comparison to CAD results, UV detector generally provided higher correlation coefficients (>0.999). However, LOQs of CAD for most of the saponins were obviously lower than those of UV detector, manifesting a higher sensitivity of CAD compared with UV detector. The RSDs for injection precision were all below 2.6% (n=3) (Table 1). The developed HPLC-CAD method proved to be accurate by the explanation of recoveries. Mean recoveries of 3 spiked concentration levels for each analyte were ranged from 85.7%-112.9%, with the RSDs within 6.8%, which could also be interpreted as intra-day precision. The method validation results of recovery and intra-day precision for each saponin could be found in Table 1S. These results demonstrated that the HPLC-CAD method was linear, sensitive, accurate and precise for quantification of saponins in notoginseng sample. Although the linearity results of CAD detector were inferior to those of UV detector, this proposed HPLC-CAD method was more sensitive, and was a reliable method for the analysis of saponins in raw and processed notoginseng.

312 **Table 1.** Method validation results of linearity test and limit of quantitation of CAD and UV detector for each saponin

Saponin	RT (min)	linearity range (mg/ml)	Calibration curve (CAD, n=7)	LOQ（CAD, ng）	Calibration curve（UV, n=7）	LOQ（UV, ng）	Injection precision (RSD, %, n=3)
R_1[a]	20.6	0.0048~0.3603	Y=2.7382X+0.0047, r=0.9994	40.0	Y=0.3783X+0.0001,r=0.9995	80.1	0.52
Rg_1	29.7	0.0142~1.0669	Y=2.5478X+0.0803, r=0.996	118.5	Y=0.5167X-0.0031,r=0.9999	118.5	0.28
Re	31.6	0.0024~0.1790	Y=3.1448X-0.0031, r=0.9998	79.5	Y=0.3871X-0.0007,r=0.9996	79.5	0.68
Rf	43.9	0.00146~0.1096	Y=3.2941X+0.0015, r=0.9994	24.4	Y=0.5165X+0.0002,r=0.9994	24.4	0.20
Rb_1	47.0	0.0104~0.7793	Y=1.8034X+0.0567, r=0.994	86.6	Y=0.3362X-0.0011,r=1.0000	86.6	0.17
20(S)-Rg_2	48.7	0.00224~0.1681	Y=3.1435X+0.0071, r=0.998	18.7	Y=0.5569X-0.0001,r=0.9999	37.4	0.96
20(S)-Rh_1	49.0	0.00291~0.2181	Y=3.1826X+0.0082, r=0.999	24.2	Y=0.6615X-0.0005,r=0.9999	48.5	0.12
20(R)-Rg_2	49.9	0.00247~0.1850	Y=3.1004X+0.0016, r=0.9992	20.6	Y=0.5309X-0.0002,r=0.9998	41.1	0.26
20(R)-Rh_1	51.0	0.00256~0.1919	Y=3.1374X+0.0028, r=0.9992	21.3	Y=0.6424X-0.0005,r=0.9999	42.6	0.41
Rb_2	51.7	0.00141~0.1056	Y=3.1191X-0.0050, r=0.9998	23.5	Y=0.3619X-0.0001,r=0.9998	23.5	0.90
Rb_3	52.6	0.00284~0.3267	Y=3.0678X-0.0043, r=0.9997	25.0	Y=0.3517X-0.0002,r=0.9999	25.0	0.43
F_1	54.9	0.00125~0.0934	Y=3.0166X-0.0034, r=0.9997	20.8	Y=0.5650X-0.0002,r=0.9999	83.0	0.12
R_d	56.5	0.00238~0.1787	Y=2.6787X+0.0044, r=0.998	19.9	Y=0.4191X-0.0003,r=1.0000	39.7	0.22
F_2	63.6	0.00158~0.1188	Y=2.9717X+0.0014, r=0.999	13.2	Y=0.5272X-0.0003,r=1.0000	26.4	0.45
20(S)-Rg_3	66.5	0.00082~0.0616	Y=3.5249X-0.0012, r=0.9995	13.7	Y=0.5836X-0.0001,r=1.0000	13.7	0.37
20(R)-Rg_3	67.1	0.00057~0.0426	Y=3.4383X-0.0026, r=0.995	9.5	Y=0.5329X+0.0001,r=0.9991	9.5	0.56
20(S)-PPT	69.2	0.00164~0.1230	Y=3.2837X+0.0032, r=0.999	13.7	Y=0.9282X-0.0005,r=0.9999	54.7	0.30
CK	72.8	0.00101~0.0757	Y=2.9136X-0.0030, r=0.9997	33.6	Y=0.5789X-0.0000,r=0.999	33.6	0.75
20(S)-Rh_2	74.3	0.00078~0.0581	Y=3.9694X-0.0028, r=0.9998	12.9	Y=0.7483X-0.0003,r=1.0000	25.8	1.05
20(R)-Rh_2	74.6	0.00049~0.0364	Y=4.0073X-0.0036, r=0.9992	8.1	Y=0.6942X+0.0002,r=1.0000	16.2	0.24
20(S)-PPD	86.0	0.00111~0.0831	Y=3.6304X-0.0017, r=0.9996	18.5	Y=0.8634X-0.0005,r=0.9999	36.9	0.26
20(R)-PPD	86.6	0.00066~0.0492	Y=3.6493X-0.0032, r=0.9996	10.9	Y=0.8227X-0.0002,r=0.9999	21.9	2.57

313 a. More data could be found in the Supplementary material.

314

315

316 3.4 Structural identification of saponins in notoginsengs

317 LC-QTOFMS analysis both in ESI negative and positive mode was utilized for structural
318 identification of the saponins in notoginseng. For those saponins with authentic reference
319 standards available in this experiment, their retention times of chromatographic peaks were further
320 confirmed. The addition of 0.01% formic acid in mobile phase A which facilitated the ionization
321 of saponins would not affect the retention times of TIC chromatographic peaks, and the retention
322 time differences were less than 0.7% for the all saponins with authentic reference standards.
323 Results suggested that ESI negative mode provided stronger mass spectrometric signals for PPD
324 and PPT type saponins under this chromatographic condition. Most of the saponins existed in
325 notoginseng could acquire $[M-H]^-$ and/or $[M+HCOO]^-$ under ESI negative mode, which provided
326 explicit information on molecular weights of those undetermined saponins. Moreover, mass peaks
327 of m/z 459 and m/z 475, which represent the ion fragments of $[PPD-H]^-$ and $[PPT-H]^-$,
328 respectively, always appeared in MS spectrograms in ESI negative mode for PPD and PPT type
329 saponins.

330 The glycosyl substitutents in notoginseng mainly include glucose (GLC), xylose (Xyl),
331 arabinose (Ara), rhamnose (Rha), and mannose (Man), etc. For instance, two most characteristic
332 ion peaks for Glc in ESI negative mode are m/z 161, i.e., $[(Glc-H_2O)-H]^-$, and m/z 101, which
333 represents the fragment as a result of a neutral loss of $C_2H_4O_2$ group from m/z 161. Furthermore,
334 the structure transformation of saponins based on collision-induced dissociation provides detailed
335 information on the identification of glycosyl substitutents. For example, if there is a loss of
336 m/z162, it usually indicates the loss of a $Glc-H_2O$ fragment. Nevertheless, the loss of m/z146, 132
337 and 324 practically suggest the loss of $Rha-H_2O$, $Xyl-H_2O$ or $Ara-H_2O$, and $Glc-Glc-2H_2O$
338 fragment, respectively.

339 Moreover, ESI positive mode was also a good means for the identification of $[M+H]^+$ and
340 $[M+Na]^+$ for those saponins with fairly weak molecular ion peaks in ESI negative mode. Besides,
341 some mass spectrometric fragments collected in ESI positive mode provided essential clues for the
342 structural confirmation of saponin agylcons. For instance, the existing of a series of fragments
343 including m/z 443, 425 and 407 practically indicate the ion fragments of $[PPD-H_2O+H]^+$,
344 $[PPD-2H_2O+H]^+$ and $[PPD-3H_2O+H]^+$, respectively. However, the occurrence of the series of m/z

441, 423 and 405 imply the ion fragments of $[PPT\text{-}2H_2O+H]^+$, $[PPT\text{-}3H_2O+H]^+$ and $[PPT\text{-}3H_2O+H]^+$, respectively.

Raw and processed notoginseng sample solutions were analyzed by LC-ESI-QTOFMS. At an analytical level, a total of 16 saponins were identified in raw notoginseng, and 25 more saponins were found in processed samples. The structures of the saponins in raw and processed *P. notoginseng* are given in Fig. 2, and the observed precursor and product ions of saponins are listed in Table 2a and 2b. What should be mentioned is that different medicinal part, including main root, branch root, rhizome and root hair, of notoginseng did not show difference in the saponin components. The differences of saponin molecular weights obtained by MassHunter software and the results inferred from molecular formulas were all below 4ppm. The structure skeletons of the saponins in notoginseng included PPD, PPT, C-20 dehydrated PPD, C-20 dehydrated PPT, 25-OH PPD, and 25-OH PPT. If the saponins are classified by the aglycons, 16 PPD type saponins were found, where 5 original saponins in raw samples and 11 more secondary saponins in processed samples were identified. Nevertheless, 25 PPT type saponins were discovered, in which 11 original saponins and 24 secondary saponins were figured out.

360 **Table 2a.** Precursor and product ions of saponins in raw notoginsengs using LC-QTOFMS

No.	Peak Identification	R_t (min)	Theoretical accurate mass (m/z)	Experimental (m/z) (ESI-) or (ESI+)	Mass accuracy (ppm)	CID (m/z)
1	20-O-Glucoginsenoside R_f	9.0	961.5378[M-H]$^-$	961.5372[M-H]$^-$	0.62	637.4300[M-H-2(Glc-H_2O)]$^-$,475.3781[M-H-3(Glc-H2O)]$^-$, 323.1008[(Glc-Glc)-2H_2O-H]$^-$, 161.0459[(Glc-H_2O)-H]$^-$, 101.0243[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$ 441.3726[PPT-2H_2O+H]$^+$, 423.3621[PPT-3H_2O+H]$^+$, 405.3516[PPT-4H_2O+H]$^+$
2	Notoginsenoside R_3	15.3	961.5378[M-H]$^-$	961.5372[M-H]$^-$	0.62	799.4883[M-H-(Glc-H_2O)]$^-$, 637.4343[M-H-2(Glc-H_2O)]$^-$, 475.3793[M-H-3(Glc-H_2O)]$^-$, 161.0453[(Glc- H_2O)-H]$^-$, 101.0241[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$ 441.3726[PPT-2H_2O+H]$^+$, 423.3622[PPT-3H_2O+H]$^+$, 405.3516[PPT-4H_2O+H]$^+$
3	Notoginsenoside R_1*	18.0	931.5272[M-H]$^-$	931.5270[M-H]$^-$	0.21	799.4885[M-H-(Xyl-H_2O)]$^-$, 637.4349[M-H-(Xyl-H_2O)-(Glc-H_2O)]$^-$, 475.3835[M-H-(Xyl-H_2O)-2(Glc-H_2O)]$^-$, 161.0466[(Glc-H_2O)-H]$^-$, 101.0253[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$
4	Ginsenoside Rg_1*	25.6	845.4904[M+HCOO]$^-$	845.4913[M+HCOO]$^-$	1.06	799.4874[M-H]$^-$, 637.4352[M-H-(Glc-H_2O)]$^-$, 475.3816[M-H-2(Glc-H_2O)]$^-$, 161.0455[(Glc-H_2O)-H]$^-$, 101.0248[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$
5	Ginsenoside Re*	27.3	945.5428[M-H]$^-$	945.5426[M-H]$^-$	0.21	799.4885[M-H-(Rha-H_2O)]$^-$, 783.4897[M-H-(Glc-H_2O)]$^-$, 637.4327[M-H-(Rha-H_2O)-(Glc-H_2O)]$^-$, 475.3811[M-H-(Rha-H_2O)-2(Glc-H_2O)]$^-$, 101.0263[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$
6	Malonyl-ginsenoside Rg_1	34.9	885.4853[M-H]$^-$	885.4855[M-H]$^-$	0.23	799.4822[M-H-Mal]$^-$, 781.4746[M-H-Mal-H_2O]$^-$, 679.4442[M-H-(Glc-H_2O)]$^-$, 637.4328[M-H-Mal-(Glc-H_2O)]$^-$, 475.3798[M-H-Mal-2(Glc-H2O)]$^-$, 161.0451[(Glc- H_2O)-H]$^-$, 101.0243[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$ 441.3729[PPT-2H_2O+H]$^+$, 423.3624[PPT-3H_2O+H]$^+$, 405.3517[PPT-4H_2O+H]$^+$

7	Yesanchinoside D	37.1	887.5010[M+HCOO]$^-$	887.5012[M+HCOO]$^-$	0.23	841.4956[M-H]$^-$, 799.4872[M-H-$COCH_2$]$^-$, 781.4737[M-H-$COCH_2$-H_2O]$^-$, 637.4217[M-H-$COCH_2$-(Glc-H_2O)]$^-$, 619.4226[M-H-$COCH_2$-(Glc-H_2O)-H_2O]$^-$, 475.3801[M-H-$COCH_2$-2(Glc-H_2O)]$^-$, 161.0453[(Glc-H_2O)-H]$^-$, 101.0246[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$ 441.3727[PPT-2H_2O+H]$^+$, 423.3622[PPT-3H_2O+H]$^+$, 405.3517[PPT-4H_2O+H]$^+$
8	Notoginsenoside R_4	41.2	1239.6379[M-H]$^-$	1239.6373[M-H]$^-$	0.48	1107.5932[M-H-(Xyl-H_2O)]$^-$, 1077.5821[M-H-(Glc-H_2O)]$^-$, 945.5419[M-H-(Xyl-H_2O)-(Glc-H_2O)]$^-$, 783.4915[M-H-(Xyl-H_2O)-2(Glc-H_2O)]$^-$, 621.4379[M-H-(Xyl-H_2O)-3(Glc-H_2O)]$^-$ 443.3885[PPD-H_2O+H]$^+$, 425.3777[PPD-2H_2O+H]$^+$, 407.3674[PPD-3H_2O+H]$^+$, 325.1128[(Glc-Glc)-2H_2O+H]$^-$
9	Notoginsenoside Fa	43.1	1239.6379[M-H]$^-$	1239.6372[M-H]$^-$	0.56	1107.5947[M-H-(Xyl-H_2O)]$^-$, 945.5433[M-H-(Xyl-H_2O)-(Glc-H_2O)]$^-$, 783.4858[M-H-(Xyl-H_2O)-2(Glc-H_2O)]$^-$, 621.4373[M-H-(Xyl-H_2O)-3(Glc-H_2O)]$^-$, 161.0452[(Glc- H_2O)-H]$^-$, 101.0241[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$ 443.3886[PPD-H_2O+H]$^+$, 425.3779[PPD-2H_2O+H]$^+$, 407.3674[PPD-3H_2O+H]$^+$, 325.1129[(Glc-Glc)+H-2H_2O]$^+$
10	20(S)-Notoginsenoside R_2	45.0	769.4744[M-H]$^-$	769.4748[M-H]$^-$	0.52	637.4339[M-H-(Xyl-H_2O)]$^-$ 475.3789[M-H-(Xyl-H_2O)-(Glc-H_2O)]$^-$, 161.0449[(Glc- H_2O)-H]$^-$, 101.0244[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$ 459.3834[PPT-H_2O+H]$^+$, 441.3728[PPT-2H_2O+H]$^+$, 423.3625[PPT-3H_2O+H]$^+$, 405.3517[PPT-4H_2O+H]$^+$
11	Ginsenoside Rb_1*	45.8	1107.5957[M-H]$^-$	1107.5960[M-H]$^-$	0.27	945.5443[M-H-(Glc-H_2O)]$^-$, 927.5267[M-H-(Glc-H_2O)-H_2O]$^-$, 783.4933[M-H-2(Glc-H_2O)]$^-$, 765.4803[M-H-2(Glc-H_2O)-H_2O]$^-$, 621.4370[M-H-3(Glc-H_2O)]$^-$, 459.3871[M-H-4(Glc-H_2O)]$^-$, 423.4257[M-H-4(Glc-H_2O)-2H_2O]$^-$, 323.1072 [(Glc-Glc)-2H_2O-H]$^-$

12	20(S)-Ginsenoside Rg_2*	47.5	829.4955 $[M-HCOO]^-$	829.4962 $[M-HCOO]^-$	0.84	783.4931$[M-H]^-$, 637.4324$[M-H-(Rha-H_2O)]^-$, 475.3810$[M-H-(Rha-H_2O)-(Glc-H_2O)]^-$, 161.0451$[(Glc-H_2O)-H]^-$, 101.0245$[(Glc-H_2O)-H-C_2H_4O_2]^-$
13	20(S)-Ginsenoside Rh_1*	47.9	683.4376$[M+HCOO]^-$	683.4387$[M+HCOO]^-$	1.61	637.4359$[M-H]^-$, 475.3820$[M-H-(Glc-H_2O)]^-$, 161.0450$[(Glc-H_2O)-H]^-$, 101.0250$[(Glc-H_2O)-H-C_2H_4O_2]^-$
14	Ginsenoside F_1*	53.6	683.4376$[M+HCOO]^-$	683.4385$[M+HCOO]^-$	1.31	637.4334$[M-H]^-$, 475.3790$[M-H-(Glc-H_2O)]^-$, 161.0442$[Glc-H_2O-H]^-$, 101.0250$[(Glc-H_2O)-H-C_2H_4O_2]^-$
			661.4266$[M+Na]^+$	661.4290$[M+Na]^+$	3.63	459.3834$[PPT-H_2O+H]^+$, 441.3728$[PPT-2H_2O+H]^+$, 423.3623$[PPT-3H_2O+H]^+$, 405.3517$[PPT-4H_2O+H]^+$
15	Ginsenoside Rd*	55.4	945.5428$[M-H]^-$	945.5430 $[M-H]^-$	0.21	783.4950$[M-H-(Glc-H_2O)]^-$, 621.4379$[M-H-2(Glc-H_2O)]^-$, 161.0464$[(Glc-H_2O)-H]^-$, 101.0247$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 443.3855$[PPD-H_2O+H]^+$, 425.3784$[PPD-2H_2O+H]^+$, 407.3677$[PPD-3H_2O+H]^+$, 325.1129$[(Glc-Glc)+H-2H_2O]^+$
16	Gypenoside-XVII	57.4	945.5428$[M-H]^-$	945.5426$[M-H]^-$	0.21	621.4857$[M-H-2(Glc-H_2O)]^-$, 475.8270$[M-H-3(Glc-H_2O)]^-$, 323.0988$[(Glc-Glc)-2H_2O]^-$, 161.0456$[(Glc-H_2O)-H]^-$, 101.0242$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 443.3887$[PPD-H_2O+H]^+$, 425.3779$[PPD-2H_2O+H]^+$, 407.3674$[PPD-3H_2O+H]^+$, 325.1128$[(Glc-Glc)+H-2H_2O]^+$

361 * These compounds have been further confirmed by the peak retention time of authentic reference standard

362

363 **Table 2b.** Precursor and product ions of saponins in processed notoginsengs using LC-QTOFMS (Saponins originally existed in raw notoginseng are not included in
364 this table)

No.	Peak Identification	R_t (min)	Theoretical accurate mass (m/z)	Experimental (m/z) (ESI-) or (ESI+)	Mass accuracy (ppm)	CID (m/z)
17	20(S)-25-OH Ginsenoside Rh_1	11.8	701.4482 $[M\text{-}HCOO]^-$	701.4490$[M\text{-}HCOO]^-$	1.14	655.4433$[M\text{-}H]^-$, 493.3905$[M\text{-}H\text{-}(Glc\text{-}H_2O)]^-$, 161.0454$[(Glc\text{-}H_2O)\text{-}H]^-$, 101.0245$[(Glc\text{-}H_2O)\text{-}H\text{-}C_2H_4O_2]^-$ 639.4479$[M+H\text{-}H_2O]^+$, 477.3942$[M+H\text{-}H_2O\text{-}(Glc\text{-}H_2O)]^+$, 459.3841$[PPT\text{-}H_2O+H]^+$, 441.3730 $[PPT\text{-}2H_2O+H]^+$, 423.3623$[PPT\text{-}3H_2O+H]^+$, 405.3518$[PPT\text{-}4H_2O+H]^+$
18	20(R)-25-OH Ginsenoside Rh_1	15.0	701.4482$[M\text{-}HCOO]^-$	701.4492$[M\text{-}HCOO]^-$	1.43	655.4427$[M\text{-}H]^-$, 493.3907 $[M\text{-}H\text{-}(Glc\text{-}H_2O)]^-$, 161.0453$[(Glc\text{-}H_2O)\text{-}H]^-$, 101.0246$[(Glc\text{-}H_2O)\text{-}H\text{-}C_2H_4O_2]^-$ 639.4481$[M+H\text{-}H_2O]^+$, 477.3942$[M+H\text{-}H_2O\text{-}(Glc\text{-}H_2O)]]^+$, 441.3730 $[PPT\text{-}2H_2O+H]^+$, 423.3623$[PPT\text{-}3H_2O+H]^+$, 405.3521$[PPT\text{-}4H_2O+H]^+$
19	20(S)-Rh_1 (Man as glycosyl substituent)	39.4	683.4376$[M\text{-}HCOO]^-$	683.4385$[M\text{-}HCOO]^-$	1.32	637.4332$[M\text{-}H]^-$, 475.3811$[M\text{-}H\text{-}(Man\text{-}H_2O)]^-$, 161.0455$[Man\text{-}H]^-$, 101.0242$[Man\text{-}H\text{-}C_2H_4O_2]^-$
20	20(R)-Rh_1 (Man as glycosyl substituent)	40.5	683.4376 $[M\text{-}HCOO]^-$	683.4384$[M\text{-}HCOO]^-$	1.17	637.4333 $[M\text{-}H]^-$, 475.3771 $[M\text{-}H\text{-}(Man\text{-}H_2O)]^-$, 161.0447$[Man\text{-}H]^-$, 101.0235$[Man\text{-}H\text{-}C_2H_4O_2]^-$
21	20(R)-Ginsenoside Rg_2*	48.4	829.4955 $[M\text{-}HCOO]^-$	829.4953$[M\text{-}HCOO]^-$	0.24	783.4926 $[M\text{-}H]^-$, 637.4548$[M\text{-}H\text{-}(Rha\text{-}H_2O)]^-$, 161.0444$[Man\text{-}H]^-$ 807.4870$[M+Na]^+$, 639.4472$[M+H\text{-}(Rha\text{-}H_2O)]^+$, 477.3940$[M+H\text{-}(Rha\text{-}H_2O)\text{-}(Glc\text{-}H_2O)]^+$, 441.3729$[PPT\text{-}2H_2O+H]^+$, 423.3626 $[PPT\text{-}3H_2O+H]^+$, 405.3519$[PPT\text{-}4H_2O+H]^+$
22	20(R)-Ginsenoside Rh_1*	49.5	683.4376$[M\text{-}HCOO]^-$	683.4386M-$HCOO]^-$	1.346	637.4361 $[M\text{-}H]^-$, 475.3836$[M\text{-}H\text{-}(Glc\text{-}H_2O)]^-$, 161.0456$[(Glc\text{-}H_2O)\text{-}H]^-$, 101.0248$[(Glc\text{-}H_2O)\text{-}H\text{-}C_2H_4O_2]^-$ 661.4294$[M+Na]^+$, 621.4369$[M+H\text{-}H_2O]^+$, 603.4265$[M+H\text{-}2H_2O]^+$, 459.3834$[M+H\text{-}H_2O\text{-}(Glc\text{-}H_2O)]^+$, 441.3731$[PPT\text{-}2H_2O+H]^+$, 423.3629 $[PPT\text{-}3H_2O+H]^+$, 405.3519$[PPT\text{-}4H_2O+H]^+$

23	25-OH Ginsenoside Rg_3	50.9	801.5006$[M-H]^-$	801.5008$[M-H]^-$	0.25	639.4473$[M-H-(Glc-H_2O)]^-$, 477.3960$[M-H-2(Glc-H_2O)]^-$, 101.0245$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 785.5064$[M+H-H_2O]^+$, 767.4949$[M+H-2H_2O]^+$, 749.4846$[M+H-3H_2O]^+$, 623.4524$[M+H-H_2O-(Glc-H_2O)]^+$, 461.3992$[M+H-H_2O-2(Glc-H_2O)]^+$, 443.3884$[PPD+H-H_2O]^+$, 425.3780$[PPD+H-2H_2O]^+$, 407.3674$[PPD+H-3H_2O]^+$, 325.1987$[Glc-Glc-2H_2O+H]^+$
24	Gypenoside LXXV	60.5	783.4900$[M-H]^-$ 785.5046$[M+H]^+$	783.4904$[M-H]^-$ 785.5063$[M+H]^+$	0.51 2.16	621.4378$[M-H-(Glc-H_2O)]^-$, 459.3802$[M-H-2(Glc-H_2O)]^-$, 161.0451$[(Glc-H_2O)-H]^-$, 101.0246$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 623.4525$[M+H-(Glc-H_2O)]^+$, 461.3991$[M+H-2(Glc-H_2O)]^+$, 443.3877$[PPD+H-H_2O]^+$, 425.3775$[PPD+H-2H_2O]^+$, 407.3674$[PPD+H-3H_2O]^+$, 325.1136$[(Glc-Glc)-2H_2O+H]^+$
25	Gypenoside LXXV isomer	61.2	783.4900$[M-H]^-$ 785.5046$[M+H]^+$	783.4902$[M-H]^-$ 785.5057$[M+H]^+$	0.26 1.40	621.4403 $[M-H-(Glc-H_2O)]^-$, 161.0441$[(Glc-H_2O)-H]^-$, 101.0252$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 623.4523$[M+H-(Glc-H_2O)]^+$, 461.3986$[M+H-2(Glc-H_2O)]^+$, 443.3875$[PPD+H-H_2O]^+$, 425.3775 $[PPD+H-2H_2O]^+$, 407.3673$[PPD+H-3H_2O]^+$, 325.1132$[(Glc-Glc)-2H_2O+H]^+$
26	Notoginsenoside T_5	61.5	751.4638$[M-H]^-$	751.4641$[M-H]^-$	0.40	619.4225$[M-H-(Xyl-H_2O)]^-$, 457.3687$[M-H-(Xyl-H_2O)-(Glc-H_2O)]^-$, 161.0452$[(Glc-H_2O)-H]^-$, 101.0247$[(Glc- H_2O)-H-C_2H_4O_2]^-$
27	Notoginsenoside U	62.2	799.4849 $[M-H]^-$ 801.4995$[M+H]^+$	799.4820$[M-H]^-$ 801.4994$[M+H]^+$	3.63 0.12	477.3940$[M+H-2(Glc-H_2O)]^+$, 459.3833$[PPT-H_2O+H]^+$, 441.3730$[PPT-2H_2O+H]^+$
28	Notoginsenoside T_5 isomer	62.4	751.4638$[M-H]^-$	751.4642$[M-H]^-$	0.53	619.4218$[M-H-(Xyl-H_2O)]^-$, 161.0463$[(Glc-H_2O)-H]^-$ 441.3728$[PPT-2H_2O+H]^+$, 423.3624$[PPT-3H_2O+H]^+$, 405.3514$[PPT-4H_2O+H]^+$
29	Ginsenoside F_4	62.8	765.4795$[M-H]^-$ 789.4759 $[M+Na]^+$	765.4797$[M-H]^-$ 789.4765$[M+Na]^+$	0.26 0.86	619.4268$[M-H-(Rha-H_2O)]^-$, 161.0460$[(Glc-H_2O)-H]^-$, 101.0246$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 621.4333$[M+H-(Rha-H_2O)]^-$, 441.3728 $[PPT-2H_2O+H]^+$,

						423.3624[PPT-3H_2O+H]$^+$, 405.3514[PPT-4H_2O+H]$^+$
30	Ginsenoside RK_3	63.4	665.4270[M+HCOO]$^-$	665.4280[M+HCOO]$^-$	1.50	619.4590[M-H]$^-$, 457.3739[M-H-(Glc-H_2O)]$^-$, 457.3739[PPT-H_2O]$^-$, 161.0452[(Glc-H_2O)-H]$^-$, 101.0246[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$
			621.4361[M+H]$^+$	621.4368[M+H]$^+$	1.13	603.4264[M+H-H_2O]$^+$,441.3731[M+H-H_2O-(Glc-H_2O)]$^+$, 441.3731[PPT-2H_2O+H]$^+$, 423.3629[PPT-3H_2O+H]$^+$, 405.3521[PPT-4H_2O+H]$^+$
31	Ginsenoside Rh_4	64.4	665.4270[M+HCOO]$^-$	665.4280[M+HCOO]$^-$	1.50	619.4200[M-H]$^-$, 457.3702[M-H-(Glc-H_2O)]$^-$, 457.3702[PPT-H_2O]$^-$, 161.0452[Glc-H_2O-H]$^-$, 101.0248[Glc-H_2O-H-$C_2H_4O_2$]$^-$
			621.4361[M+H]$^+$	621.4365[M+H]$^+$	0.64	603.4263[M+H-H_2O]$^+$, 441.3732[M+H-H_2O-(Glc-H_2O)]$^+$, 441.3732[PPT-2H_2O+H]$^+$, 423.3629[PPT-3H_2O+H]$^+$, 405.3520[PPT-4H_2O+H]$^+$
32	20(S)-Ginsenoside Rg_3*	65.4	783.4900[M-H]-	783.4905[M-H]-	0.64	621.4385[M-H-(Glc-H2O)]$^-$, 459.3867[M-H-2(Glc-H2O)]$^-$, 323.1867[(Glc-Glc)-2H_2O-H]$^-$, 161.0458[(Glc-H_2O)-H]$^-$, 101.0248[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$
33	20(R)-Ginsenoside Rg_3*	65.9	783.4900[M-H]-	783.4904[M-H]-	0.51	621.4392[M-H-(Glc-H_2O)]-, 459.3870[M-H-2(Glc-H_2O)]-, 161.0449[(Glc-H_2O)-H]$^-$, 101.0248[(Glc-H_2O)-H-$C_2H_4O_2$]$^-$
34	Unknown 1#	68.1	975.7623[M+Na]$^+$	975.7642[M+Na]$^+$	1.95	953.7817 [M+H]$^+$, 499.3755[M+H-PPT]$^+$, 477.394[PPT+H-H_2O]$^+$, 459.3840[PPT+H-H_2O]$^+$, 441.3736[PPT+H-2H_2O]$^+$, 423.3629[PPT+H-3H_2O]$^+$, 405.3523[PPT+H-4H_2O]$^+$, 147.1167[(Rha-H_2O)+H]$^+$
						805.9862[M-H-(Rha-H_2O)]$^-$, 475.3794[PPT-H]$^-$
35	Unknown 2#	68.5	975.7623[M+Na]$^+$	975.7639[M+Na]$^+$	1.64	953.7817 [M+H]$^+$, 499.3764[M+H-PPT]$^+$, 459.3840[PPT+H-H_2O]$^+$, 441.3736[PPT+H-2H_2O]$^+$, 423.3628[PPT+H-3H_2O]$^+$, 405.3520[PPT+H-4H_2O]$^+$, 147.1167[(Rha-H_2O)+H]$^+$
						805.9878[M-H-(Rha-H_2O)]$^-$, 475.3808[PPT-H]$^-$

36	Ginsenoside RK_1	71.4	765.4795$[M-H]^-$	765.4801$[M-H]^-$	0.78	603.4288$[M-H-(Glc-H_2O)]^-$, 161.04534$[(Glc-H_2O)-H]^-$, 101.0246$[(Glc-H_2O)-H-C_2H_4O_2]^-$ 789.4770$[M+Na]^+$, 767.4950$[M+H]^+$, 605.4417$[M+H-(Glc-H_2O)]^+$, 587.4312$[M+H-(Glc-H_2O)-H_2O]^+$, 477.3947$[PPT+H]^+$, 459.3835$[PPT+H-H_2O]^+$, 443.3887$[M+H-2(Glc-H_2O)]^+$, 443.3887$[PPD+H-H_2O]^+$, 425.3784$[PPD+H-2H_2O]^+$, 407.3677$[PPD+H-3H_2O]^+$, 325.1131$[2(Glc-H_2O)+H]^+$
37	Ginsenoside Rg_5	71.9	765.4795$[M-H]^-$	765.4803$[M-H]^-$	1.05	603.4277[M-H-(Glc-H2O)]-, 323.0994[(Glc-Glc)-2H2O+H]+, 161.0454[(Glc-H2O)-H]-, 101.0245[(Glc-H2O)-H-C2H4O2]- 789.4770[M+Na]+, 767.4950[M+H]+, 605.4416[M+H-(Glc-H2O)]+, 587.4315[M+H-(Glc-H2O)-H2O]+, 477.3947[PPT+H]+, 459.3836[PPT+H-H2O]+, 443.3886[M+H-2(Glc-H2O)]+, 443.3886[PPD+H-H2O]+, 425.3784[PPD+H-2H2O]+, 407.3678[PPD+H-3H2O]+, 325.1131[(Glc-Glc)-2H2O)+H]+
38	20(S)-Ginsenoside Rh_2*	73.2	667.4427$[M-HCOO]^-$	667.4438$[M-HCOO]^-$	1.65	621.4388$[M-H]^-$, 459.3904$[M-H-(Glc-H_2O)]^-$, 161.0458$[(Glc-H_2O)-H]^-$, 101.0241$[(Glc-H_2O)-H-C_2H_4O_2]^-$
39	20(R)-Ginsenoside Rh_2*	73.6	667.4427$[M-HCOO]^-$	667.4439$[M-HCOO]^-$	1.80	621.4385[M-H]-, 459.3869$[M-H-(Glc-H_2O)]^-$, 161.0459$[(Glc-H_2O)-H]^-$, 101.0242$[(Glc-H_2O)-H-C_2H_4O_2]^-$
40	Ginsenoside RK_2	80.0	649.4321$[M-HCOO]^-$	649.4331$[M-HCOO]^-$	1.54	603.4266$[M-H]^-$, 161.0457$[(Glc-H_2O)-H]^-$
			605.4412[M+H]+	605.4392$[M+H]^+$	3.30	443.3892$[M+H-(Glc-H_2O)]^+$, 443.3892$[PPD+H-H_2O]^+$, 425.3787$[PPD+H-2H_2O]^+$, 407.3682$[PPD+H-3H_2O]^+$
41	Ginsenoside Rh_3	80.6	649.4321$[M-HCOO]^-$	649.4329$[M-HCOO]^-$	1.23	603.4292$[M-H]^-$, 161.0457$[(Glc-H_2O)-H]^-$, 101.0239$[(Glc-H_2O)-H-C_2H_4O_2]^-$
			605.4412$[M+H]^+$	605.4412$[M+H]^+$	0.00	443.3888$[M+H-(Glc-H_2O)]^+$, 443.3888$[PPD+H-H_2O]^+$, 425.3785$[PPD+H-2H_2O]^+$, 407.3681$[PPD+H-3H_2O]^+$

The peaks at retention times of 68.1min and 68.5min are a pair of PPT type saponin isomers, with a –Rha glycosyl substitute in their structures.

During the processing of notoginseng, the two most common routes to produce secondary saponins were (1) deglycosylation and (2) dyhydration at C-20 of their aglycons. For example, the deglycosylation of one -Glc at C-20 of ginsenoside Rd forms Rg_3, and the loss of one -Glc at C-3 of Rg_3 produces Rh_2. Moreover, notoginsenoside T_5, ginsenoside F_4, Rk_3/Rh_4, RK_1/Rg_5, and Rk_2/Rh_3 are the C-20 dehydrated products of notoginsenoside R_2, 20(S)-ginsenoside Rg_2, 20(S)/(R)-Rh_1, 20(S)/(R)-Rg_3 and 20(S)/(R)-Rh_2, respectively.

3.5. Impact of structural features on CAD responses of PPD and PPT ginsenosides

Since the intercept value in linearity test was far smaller than the corresponding slope value for both CAD (<3%) and UV (<0.6%) detector, the slope could be regarded as the RF for each analyte. In virtue to its feature of being a universal detector where the RF value is theoretically independent of the analyte's chemical structure, CAD presented far smaller RF differences than UV detector. However, there was still little variation of the CAD RFs of each saponin in the process of nebulization in CAD detector, leading to a narrow range of CAD RFs. Based on the difference of CAD RFs of authentic reference standards, the rules of CAD response over structural features of PPD and PPT ginsenosides were discovered. Nevertheless, UV RFs were not found to have relevance to saponin structures.

It was found that the optical rotations at C-20 had no influence on CAD responses (Table 2Sa). 20(S)-epimers of Rg_2, Rh_1, Rg_3, Rh_2 and PPD all exhibited very little variation (<2.3%) on CAD RFs compared with their corresponding 20(R)-epimers. Data also showed that CAD RFs of PPD saponins were generally higher (6%-26%) than those of PPT saponins (Table 2Sb). The glycosyl substituent at C-3 position of PPD saponins had little impact on CAD RFs (Table2Sc). For example, 20(S)-PPD, 20(S)-Rh_2 and 20(S)-Rg_3 all bear -H at C-20, and the substituents at C-3 are -H, -Glc, and -Glc-Glc, respectively. The variations of CAD RFs of these above three saponins were below 15%. If these variations are ignored, the CAD RFs of all PPD saponins with the same C-20 substituent while different C-3 substituents could be considered as a constant value.

The glycosyl substituent at C-6 position of PPT saponins had a little influence on CAD RFs (Table 2Sd). For example, when -H is fixed at C-20, the variations of CAD RFs of 20(S)-PPT, Rh1, Rf, and Rg_2, which bears -H, -Glc, -Glc-Glc and -Glc-Rha at C-6, respectively, were within 5%. However, if -Glc is fixed at C-20, cases were complicated. When C-6 substituent changed

from -H to -Glc, CAD RF decreased by 18%. When C-6 substituent changed from -Glc to -Glc-Xyl, i.e., one more five-carbon sugar was added, CAD RF increased by 7%. When C-6 substituent changed from -Glc to -Glc-Rha, i.e., one more six-carbon sugar was added, CAD RF increased by 23%.

For PPT saponins, once the C-6 substituent was fixed, the change of C-20 substituent from -H to -Glc caused less than 25% of the variation of CAD RFs (Table 2Se). Nevertheless, different glycosyl substituents at C-20 caused relatively greater changes speaking of PPD saponins. It indicated that once C-3 substituent is fixed, the adding of one more six-carbon sugar, i.g. -Glc, causd the reduction of CAD from 25% to 49%. Furthermore, the addition of one more five-carbon sugar, i.g., -Xyl and -Ara, led to a 72% increase of CAD RFs (Table 2Sf).

3.6 Prediction of CAD RF values of saponins with known structure

Now that the impact of structural features of PPD and PPT saponins on their CAD RFs has been discovered, the CAD RF values of those saponins without authentic reference standards available could be predicted. The assigned RF values were further employed to determine the overall saponin content in raw and processed notogingseng samples. To make things easier, PPT saponins can be divided into two groups according to their CAD RF values: (1) high polar saponins, including R_1 and Rg_1, with retention times prior to 30 min; (2) medium and low polar saponins, including F_1, Rg_2, Rh_1, Re, Rf and PPT, with retention times from 30 min to 69 min. CAD RFs in the former group were 2.55 and 2.74, with the variation less than 8%. Nevertheless, CAD RFs in the latter group were ranging from 3.02 to 3.29, with the variation within 10%. This classification made the CAD RF assignments easier when PPT saponins were concerned: if the saponin was of high polarity (Rt<30min), CAD RF was assigned to be 2.64, i.e., the average value of 2.55 and 2.74; if the polarity of the saponin was medium or low (Rt>30min), CAD RF was assigned as 3.16, i.e., the average value of 3.02 and 3.29. Since C-20 dehydrated and 25-OH hydrated PPD and PPT saponins all belong to secondary saponins which may only appear in processed notoginsengs with quite low contents, their authentic reference standards were difficult to obtain. The assignment rule of CAD RFs of these PPD/PPT saponin derivatives was regarded as the same as regular PPD/PPT saponins. The prediction of CAD RF values of the saponins without authentic reference standard in raw and processed notoginsengs is presented in Table 3a and 3b.

425 **Table 3a.** The retention times, aglycons, glycosyl substituents and the prediction of CAD RF values of saponins without authentic reference standard in raw
426 notoginsengs

RT(min)	Saponin	Aglycon	C-3 substituent	C-6 substituent	C-20 substituent	Predicted CAD RF	Comments on the prediction of CAD RF
10	20-*O*-glucoginsenoside R_f	PPT	-H	-Glc^{2}-^{1}Glc	-Glc	2.64	High polar PPT type ginsenoside
17	Notoginsenoside R_3	PPT	-H	-Glc	-Glc^{6}-^{1}Glc	2.64	
36.9	Malonyl-ginsenoside Rg_1	PPT	-H	-Glc^{6}-^{1}Malony	-Glc	3.16	Medium/low polar PPT type ginsenoside
38.4	Yesanchinoside D	PPT	-H	- Glc^{6}-Ac	-Glc	3.16	
42.3	Notoginsenoside R_4	PPD	-Glc^{2}-^{1}Glc	/	-Glc^{6}-$^{1}Glc^{6}$-^{1}Xyl	2.07	[a]
44.2	Notoginsenoside Fa	PPD	-Glc^{2}-$^{1}Glc^{2}$-^{1}Xyl	/	-Glc^{6}-^{1}Glc	1.80	[b]
46.1	Notoginsenoside 20(S)-R_2	PPT	-H	-Glc^{6}-^{1}Glc	-H	3.16	Medium/low polar PPT type ginsenoside
58	Gypenoside XVII	PPD	-Glc	/	-Glc^{6}-^{1}Glc	1.98	[c]

427 a. Notoginsenoside R_4 has one more -Xyl at C-6 substituent compared with Rb1. Since the addition of one -Xyl to C-6 in PPD ginsenosides causes 15% increasing of
428 CAD RF, the CAD RF of R_4 was assigned as 2.07 (=1.80×115%).
429 b. Notoginsenoside Fa has one more -Xyl at C-3 substituent compared with Rb1. Sine the glycosyl substituent at C-3 position of PPD type ginsenosides had little impact
430 on CAD RF, the CAD RF of Fa was assigned as that of Rb_1.

c. Gypenoside-XVII has one more –Glc at C-20 substituent compared with F2. Since The change of –Glc to –Glc-Glc at C-20 in PPD ginsenosides causes 50% decreases of CAD RF, the CAD RF of Gypenoside-XVII was assigned as 1.98(=2.97÷150%).

Table 3b. The retention times, aglycons, glycosyl substituents and the prediction of CAD RF values of saponins without authentic reference standard in processed notoginsengs (saponins already listed in Table 3a are not included)

RT(min)	Saponin	Aglycon	C-3 substituent	C-6 substituent	C-20 substituent	Predicted CAD RF	Comments on the prediction of CAD RF
13.3	25-OH-20(S)-Rh_1	25-OH PPT	-H	-Glc	-H	2.64	high polar PPT ginsenoside
17.0	25-OH-20(R)-Rh_1	25-OH PPT	-H	-Glc	-H	2.64	
40.3	20(S)-Rh_1 isomer	PPT	-H	-Mannose	-H	3.16	Medium/low polar PPT ginsenoside
41.4	20(R)-Rh_1 isomer	PPT	-H	-Mannose	-H	3.16	
52.3	25-OH Rg_3	25-OH PPD	-Glc	/	-Glc	3.48	The average RF value of 20(S)-Rg_3 and 20(R)-Rg_3
59.3	Gypenoside LXXV	PPD	-H	/	-Glc^{6}-^{1}Glc	1.94	[a]
61.5	Gypenoside LXXV isomer	PPD	-H	/	-Glc-Glc （linkage not sure）	1.94	
62.3	Notoginsenoside T_5	C-20 dehydrated PPT	-H	-Glc^{6}-^{1}Xyl	-H	3.16	
63.2	Nnotoginsenoside U	PPT	-H	-H	-Glc^{6}-^{1}Glc	3.16	Medium/low polar PPT ginsenoside
63.4	Notoginsenoside T_5 isomer	C-20 dehydrated PPT	-H	-Glc^{6}-^{1}Xyl	-H	3.16	
63.9	Ginsenoside F_4	C-20	-H	-Glc^{6}-^{1}Rha	-H	3.16	

		dehydrated PPT					
64.5	Ginsenoside RK_3	C-20 dehydrated PPT	-H	-Glc	-H	3.16	
65.4	Ginsenoside Rh_4	C-20 dehydrated PPT	-H	-Glc	-H	3.16	
69.1	Unknown 1	PPT	-H	Unknown	Unknown	3.16	
69.5	Unknown 2	PPT	-H	Unknown	Unknown	3.16	
72.4	GinsenosideRK_1	C-20 dehydrated PPD	-Glc^{2}-^{1}Glc	/	-H	3.48	
73.1	Ginsenoside Rg_5	C-20 dehydrated PPD	-Glc^{2}-^{1}Glc	/	-H	3.48	The average RF value of 20(S)-Rg_3 and 20(R)-Rg_3 [b]
81.0	Ginsenoside RK_2	C-20 dehydrated PPD	-Glc	/	-H	3.99	
81.5	Ginsenoside Rh_3	C-20 dehydrated PPD	-Glc	/	-H	3.99	The average RF value of 20(S)-Rh_2 and 20(R)-Rh_2 [c]

a. Gypenoside LXXV has one more -Glc at C-20 substituent compared with CK. Since the change of –Glc to –Glc-Glc to C-20 in PPD ginsenosides causes 50% decreasing of CAD RF, the CAD RF of Gypenoside LXXV was assigned as 1.94 (=2.91÷150%).

b. Ginsenoside RK_1 and Rg_5 are C-20 dehydrated Rg_3

c. Ginsenoside RK_2 and Rh_3 are C-20 dehydrated Rh_2

439 3.7 Validation results of the prediction of CAD RFs

440 The predicted CAD RFs values for notoginsenoside R_1 and ginsenoside Rg_1 are 2.64, since
441 they are both high polar PPT type saponins. And the predicted CAD RFs for the medium or low
442 polar PPT type saponins with authentic reference standard, i.e., ginsenoside Re, Rf,
443 20(S)-/20(R)-Rh_1, 20(S)-/20(R)-Rg_2, F_1, and 20(S)-PPT, are all assigned as 3.16. The differences
444 between the predicted RFs and CAD slopes of linear regression for all the ten PPT saponins are
445 between 0.48% and 4.75%, indicating the accuracy of our prediction. Moreover, the retention time
446 of gypenoside XVII was 58.9 min, which was in accordance with the retention time, i.e., 58.0 min,
447 based on our identification by the combination of LC-QTOFMS and HPLC-CAD analysis (Table
448 2a). The linear regression equation was Y=2.0744X+0.00818, r=0.999 (n=6). The difference
449 between the CAD slope of linear regression and the predicted RF value was 4.77%, suggesting
450 that our prediction was reasonable and accurate. Although it is very difficult to get all the authentic
451 reference standards in raw and processed notoginsengs at this stage, our prediction is a quite easy,
452 accurate and stable method to determine the complex saponin contents in this herb according to
453 our validation results.

454 3.8 Determination of saponins in raw and processed notoginseng

455 Notoginseng saponins are the main component of *P. notoginseng*. Besides, volatile oils,
456 polysaccharides, dencichine and flavonoids are also contained in this herb. Among these species,
457 volatile oils do not have CAD signals because CAD can only detect non-volatile or semi-volatile
458 compounds. Polysaccharides and dencichine are of high polarity, and may have very short
459 retention times or could even hardly be retained on C18 column. Moreover, the contents of
460 flavonoids are quite low in the underground parts of notoginseng. As long as the assigned saponin
461 peaks without authentic reference standards in CAD chromatograms are one-to-one corresponding
462 to the ones identified by LC-QTOFMS, the low content of flavonoids would not interfere with the
463 detection of saponins in this method.

464 The content of saponins in raw notoginseng from different cultivated places, and of various
465 sizes and medicinal parts were calculated. Moreover, saponin content in processed notoginseng
466 with diverse processing procedures were also evaluated and compared. The powder (80 mesh) of
467 corresponding notoginseng sample was employed to prepare sample solutions. For the saponins

with authentic reference standards available in this experiment, a simple internal standard method was performed to calculate the content of saponins. However, if the saponin had no authentic reference standard, its CAD RF was predicted and calculated based on the structure identified using LC-QTOFMS. And then an internal standard method could easily be carried out. To better compare the saponin content in different kinds of notoginsengs, the water content of each batch of notoginseng powder was previously determined by Karl Fisher titration, and the final results were calculated based on water-free basis. The saponins contents in raw notoginsengs are listed in Table 3S. Our results conformed to the data presented in previous literature.[30]

The total amount of all saponins, the total amount of ginsenoside Rg_1, Rb_1 and R_1, and the ratio of PPD *vs.* PPT saponins were compared in raw *P. notoginseng*. Take 120 head raw *P. notoginseng* as an example, the saponins with content greater than 1 mg/g were as follows: R_1, Rg_1, Re, malonyl Rg_1, R_4, Fa, 20(S)-R_2, Rb_1, 20(S)-Rh_1 and Rd, among which R_1, Rg_1 and Rd are considered to be the three most representative saponins, since the total amount of R_1, Rg_1 and Rd is used to evaluate the quality of raw *P. notoginseng* in Chinese Pharmacopoeia (Chp). The total amount of R_1, Rg_1 and Rb_1 accounted for 74% to 81% of the total saponin content no matter of what herbal size, cultivate place or medicinal part, with fairly low variation (RSD 3.0% , n=15). It can be concluded that the total amount of these three saponins can be used to represent the total saponin content, and tedious determination of total amount is unnecessary. In Chp, the total amount of these three saponins should be no less than 5.0% (50 mg/g). Based on our results, all but the saponin contents in the main root of countless head notoginseng met the requirement in Chp. Interestingly, our study exhibited that the saponin contents were not always proportionate to the size of main root. To our surprise, 40-head, not 20 or 30-head, *P. notoginseng* possessed the highest total saponin content no matter where the cultivated place was. In addition, 40-head *P. notoginseng* also exhibited the highest PPD/PPT ratios, which were 0.984 and 0.912 for the notoginseng from Yunnan and Guangxi, respectively. The PPD/PPT ratios in main root of *P. notoginseng* from Yunnan were basically higher than those from Guangxi for the corresponding sizes greater than 120 heads. Literature has mentioned that PPD/PPT ratio could be used as a tool to distinguish the types of ginseng. [31] Thus, we tried to find the relationship between PPD/PPT ratios and total saponin contents. We correlated these two results obtained from different heads of

notoginsengs (Yunnan), a correlation coefficient of 0.614 was calculated from the linear regression (Fig. 1S). Although the linear correlation was not good enough, thereis still some trend that the PPD/PPT ratio has relation to total saponins at least in the case of different size of main root. This result suggested that PPD/PPT ratio could be regarded as a parameter to determine the quality of *P. notoginseng*. Moreover, the total saponin amount in different medicinal parts decreased in the following order: 40 or 60 head main root≈rhizome>branch root>>root hair. What should be mentioned is that the total amount of R_1, Rg_1 and Rb_1 in root hair was 52.5mg/g, which was only 5% above the qualified line of notoginseng in Chp 2010. These results conformed to the description of Sanqi in Chp 2010, in which only the main root, branch root and rhizome are included.

Chan *et. al* firstly introduced the term "biomarker" into steamed notoginseng.[32] Here, the "biomarker" means the compounds only existed in steamed notoginseng, or those of quite high content in steamed notoginseng yet of extremely low content in raw herbs. The concept of biomarker could be successfully utilized to differentiate raw and processed notoginseng. In our study, 25 secondary saponins were found in processed notoginseng, among which Rk_3, Rh_4, 20(S)-/20(R)-Rg_3, RK_1 and Rg_5 were those with the highest amount. In the processed notoginseng which has been steamed for 3hrs, the content of these above saponins were no less than 0.5 mg/kg. Thus, these 6 saponins were designated as biomarkers in processed notoginseng at the analytical level in this experiment. Researches have shown that ginsenoside Rk_3, Rh_4, Rg_3, RK_1 and Rg_5 are proven to be biologically potent in anti-tumor activities and in cardiovascular systems.[33-41] Thus, the function difference of processed notoginseng compared with the raw herbs was basically due to the difference of compound basis, in which the biomarkers may have major contributions to the pharmacological activities of processed notoginseng.

Steaming is a most frequently used processing method for this herb, and steaming at 100 ℃ for 3 hrs has been set as the provincial standard for processed notoginseng powder in Yunnan, China, since Apr. 1, 2013. The contents of the biomarkers and total secondary saponins increased basically with the increasing of steaming time, and the data of the content of all saponins are shown in Table 4Sa, 4Sb, 4Sc, 4Sd and 4Se. However, the increasing rate of the biomarker content in steamed notoginseng from 3hr to 4hr was not obvious, indicating that a 3 to 4-hr steaming time

526 is enough, while a longer steaming time may not always lead to significantly greater contents of
527 secondary saponins. Thus, steaming for 3hrs can be regarded as the beginning of the platform of a
528 relatively constant content of biomarkers. Furthermore, the ratio of secondary *vs.* original saponins
529 of 3hr-steamed notoginsengs was the highest among the steamed samples, proving that 3hr could
530 be regarded as the best steaming time for notoginseng based on our results. Moreover, steaming is
531 also a cost efficient way for processing notoginseng.

532 Except for steaming, frying and stewing of notoginseng are two other traditional processing
533 procedures in Chinese culture. However, our results showed that the contents of biomarkers and
534 total secondary saponins were quite low compared with those in 3hr-steamed notoginseng,
535 indicating that frying or stewing may not be an appropriate way for processing notoginseng. In
536 recent literatures, baking and autoclaving are two techniques to process notoginsengs. [42,43]
537 Apparently, with the increasing of temperature and time, the contents of biomarkers and secondary
538 saponins increased dramatically. Given the same temperature (100 ℃) and processing time (24
539 hr), the contents of biomarkers and secondary saponins in autoclaved notoginseng were of about
540 10 times compared with those in the baked sample, indicating that pressure was an important
541 parameter for the generation of secondary saponins. According to our data, the biomarkers and
542 secondary saponin contents in baked notoginseng under 100 ℃ for as long as 48hrs were just
543 comparable to those in 3hr-steamed sample, suggesting that the humidity in the processing
544 procedure was also essential. Thus, four major parameters, ie., humidity, temperature, time and
545 pressure, should be taken into consideration on the journey of seeking for the best processing
546 procedure for notoginseng. The saponin contents in the processed notoginsengs which has been
547 baked at 120 ℃ for 24 hrs, autoclaved at 100 ℃ for 4 to 6 hrs, and autoclaved at 120 ℃ for 2
548 hrs are comparable to those in 3hr-steamed notoginsengs. Although baking is easy to achieve, a
549 relatively longer processing time leads to a low cost efficiency in this case. In the case of
550 autoclaving, the 6 hr-autoclaving at 120 ℃ and the 18 hr-autoclaving at 100 ℃ led to the
551 increasing of the amount of secondary saponins by 3 to 5 times. However, the advantage of
552 steaming at high pressure at 100 ℃ over ordinary steaming for a relatively short period of
553 processing time, i.e., less than 6 hrs, was not very obvious. The autoclaving at 120 ℃ for a

554 comparatively shorter period could produce a considerable amount of secondary saponins.
555 However, the equipment of autoclave needs special attention for operation, and could not be
556 implemented in household. To sum up, steaming for 3 hrs was confirmed to be an easy and cost
557 efficient method for the processing of notoginseng. Nevertheless, autoclaving for a relatively
558 longer period could be an economic and efficient way to prepare and isolate secondary saponins
559 with potent pharmacological effects which are not existed in raw notoginsengs.

560 **4. Conclusions**

561 In this study, an in-depth analysis of the saponin components and saponin contents in raw and
562 processed *Panax notoginseng* was performed. A gradient eluted HPLC method using acetonitrile
563 and water as mobile phases coupled with charged aerosol detector was established and validated to
564 determine 22 PPD and PPT saponins and aglycons simultaneously in notoginseng. Since the
565 discrepancy of CAD RFs of saponins is quite narrow, the impact of the structural features,
566 including the type of aglycon, the optical rotations at C-20, the glycosyl substituent and the
567 glycosyl linkage of different PPD and PPT saponins on their CAD RFs was discovered. Moreover,
568 the structures of saponins existed in raw and processed notoginseng were extensively identified
569 using LC-QTOFMS. At the analytical level, 16 original and 25 secondary saponins were detected
570 in raw and processed notoginseng, respectively. Since the saponins in raw or in processed
571 notoginsengs were predominately dammarane-type tetracyclic triterpenoid saponins, the impact
572 rules were successfully applied to predict the CAD RFs of saponins, and then a simple internal
573 standard method could be carried out to determine the content of each saponin in notoginsengs. An
574 investigation on saponin contents in raw *P. notoginseng* of different sizes, growing places,
575 medicinal parts, and those in processed notoginseng of different processing procedures, *i.e.*,
576 steaming, baking, autoclaving, stewing and frying was then implemented. Our results indicated
577 that 40 head main root of notoginseng possessed the greatest quantity of total saponins among raw
578 herbs. The total saponins in raw *P. notoginseng* from Yunnan were greater that those from Guangxi
579 for the corresponding size. Moreover, the total saponins in different medicinal parts were
580 decreased in the following order: 40 or 60 head main root≈rhizome>branch root>>root hair.
581 Ginsengnoside Rk_3, Rh_4, 20(S)-/20(R)-Rg_3, RK_1 and Rg_5 were set as biomarkers, as these were the
582 6 most abundant secondary saponins existed in processed notoginseng, and may be responsible for

583 the main difference of pharmacological functions between the raw and processed herbs. As a result,
584 steaming for 3hrs was proven to be the best processing method as it produces fairly high amount
585 of biomarkers and secondary saponins in processed samples for a relatively shorter period and in a
586 cost effective and convenient way. The authentic reference standards of quite a few saponins in
587 notoginsengs could not be obtained commercially, especially those secondary saponins in
588 processed herb. The major advantage of this HPLC-CAD method over previous established
589 QAMS methods is that it is not necessary to get relative correction factors for each saponin
590 experimentally, but the CAD response factor could just be predicted theoretically according to our
591 data. Thus, this article has for the first time provided an easy and reliable method to evaluate the
592 content of those saponins and the quality of raw and processed notoginseng which had never been
593 extensively studied in previous literatures.
594

595 **Acknowledgements**

596 This work was financially supported by the Project of Shanghai Committee of Science and
597 Technology (13401900300) and the Foundation of the Ministry of Education of China
598 (NCET-10-0944).

599 **References**

1 S.Y. Han, H.X. Li, X. Ma, K. Zhang, Z.Z. Ma, Y. Jiang, P.F. Tu, Evaluation of the anti-myocardial ischemia effect of individual and combined extracts of *Panax notoginseng* and *Carthamustinctorius* in rats, *J. Ethnopharmacol.*, 2013, **145**, 722.

2 P.W. Wang, J.G. Cui, X.Y. Du, Q.B. Yang, C.L. Jia, M.Q. Xiong, X.T. Yu, L. Li, W.J. Wang, Y. Chen, T. Zhang, *Panax notoginseng* saponins (PNS) inhibits breast cancer metastasis, *J. Ethnopharmacol.*, 2014, **154**, 663.

3 W. Xia, C.H. Sun, Y. Zhao, L.J. Wu, Hypolipidemic and antioxidant activities of Sanchi (Radix Notoginseng) in rats fed with a high fat diet, *Phytomedicine*, 2011, **18**, 516

4 C.Y. Yang, J. Wang, Y. Zhao, L. Shen, X. Jiang, Z.G. Xie, N. Liang, L. Zhang, Z.H. Chen, Anti-diabetic effects of *Panax notoginseng* saponins and its major anti-hyperglycemic components, *J. Ethnopharmacol.*, 2010, **130**, 231.

5 State Administration of Traditional Chinese Medicine (People's Republic of China). Zhong Hua Ben Cao Jin Xuan Ben. vol. 1. Shanghai: Shanghai Science and Technology Publishers; 1996.

6 D.F. Toh, D.N. Patel, E.C.Y. Chan, A. Teo, S.Y. Neo, H.L. Koh, Anti-proliferative effects of raw

and steamed extracts of *Panax notoginseng* and its ginsenoside constituents on human liver cancer cells, *Chin. Med.*, 2011, **6**, 4.

7 S. Sun, C.Z. Wang, R. Tong, X.L. Li, A. Fishbein, Q. Wang, T.C. He, W. Du, C.S. Yuan, Effects of steaming the root of *Panax notoginseng* on chemical composition and anticancer activities, *Food Chem.*, 2010, **118**, 307

8 S. Sun, L.W. Qi, G.J. Du, S.R. Mehendale, C.Z. Wang, C.S. Yuan. Red notoginseng: Higher ginsenoside content and stronger anticancer potential than Asian and American ginseng, *Food Chem.*, 2011, **125**, 1299.

9 A.J. Lau, D.F. Toh, T.K. Chua, Y.K. Pang, S.O. Woo, H.L. Koh, Antiplatelet and anticoagulant effects of *Panax notoginseng*: Comparison of raw and steamed *Panax notoginseng* with *Panax ginseng* and *Panax quinquefolium*, *J. Ethnopharmacol.*, 2009, **125**, 380.

10 C.Z. Wang, E. McEntee, S. Wicks, J.A. Wu, C.S. Yuan, Phytochemical and analytical studies of *Panax notoginseng* (Burk.), *J. Nat. Med.*, 2006, **60**, 97.

11 C.Q. Wang, X.H. Jia, S. Zhu, K. Komatsu, X. Wang, S.Q. Cai, A systematic study on the influencing parameters and improvement of quantitative analysis of multi-component with single marker method using notoginseng as research subject, *Talanta*, 2015, **134**, 587.

12 C.J.S. Lai, T. Tan, S.L. Zeng, L.R. Xu, L.W. Qi, E.H. Liu, P. Li, An enzymatic protocol for absolute quantification of analogues: application to specific protopanoxadiol-type ginsenosides, *Green Chem.*, 2015, **17**, 2580.

13 C.J.S. Lai, T. Tan, S.L. Zeng, X. Dong, E.H. Liu, P. Li, Relative quantification of multi-components in *Panax* notoginseng (Sanqi) by high-performance liquid chromatography with mass spectrometry using mobile phase compensation, *J. Pharmaceut. Biomed.*, 2015, **102**, 150.

14 R.W. Dixon, D.S. Peterson, Development and testing of a detection method for liquid chromatography based on aerosol charging, *Anal. Chem.*, 2002, **74**, 2930.

15 U. Holzgrabe, C.J. Nap, S. Almeling, Control of impurities in L-aspartic acid and L-alanine by high-performance liquid chromatography coupled with a corona charged aerosol detector, *J. Chromatogra. A,* 2010, **1217**, 294.

16 Z. Long, Z.M Guo, X.D. Liu, Q. Zhang, X.g. Liu, Y. Jin, L.N Liang, H.S. Li, J. Wei, N.P. Wu, A sensitive non-derivatization method for apramycin and impurities analysis using hydrophilic interaction liquid chromatography and charged aerosol detection, *Talanta*, 2016, **146**, 423.

17 I. Marquez-Sillero, S. Cardenas, M. Valcarcel, Determination of water-soluble vitamins in infant milk and dietary supplement using a liquid chromatography on-line coupled to a corona-charged aerosol detector, *J. Chromatogr. A*, 2013, **1313**, 253.

18 M. Plaza, J. Kaiuki, C. Turner, Quantification of individual phenolic compounds' contribution to antioxidant capacity in apple: a novel analytical tool basedon liquid chromatography with diode array, electrochemical, and charged aerosol detection, *J. Agric. Food Chem.*, 2014, **62**,

409.

19 M. Poplawska, A. Blazewicz, K. Bukowinska, Z. Fijalek, Application of high-performance liquid chromatography with charged aerosol detection for universal quantitation of undeclared phosphodiesterase-5 inhibitors in herbal dietary supplements, J. Pharm.Biomed. Anal., 2013, 84, 232.

20 L.M. Nair, J.O. Werling, Aerosol based detectors for the investigation of phos-pholipid hydrolysis in a pharmaceutical suspension formulation, *J. Pharm.Biomed. Anal.*, 2009, **49**, 95.

21 M. Poplawska, A. Blazewicz, P. Zolek, Z. Fijalek, Determination of flibanserinand tadalafil in supplements for women sexual desire enhancement usinghigh-performance liquid chromatography with tandem mass spectrometer, diode array detector and charged aerosol detector, *J. Pharm. Biomed. Anal.*, 2014, **94**, 45.

22 Q. Zhou, M.T. Chen, L.H. Zhu, H.Q. Tang, Determination of perfluorinated carboxylic acids in water using liquid chromatography coupled to a corona-charged aerosol detector, *Talanta*, 2015, **136**, 35.

23 L. Wang, W.S. He, H.X. Yan, Y. Jiang, K.S. Bi, P.F. Tu, Performance evaluation of charged aerosol and evaporative light scattering detection for the determination of ginsenosides by LC, *Chromatographia*, 2009, **70**, 603.

24 C.C. Bai, S.Y. Han, X.Y. Chai, Y. Jiang, P. Li, P.F. Tu, Sensitive determination of saponins in Radix et rhizoma notoginseng by charged aerosol detector coupled with HPLC, *J. Liq. Chromatogr. R. T.*, 2008, **32**, 242.

25 A. Stojanovic, M. Lammerhofer, D. Kogelnig, S. Schiesel, M. Sturm, M. Galanski, R. Krachler, B. Keppler, W. Lindner, Analysis of quaternary ammoniumand phosphonium ionic liquids by reversed-phase high-performance liquidchromatography with charged aerosol detection and unified calibration, *J.Chromatogr. A*, 2008, **1209**, 179.

26 P.H. Gamache, R.S. McCarthy, S.M. Freeto, D.J. Asa, M.J. Woodcock, K. Laws, R.O.Cole, HPLC analysis of non-volatile analytes using charged aerosol detection, *LC-GC Eur*., 2005, **18**, 345.

27 S. Matsuyama, Y. Orihara, S. Kinugasa, H. Ohtani, Effects of densities of brominated flame retardants on the detection response for HPLC analysis with acorona-charged aerosol detector, *Anal. Sci*., 2015, **31**, 61.

28 L.E. Magnusson, D.S. Risley, J.A. Koropchak, Aerosol-based detectors for liquid chromatography, *J.Chromatogr. A*., 2015, **1421**, 68.

29 S. Almeling, D. Ilko, U. Holzgrabe, Charged Aerosol detection in pharmaceutical analysis, *J. Pharm.Biomed. Anal.*, 2012, **69**, 50.

30 S.P. Li, C.F. Qiao, Y.W. Chen, J. Zhao, X.M. Cui, Q.W. Zhang, X.M. Liu, D.J. Hu, A novel strategy with standardized reference extract qualification and single compound quantitative evaluation for quality control of *Panax notoginseng* used as a functional food, *J. Chromatogr. A*, 2013, **1331**, 302.

31 C.Y. Liu, J.G. Song, P.F. Li, H.S. Yu, F.X. Jin, Ginsenoside contents in three different ginseng. *Journal of Dalian Polytechnic University*,2011, **30**, 79-82.

32 E.C.Y. Chan, S.L. Yap, A.J. Lau, Ultra-performance liquid chromatography/time-of-flight mass spectrometry based metabolomics of raw and steamed *Panax notoginseng*, *Rapid Commun. Mass SP*., 2007, **21**, 519.

33 H.H. Kwok, G.L. Guo, J.K. Lau, Y.K. Cheng, J.R. Wang, Z.H. Jiang, M.H. Keung, N.K. Mak, P.Y.K. Yue, R.N.S. Wong, Stereoisomers ginsenosides-20(S)-Rg_3 and -20(R)-Rg_3 differentially induce angiogenesis through peroxisome proliferator-activated receptor-gamma, *Biochem. Pharmacol.*, 2012, **83**, 893.

34 R. Wu, Q. Ru, L. Chen, B. Ma, C. Li, Stereospecificity of ginsenoside Rg3 in the promotion of cellular immunity in hepatoma H22-bearing mice, *J. Food Sci.*, 2014, **79**, H1430.

35 J.G. Lee, Y.Y. Lee, S.Y. Kim, J.S. Pyo, H. S. Yun-Choi, J.H. Park, Platelet antiaggregating activity of ginsenosides isolated from processed ginseng, *Pharmazie*, 2009, **64**, 602.

36 H.K. Ju, J.G. Lee, M.K. Park, S.J. Park, C.H. Lee, J.H. Park, C.H. Lee, J. H. Park, S.W. Kwon, Metabolomic investigation of the anti-platelet aggregation activity of ginsenoside Rk_1 reveals attenuated 12-HETE production, *J. Proteome Res.*, 2012, **11**, 4939.

37 J.S. Kim, E.J. Joo, J. Chun, Y.W. Ha, J.H. Lee, Y. Han, Y.S. Kim, Induction of apoptosis by ginsenoside Rk1 in SK-MEL-2-human melanoma, *Arch. Pharm. Res*., 2012, **35**, 717.

38 Y.S. Maeng, S. Maharjan, J.H. Kim, J.H. Park, Y.Y. Suk, Y.M. Kim, et al. Rk_1, a ginsenoside, is a new blocker of vascular leakage acting through actin structure remodeling, *PLOS ONE*, 2013, **8**, Article number e68659.

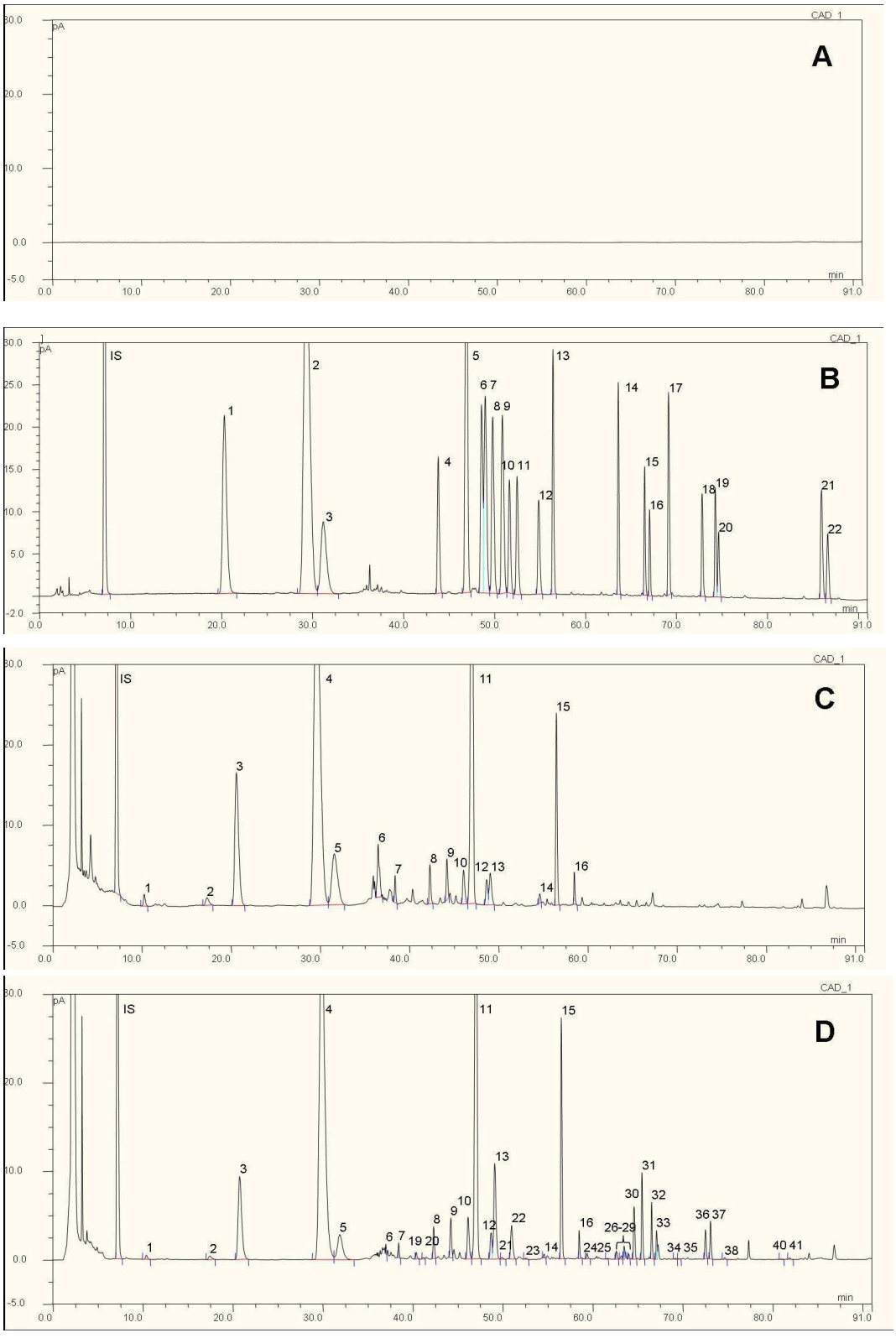
39 J. Sun, G. Sun, X. Meng, H. Wang, M. Wang, M. Qin, B. Ma, Y. Luo, Y. Yu, R. Chen, Q. Ai, X. Sun, Ginsenoside RK_3 prevents hypoxia-reoxygenation induced apoptosis in H9c2 cardiomyocytes via AKT and MAPK pathway, *Evid.-Based Compl. Alt.*, 2013, Article number 690190.

40 N.I. Baek, D.S. Kim, Y. H. Lee, J. D. Park, C.B. Lee, S.I. Kim, Ginsenoside Rh_4, a genuine dammarane glycoside from Korean red ginseng, *Planta Med*., 1996, **62**, 86.

41 S.J. Kim, A.K. Kim, Anti-breast cancer activity of Fine Black ginseng (*Panax ginseng Meyer*) and ginsenoside Rg5, *J. Ginseng Res.*, 2015, **39**, 125e134.

41 D.F. Toh, L.S. New, H.L. Koh, E.C.Y. Chan, Ultra-high performance liquid chromatography/time-of-flight mass spectrometry (UHPLC/TOFMS) for time-dependent profiling of raw and steamed *Panax notoginseng, J. Pharm.Biomed. Anal.*, 2010, **52**, 43.

43 D.Wang, P.Y. Liao, H.T. Zhu, K.K. Chen, M. Xu, Y.J. Zhang, C.R. Yang, The processing of *Panax notoginsen*g and the transformation of its saponin components, *Food Chem*., 2012, **132**, 1808.



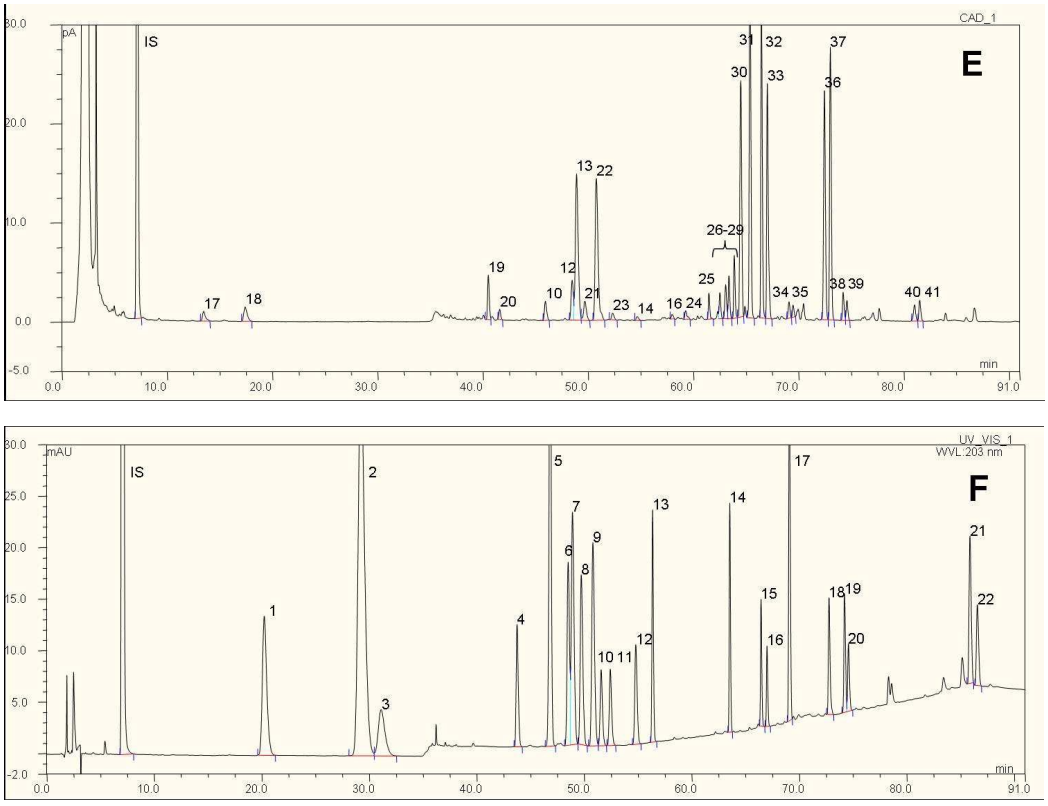
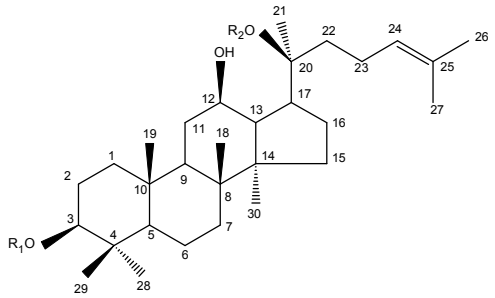
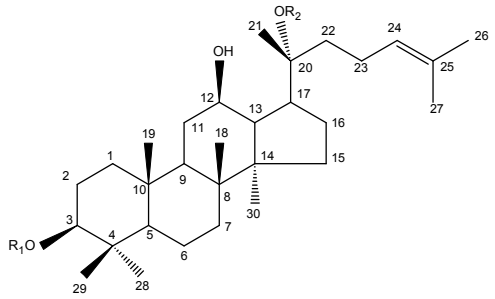


Fig. 1. Typical HPLC-CAD chromatograms of blank (A), reference standard solution (B), sample solutions of raw notoginseng (40 head, Yunnan) (C), steamed notoginseng (100℃, 3hrs) (D), and autoclaved notoginseng (120℃, 18hrs) (E), and typical HPLC-UV chromatogram of reference standard solution (F). The peak numbers denoted in reference standard solution (B) and (F) are: 1, R_1; 2, Rg_1; 3, Re; 4, Rf; 5, Rb_1; 6, 20(S)-Rg_2; 7, 20(S)-Rh_1; 8, 20(R)-Rg_2; 9, 20(R)-Rh_1; 10, Rb_2; 11, Rb_3; 12, F_1; 13, Rd; 14, F_2; 15, 20(S)-Rg_3; 16, 20(R)-Rg_3; 17, 20(S)-PPT; 18, Compound K; 19, 20(S)-Rh_2; 20, 20(R)-Rh_2; 21, 20(S)-PPD; 22, 20(R)-PPD. The peak numbers denoted in sample solution (C), (D) and (E) are: 1, 20-*O*-glucoginsenoside R_f; 2, R_3; 3, R_1; 4, Rg_1; 5, Re; 6, malonyl-ginsenoside Rg_1; 7, yesanchinoside D; 8, R_4; 9, Fa; 10, 20(S)-R_2; 11, Rb_1; 12, 20(S)-Rg_2; 13, 20(S)-Rh_1; 14, F_1; 15, Rd; 16, gypenoside XVII; 17, 20(S)-25-OH Rh_1; 18, 20(R)-25-OH Rh_1; 19, 20(S)-Rh_1 (Man as glycosyl substituent); 20, 20(R)-Rh_1 (Man as glycosyl substituent); 21, 20(R)-Rg_2; 22, 20(R)-Rh_1; 23, 25-OH Rg_3; 24, gypenoside LXXV; 25, gypenoside LXXV isomer; 26, T_5; 27, U; 28, T_5 isomer; 29, F_4; 30, RK_3; 31, Rh_4; 32, 20(S)-Rg_3; 33, 20(R)-Rg_3; 34, Unknown 1[#]; 35, Unknown 2[#]; 36, RK_1; 37, Rg_5; 38, 20(S)-Rh_2; 39, 20(R)-Rh_2; 40, RK_2; 41, Rh_3. The peak numbers denoted in sample solutions are the same as in Table 2a and 2b.

Fig. 2. Saponins observed by LC-ESI-QTOFMS in raw and processed *P. notoginseng*. Unless being specified, the default chirality of the saponins is 20(S) form. The saponins in bold are characteristic for processed samples. Glc, *β*-D-glucopyranosyl; Rha, rhamnose; Ara(p), *α*-L-Arabinose in pyranose form; Ara(f), *α*-L-Arabinose in furanose form; Xyl, xylose; Man, mannose.

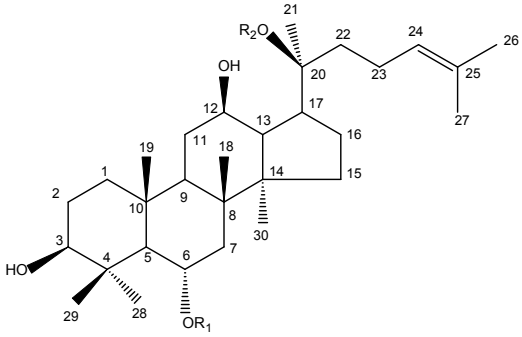


20(S)-protopanaxadiol (20(S)-PPD, R_1=R_2=H)

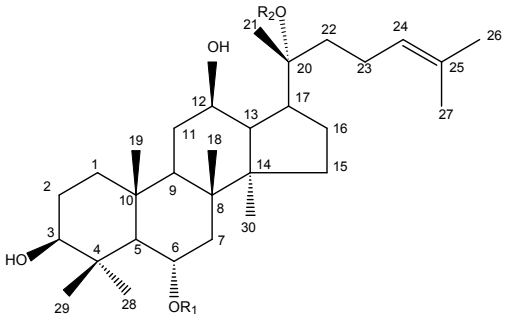


20(R)-protopanaxadiol (20(R)-PPD, R_1=R_2=H)

Compound (PPD type)	R_1	R_2	Molecular formula
Notoginsenoside R_4	-Glc-Glc (2-1)	-Glc-Glc-Xyl (6-1；6-1)	$C_{59}H_{100}O_{27}$
Notoginsenoside Fa	-Glc-Glc-Xyl (2-1；2-1)	-Glc-Glc (6-1)	$C_{59}H_{100}O_{27}$
Ginsenoside Rb_1	-Glc-Glc (2-1)	-Glc-Glc (6-1)	$C_{54}H_{92}O_{23}$
Ginsenoside Rd	-Glc-Glc (2-1)	-Glc	$C_{48}H_{82}O_{18}$
Gypenoside XVII	-Glc	-Glc-Glc (6-1)	$C_{48}H_{82}O_{18}$
Gypenoside LXXV	**-H**	**-Glc-Glc (6-1)**	$\mathbf{C_{42}H_{72}O_{13}}$
20(S)/(R)-Ginsenoside Rg_3	**-Glc-Glc (2-1)**	**-H**	$\mathbf{C_{42}H_{72}O_{13}}$
20(S)/(R)-Ginsenoside Rh_2	**-Glc**	**-H**	$\mathbf{C_{36}H_{62}O_8}$

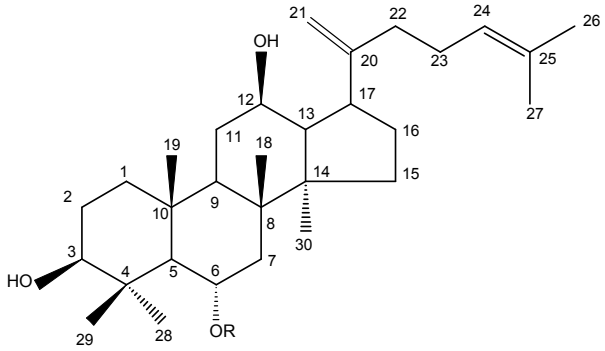


20(S)-protopanaxatriol (20(S)-PPT, R_1=R_2=H)

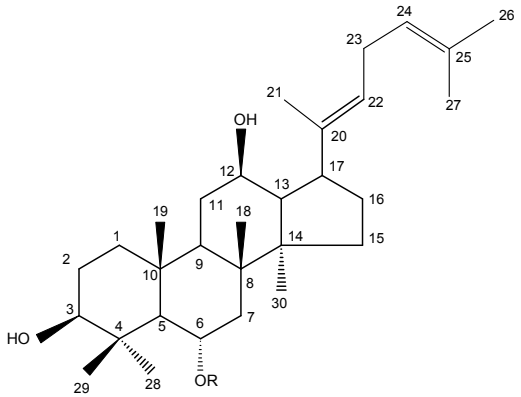


20(R)-protopanaxatriol (20(R)-PPT, R_1=R_2=H)

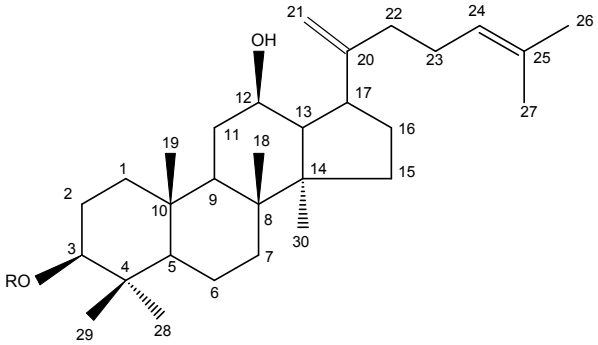
Compound (PPT type)	R_1	R_2	Molecular formula
20-*O*-Glucoginsenoside R_f	-Glc-Glc (2-1)	-Glc	$C_{48}H_{82}O_{19}$
Notoginsenoside R_3	-Glc	-Glc-Glc (6-1)	$C_{48}H_{82}O_{19}$
Ginsenoside R_1	-Glc-Xyl (2-1)	-Glc	$C_{47}H_{80}O_{18}$
Ginsenoside Rg_1	-Glc	-Glc	$C_{42}H_{72}O_{14}$
Ginsenoside Re	-Glc-Rha (2-1)	-Glc	$C_{48}H_{82}O_{18}$
Malonyl-ginsenoside Rg_1	-Glc^6-Malonyl-Glc	-Glc	$C_{45}H_{74}O_{17}$
Yesanchinoside D	-Glc-$COCH_3$	-Glc	$C_{45}H_{74}O_{15}$
20(S)-Notoginsenoside R_2	-Glc-Xyl (6-1)	-H	$C_{41}H_{70}O_{13}$
20(S)-Ginsenoside Rg_2	-Glc-Rha (2-)	-H	$C_{42}H_{72}O_{13}$
20(S)-Ginsenoside Rh_1	-Glc	-H	$C_{36}H_{62}O_9$
Ginsenoside F_1	-H	-Glc	$C_{36}H_{62}O_9$
20(S)/(R)-Ginsenoside Rh_1 (Man as glycosyl substituents)	**-Man**	**-H**	**$C_{36}H_{62}O_9$**
20(R)- Ginsenoside Rg2	**-Glc-Rha (2-1)**	**-H**	**$C_{42}H_{72}O_{13}$**
20(R)- Ginsenoside Rh1	**-Glc**	**-H**	**$C_{36}H_{62}O_9$**
Notoginsenoside U	**-H**	**-Glc-Glc (6-1)**	**$C_{42}H_{72}O_{14}$**



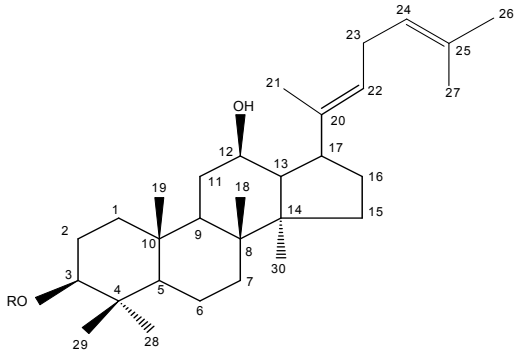
Compound (cis C-20 dehydrated PPT type)	R	Molecular formula
Notoginsenoside T_5	**-Glc-Xyl（6-1）**	**$C_{41}H_{68}O_{12}$**
Ginsenoside F_4	**-Glc-Rha（6-1）**	**$C_{42}H_{70}O_{12}$**
Ginsenoside RK_3	**-Glc**	**$C_{36}H_{60}O_8$**



Compound (trans C-20 dehydrated PPT type)	R	Molecular formula
notoginsenoside T_5 isomer	**-Glc-Xyl（6-1）**	**$C_{41}H_{68}O_{12}$**
Ginsenoside Rh_4	**-Glc**	**$C_{36}H_{60}O_8$**

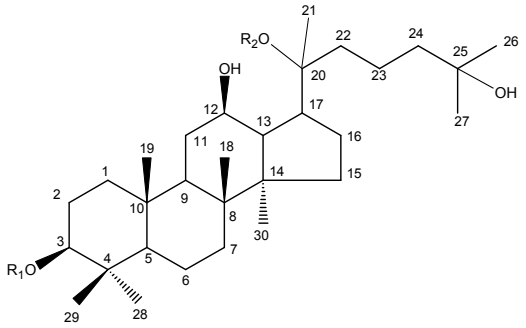


Compound (cis C-20 dehydrated PPD type)	R	Molecular formula
Ginsenoside RK_1	**-Glc-Glc（2-1）**	**$C_{42}H_{70}O_{12}$**
Ginsenoside RK_2	**-Glc**	**$C_{36}H_{60}O_7$**



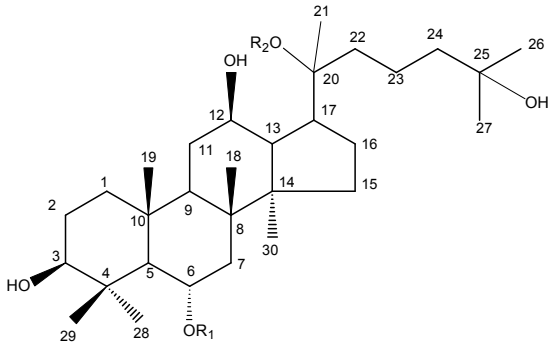
Compound	R	Molecular formula

(trans C-20 dehydrated PPD type)		
Ginsenoside Rg_5	**-Glc-Glc（2-1）**	**$C_{42}H_{70}O_{12}$**
Ginsenoside Rh_3	**-Glc**	**$C_{36}H_{60}O_7$**



20(S)/(R)-dammarane-3β, 12β, 20, 25-tetrol (25-OH PPD, R_1=R_2=H)

Compound(25-OH PPD)	R_1	R_2	Molecular formula
25-OH Ginsenoside Rg_3	**-Glc-Glc**	**-H**	**$C_{42}H_{74}O_{14}$**



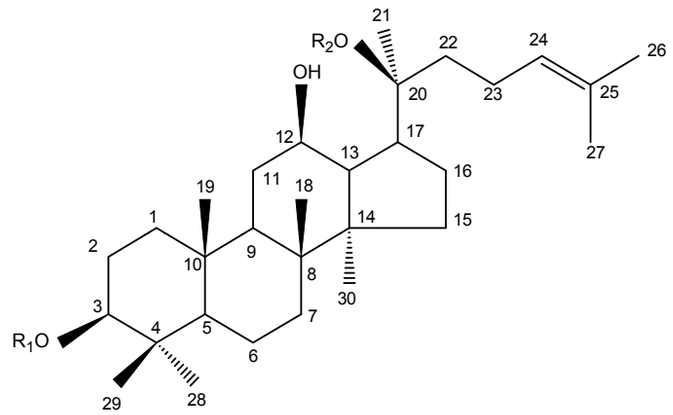
20(S)/(R)-dammarane-3β, 6α, 12β, 20, 25-pentol (25-OH PPT, R_1=R_2=H)

Compound(25-OH PPT)	R_1	R_2	Molecular formula
25-OH Ginsenoside Rh_1	**-Glc**	**-H**	**$C_{36}H_{64}O_{10}$**

Structure-based impacts of saponins on CAD response factors are discovered, and in-depth analysis of saponins in Panax notoginseng is implemented.

Panax notoginseng; Sanqi; Sanchi

Main components



Parent structure of protopanaxadiol (left) and protopanaxatriol (right) saponins