Waters
THE SCIENCE OF WHAT'S POSSIBLE.®

Using Natural Products Application Solution with UNIFI to Identify Chemical Ingredients and Deduce Possible Herbal Composition from Unknown Traditional Medicine Tablets

Lirui Qiao,[1] Rob Lewis,[2] Alex Hooper,[2] James Morphet,[2] Xiaojie Tan,[1] Kate Yu[3]
[1]Waters Corporation (Shanghai), Shanghai, China
[2]Waters MS Technology Center, Manchester, UK
[3]Waters Corporation, Milford, MA, USA

APPLICATION BENEFITS

UltraPerformance LC® (UPLC®) is combined with orthogonal quadrupole time-of-flight mass spectrometry within the Natural Products Application Solution with UNIFI to identify the unknown chemical ingredients and deduce the potential herbal composition of Traditional Medicine tablets. This comprehensive workflow enables researchers to determine the chemical and herbal composition of completely unknown samples, while significantly enhancing accuracy and efficiency.

WATERS SOLUTIONS

Natural Products Application Solution with UNIFI®

ACQUITY UPLC® I-Class System

Xevo® G2-S QTof Mass Spectrometer

UNIFI Scientific Information System

Waters Analytical Standards and Reagents

KEY WORDS

Traditional Medicine Library, ingredient analysis of unknown natural products, UPLC/QTof MS^E, identification of herbal materials, DanShen, *(Salvia miltiorrhiza)*, SanQi, *(Panax Notoginseng)*

INTRODUCTION

Traditional Medicines are known for being comprised of extremely complex elements that can include a variety of plants, extracts, minerals, and animal parts. The critical foundation for their effectiveness originates from the chemical ingredients of their raw herbal materials. In a related application note,[1] we have described how to efficiently identify chemical ingredients from samples with known plants by utilizing the Natural Products Application Solution with UNIFI and its Traditional Medicine Library. For this type of analysis, a researcher only needs to import the compounds associated with the known plants from the Traditional Medicine Library, and use them as search targets to compare with the acquired data. The result is a list of identified components that can be used for further verification. The workflow has a straightforward strategy and its analytical procedure contains simple steps.

However, the reality is that researchers often need to identify chemical ingredients and deduce possible herbal composition for a completely unknown Traditional Medicine product. This type of work is extremely difficult; many times one may not even know where to start. Because the available sample background information is close to none, even with the large amount of data that can be generated using popular approaches such as LC/MS, researchers are still challenged to narrow down their scope and to obtain meaningful information quickly.

The classic workflow for profiling the components of unknown Traditional Medicine products is this: to manually extract each individual chromatographic peak, propose possible molecular formula based on the exact mass of intact protonated or deprotonated ions, and that result is used to search online libraries to obtain potential hits. Afterwards, fragmentation pathways are deduced based on MS/MS fragment ions so that the proposed chemical structure of a target component is confirmed. This process not only requires manual intervention by researchers in almost every single step throughout the entire process, but also has very high demands for expertise levels (both in natural products and in chemistry), as the researcher must be able to find answers in oceans of information.

EXPERIMENTAL

Sample preparation

Two tablets of a TCM product were used for the analysis. After removing the coatings, they were ground into powder. 500 mg of the powder was dissolved in 50 mL MeOH/H_2O (3:1) by ultrasonic the solution for 5 minutes. The final solution was filtered through a 45 µm membrane prior to injection.

LC conditions

LC system:	ACQUITY UPLC I-Class with FTN Sample Manager
Column:	ACQUITY UPLC HSS T3 2.1 x 100 mm, 1.8 µm
Column temp.:	40 °C
Sample temp.:	15 °C
Mobile phase:	A: water (0.1% formic acid); B: acetonitrile

Gradient:

Time	Flow rate (mL/min)	Solvent A (%)	Solvent B (%)	Curves
0	0.6	90	10	Starting
1	0.6	90	10	6
12	0.6	5	95	6
14	0.6	0	100	1
17	0.6	90	10	1

MS conditions

MS system:	Xevo G2-S QTof
Acquisition range:	100-1500 Da
Scan time:	0.1 s
Acquisition mode:	MS^E, ESI- and ESI+ in resolution mode
Lock mass:	Leucine Enkephalin (LE) 1 ppm (scan for 0.3 s, interval: 15 s)
Capillary voltage:	3 kV (ESI+)/2.5 kV (ESI-)
Cone voltage:	100 V
Collision energy (eV):	low CE: 6/High CE: 20-50
Source temp.:	120 °C
Desolvation temp.:	500 °C
Cone gas flow:	30 L/h
Desolvation gas flow:	1000 L/h
Acquisition time:	17 min

Data acquisition, processing, and reporting

UNIFI Scientific Information System with Traditional Medicine Library

The Natural Products Application Solution with UNIFI provides a completely new and comprehensive strategy for solving such a problem. It utilizes the ACQUITY UPLC I-Class System and Xevo G2-S QTof MS to acquire data-independent MS^E data. These data are then searched against the integrated Traditional Medicine Library. The structures of the matched components are verified by MassFragment™ using their corresponding fragment ions. Finally, detailed information of the identified components are displayed automatically in UNIFI using preset workflow templates.

This application note describes how to use the Natural Products Application Solution with UNIFI to identify chemical ingredients and deduce possible herbal content from unknown samples using a Traditional Chinese Medicine (TCM) tablet product as an application example.

RESULTS AND DISCUSSION

UPLC and QTof MS were used for the ingredient separation and MS data acquisition of the unknown TCM tablet sample. The Natural Products Application Solution with UNIFI along with the Traditional Medicine Library was used for the data processing, which resulted in 288 components identified by having a match from the library. Among them, 37 high-level ingredients were initially verified and labeled as "confirmed" based on fragment analysis by MassFragment.

By associating the confirmed components with potential plants, it was deduced that the tablets may contain DanShen (*Salvia miltiorrhiza*) and SanQi (*Panax notoginseng*). By searching the Internet to find known TCM recipes that contain these two herbs, the chemical ingredients (listed from the Traditional Medicine Library) of related herbs from matched recipes can then be used to compare with components found from experiment data. As a result, for this example, 59 major chemical ingredients from the tablets were verified, all from DanShen and SanQi. Hence, the final conclusion was that herbal composition of this TCM product is DanShen and SanQi, which leads us to believe that this product was possibly to be the Sanqi Danshen Tablet or the Compound DanShen Tablet.

The workflow of chemical ingredient analysis with known plants using the Natural Products Application Solution with UNIFI has been described previously in detail.[1] For samples that are complete unknown, additional steps would be deducing possible herbal identities, searching online for potential known TCM recipes that contain these herbs, and, from the UNIFI Traditional Medicine Scientific Library, re-importing corresponding compounds related to potential herbs listed in the matched recipes to verify the existence of these herbs. Figure 1 shows the complete workflow of chemical and herbal ingredient identification for unknown samples.

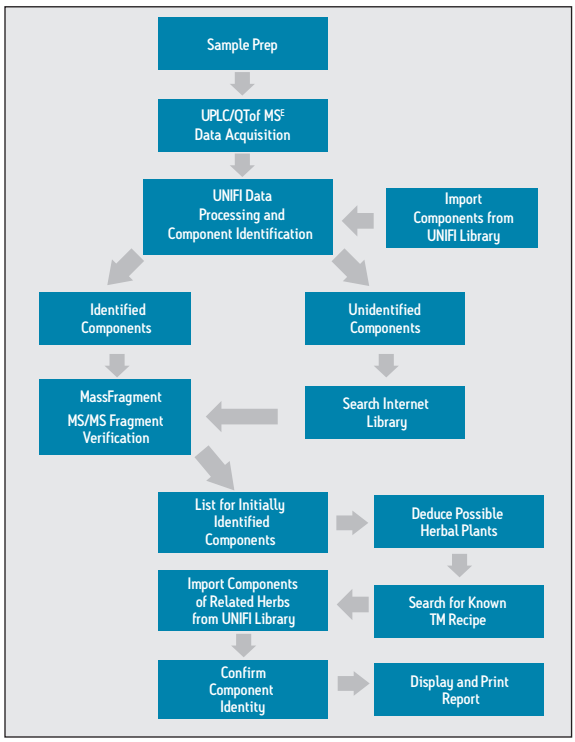

Figure 1. Complete workflow for identifying chemical and herbal ingredient for unknown samples.

Figure 2A shows the UPLC/QTof MS base peak ion (BPI) chromatogram for the unknown TCM tablet. With the UNIFI Scientific Information System, the same results can also be displayed in a 3D format, shown in Figure 2B. Compared with 2D plot (Figure 2A), the information displayed from the 3D plot is closer to the true representative of the components within the sample. It provides a directive visual profile that is much more intuitive for observing the entire chemical component distribution of the sample. For example, from Figure 2B, one can quickly conclude that the range of molecular weights of the chemical ingredients from this sample is mainly between 400 and 1000 Da. In addition, it also allows chemists to have a quick observation on the compounds' coeluting status within the entire run.

The plot displayed in Figure 2B is generated from Apex 3D image scan mode, which is unique to UNIFI. Apart from providing a direct visual effect, it helps to enhance the accuracy of the qualitative and quantitative work for future steps, and it provides major advantages in identifying and eliminating background peaks.

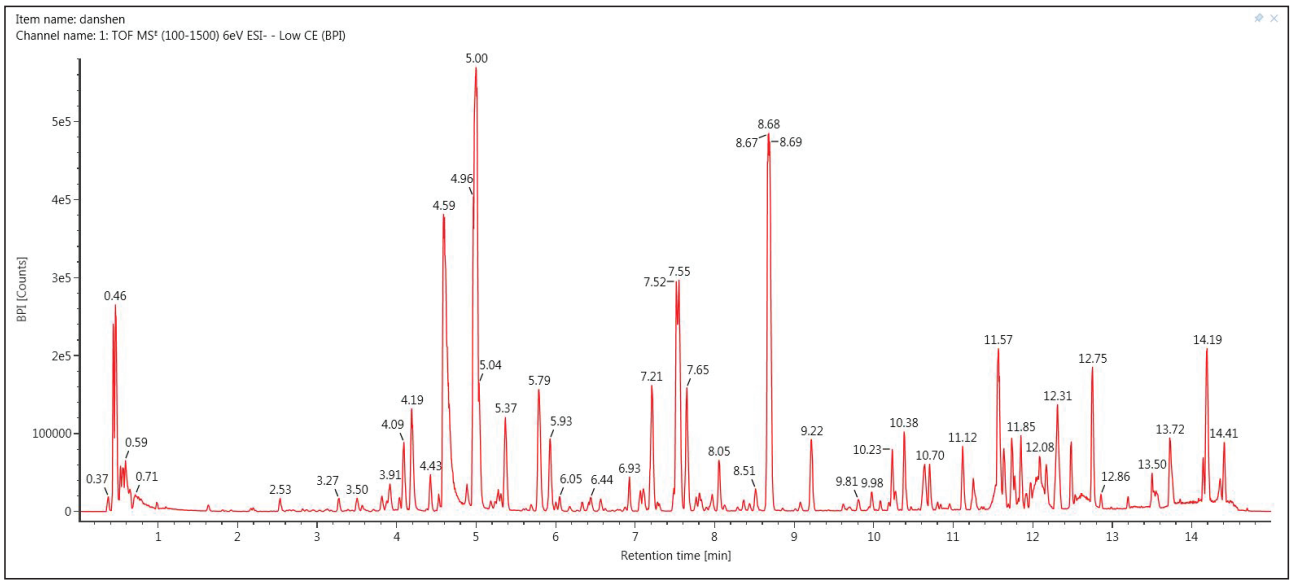

Figure 2A. UPLC/QTof MS base peak ion (BPI) chromatogram of the unknown TCM tablet.

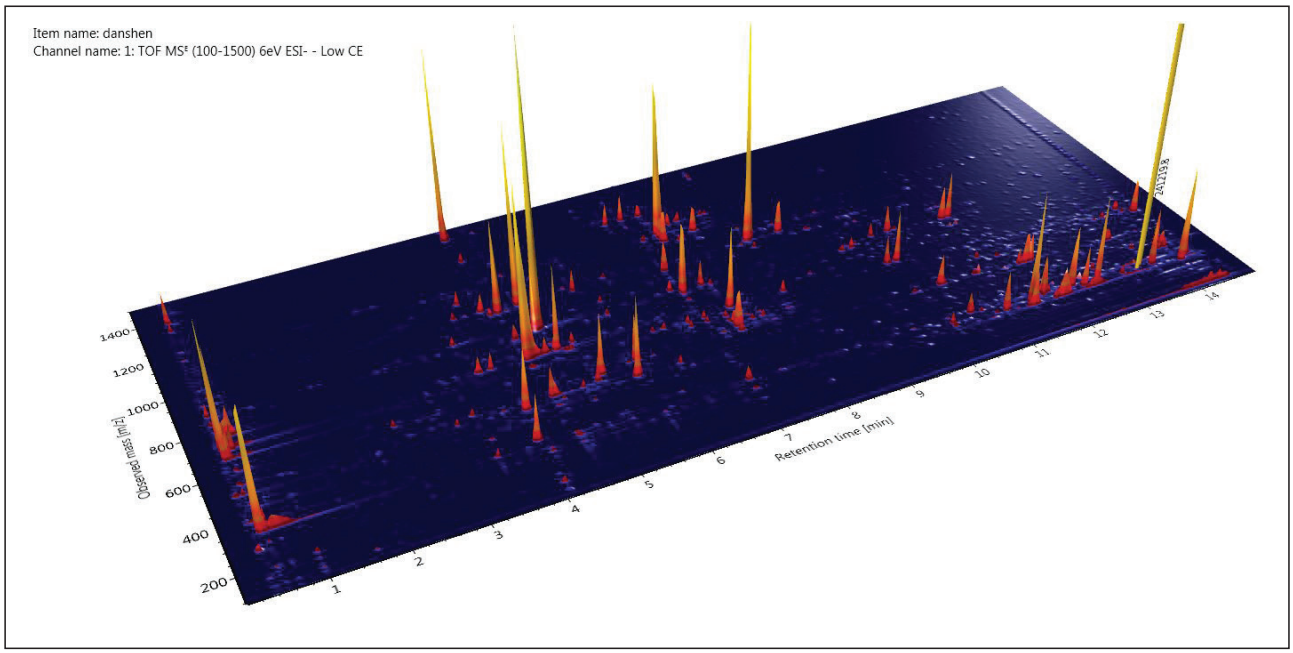

Figure 2B. 3D LC/MS plot of the unknown TCM tablet.

With the Natural Products Application Solution with UNIFI:

1. All steps are completed in automated fashion without the need for operator's intervention. These steps include chromatographic peak extraction, elemental composition determination, Traditional Medicine Library searching, fragment ion structural elucidation, and component identification.
2. Researchers only need to verify whether the fragment ion structural elucidation that was automatically provided by MassFragment is reasonable or not.
3. If a false positive is suspected, or any component that is not matched from the Traditional Medicine Library, the researcher can then initiate a manual process for further identification.

Compared with conventional research protocols, the Natural Products Application Solution with UNIFI converts a manual process of seeking meaningful targets from oceans of information into an automated workflow. This significantly reduces the blindness of the work and enhances productivity. Meanwhile, the demands for the researcher's expertise level is greatly reduced as well.

Figure 3 shows the UNIFI's results for the chemical ingredients identification of the unknown TCM tablet after data processing. The ingredient table shown in Figure 3B lists the components initially identified from the library match. It is possible to have multiple isomers corresponding to the same chromatographic peak at the same retention time. This is when researchers need to verify whether a match is reasonable by looking at the adduct ions as well as the structural elucidation of the fragmentation ions (blue icon).

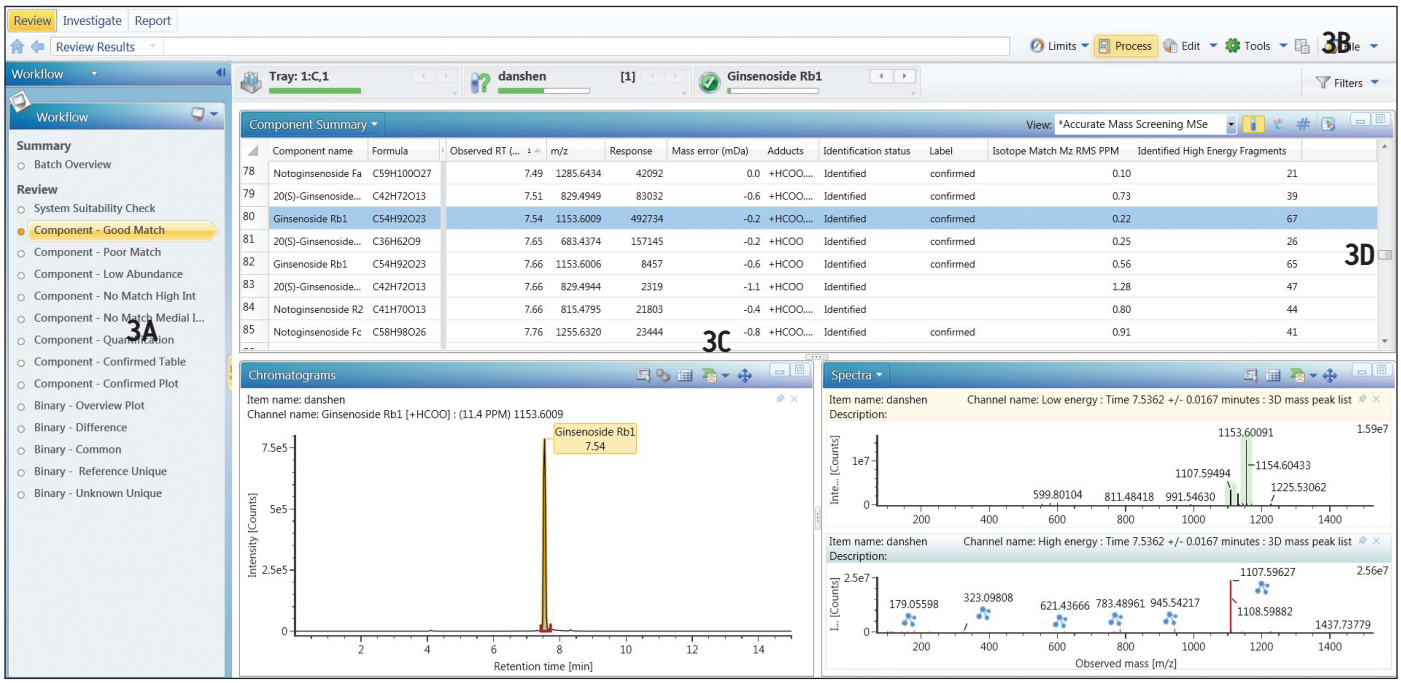

Figure 3. Chemical ingredient identification results in UNIFI for the unknown TCM tablet. 3A shows the template workflows; 3B is the component identification list; 3C is the selected ion chromatogram of single component corresponding to 3B; and 3D is the respective mass spectrum of 3C.

For example, the chromatographic peak at 7.54 minutes is automatically identified by UNIFI as ginsenoside Rb1 or Yesanchinoside E. By clicking the window represented by Figure 3D, an enlarged figure is obtained (Figure 4). Since all fragment ions have been automatically elucidated by the MassFragment, researchers can easily verify whether the fragmentation pathway is reasonable or not. In this example, the compound's cleavage started from the glycosidic bond, and ended at the formation of protopanaxadiol aglycone fragments, indicating the reasonable structure should be the ginsenoside Rb1, which was then labeled as confirmed.

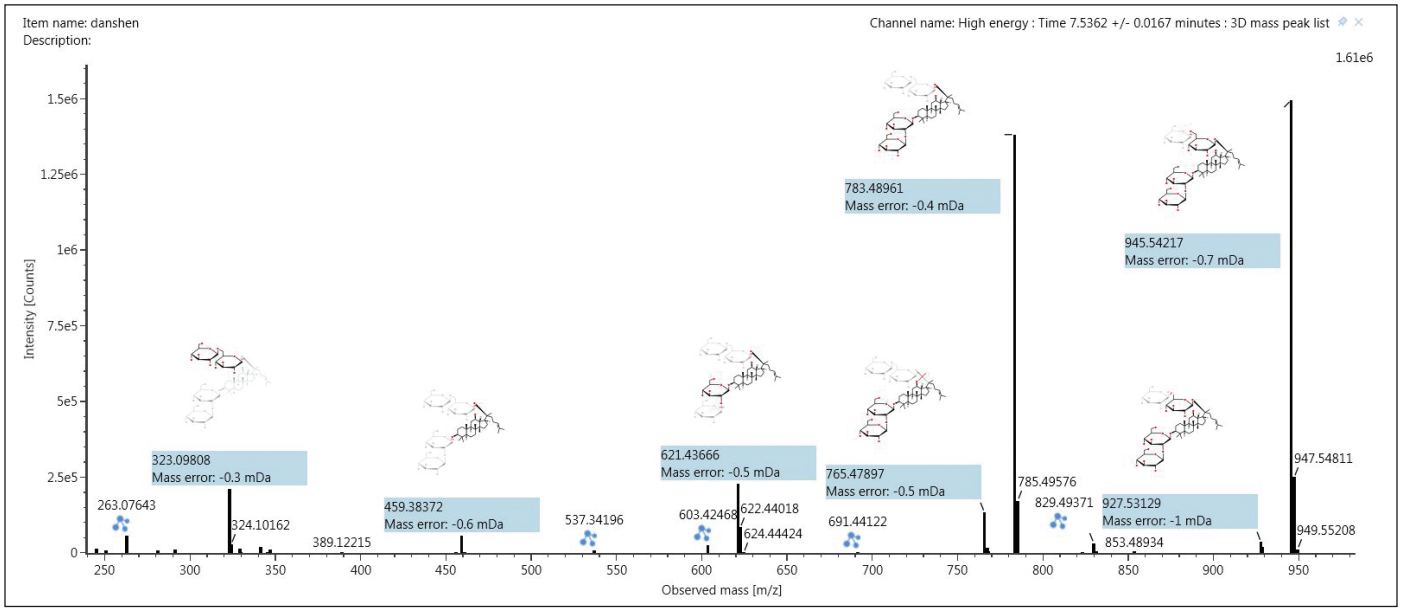

Figure 4. Structural elucidation of the fragmentation ions of ginsenoside Rb1 by MassFragment.

After component verification and confirmation as shown above, it can be observed that this unknown tablet contains chemical classes such as ginsenosides, salvia phenol, tanshinone, notoginsenoside, etc. These components are clearly associated with the herbal materials of DanShen and SanQi. Further online search indicated that some known TCM recipes that contain these two herbal ingredients could be Sanqi Danshen tablet, and Compound DanShen tablet. Of course, other recipes were also discovered, such as recipes containing American ginseng (*Panax quinquefolius*) or ShanZha (*Crataegi fructus*), etc.

Now, this research project has progressed from a non-targeted screening at the initial stage to a targeted screening, which is the chemical ingredient identification of known plants. This part of the workflow has already been well defined,[1] which is to import chemical ingredients of known herbal materials (*Salvia miltiorrhiza, Panax notoginseng, Panax quinquefolius, and Crataegi fructus*) from the Traditional Medicine Library into the target list of the UNIFI analysis method, and compare them with the experimental data.

The result was that no component was matched with the ingredients listed for the American ginseng and ShanZha (such as American ginseng saponin, gynostemma saponin, ShanZha saponin, etc., which are characteristic to these two herbs). This provides further confirmation to the initial conclusion that this tablet doesn't contain American ginseng and ShanZha. Meanwhile, all major chromatographic peaks obtained from the sample matched well with the major ingredients of DanShen and SanQi (59 key components confirmed), which is listed in Table 1.

Thus our final conclusion is this tablet was mainly composed of DanShen (*Salvia miltiorrhiza*) and SanQi (*Panax notoginseng*). The commercial product could be the Sanqi Danshen Tablet, or the Compound DanShen Tablet.

.	Component name	Formula	RT (min)	Response	m/z	Error (mDa)	Error (ppm)	Adducts	Label
1	Salvianic acid A	C9H10O5	0.99	7002	197.0451	-0.4148	-2.10	-H	confirmed
2	Protocatechuic aldehyde	C7H6O3	1.63	4283	137.0242	-0.1934	-1.41	-H	confirmed
3	Lithospermic acid	C27H22O12	3.50	19188	537.1027	-1.1279	-2.10	-H	confirmed
4	Salvianolic acid D	C20H18O10	3.57	7772	417.0820	-0.6713	-1.61	-H	confirmed
5	20-O-Glucopyranosyl ginsenoside Rf	C48H82O19	4.04	24616	1007.5421	-1.1674	-1.16	+HCOO, -H	confirmed
6	Rosmarinic acid	C18H16O8	4.09	60848	359.0770	-0.2202	-0.61	-H	confirmed
7	Salvianolic acid A	C26H22O10	4.19	129863	493.1138	-0.1842	-0.37	-H	confirmed
8	20-O-Glucopyranosyl ginsenoside Rf	C48H82O19	4.43	62970	1007.5423	-0.9179	-0.91	+HCOO, -H	confirmed
9	Notoginsenoside Fc	C58H98O26	4.51	1010	1255.6329	0.0889	0.07	+HCOO	confirmed
10	Lithospermic acid B	C36H30O16	4.60	724974	717.1453	-0.8088	-1.13	-H	confirmed
11	Notoginsenoside R1	C47H80O18	4.64	204023	977.5314	-1.2607	-1.29	+HCOO, -H	confirmed
12	Baicalin	C21H18O11	4.69	5640	445.0780	0.3632	0.82	-H	confirmed
13	Ginsenoside Rd	C48H82O18	4.97	368264	991.5475	-0.8503	-0.86	+HCOO, -H	confirmed
14	Ginsenoside Rg1	C42H72O14	4.99	1102946	845.4903	-0.0922	-0.11	+HCOO, -H	confirmed
15	Ginsenoside Rg1	C42H72O14	5.25	12929	845.4902	-0.2179	-0.26	+HCOO	confirmed
16	Monomethyl lithospermate	C28H24O12	5.37	114604	551.1197	0.2131	0.39	-H	confirmed
17	Salvianolic acid A	C26H22O10	5.60	1545	493.1141	0.0894	0.18	-H	confirmed
18	Salvianolic acid C	C26H20O10	5.79	175372	491.0990	0.6454	1.31	-H	confirmed
19	Dimethyl lithospermate	C29H26O12	5.81	10725	565.1354	0.2655	0.47	-H	confirmed
20	Dimethyl lithospermate	C29H26O12	5.93	92868	565.1357	0.5784	1.02	-H	confirmed
21	20-O-Glucopyranosyl ginsenoside Rf	C48H82O19	6.33	19364	1007.5433	0.0201	0.02	+HCOO, -H	confirmed
22	Notoginsenoside T	C64H108O31	6.62	8889	1417.6854	-0.2913	-0.21	+HCOO, -H	confirmed
23	Notoginsenoside Fa	C59H100O27	6.93	44213	1285.6447	1.2916	1.00	+HCOO, -H	confirmed
24	Ginsenoside Rg1	C42H72O14	7.06	34689	845.4906	0.2163	0.26	+HCOO	confirmed
25	Cryptoacetalide	C18H22O3	7.10	9539	285.1497	0.0427	0.15	-H	confirmed
26	Notoginsenoside T	C64H108O31	7.16	7914	1417.6854	-0.2771	-0.20	+HCOO, -H	confirmed
27	Notoginsenoside Fa	C59H100O27	7.20	83126	1285.6443	0.8786	0.68	+HCOO, -H	confirmed
28	Notoginsenoside R2	C41H70O13	7.21	219884	815.4801	0.3041	0.37	+HCOO, -H	confirmed
29	Ginsenoside Rb3	C53H90O22	7.28	20796	1137.6065	0.3184	0.28	+CH3COO	confirmed
30	Notoginsenoside S	C63H106O30	7.43	871	1387.6732	-1.9177	-1.38	+HCOO, -H	confirmed
31	Notoginsenoside Fa	C59H100O27	7.49	42092	1285.6434	-0.0362	-0.03	+HCOO, -H	confirmed
32	20(S)-Ginsenoside Rg3 (Ginsenoside Rg3)	C42H72O13	7.51	83032	829.4949	-0.5657	-0.68	+HCOO, -H	confirmed
33	Ginsenoside Rb1	C54H92O23	7.54	492734	1153.6009	-0.2281	-0.20	+HCOO, -H	confirmed
34	20(S)-Ginsenoside Rh1 (Ginsenoside Rh1)	C36H62O9	7.65	157145	683.4374	-0.2032	-0.30	+HCOO	confirmed
35	Ginsenoside Rb1	C54H92O23	7.66	8457	1153.6006	-0.5535	-0.48	+HCOO	confirmed
36	Notoginsenoside Fc	C58H98O26	7.76	23444	1255.6320	-0.8412	-0.67	+HCOO, -H	confirmed
37	Ginsenoside Rb2	C53H90O22	8.06	97093	1123.5891	-1.4610	-1.30	+HCOO, -H	confirmed
38	20(S)-Ginsenoside Rh1 (Ginsenoside Rh1)	C36H62O9	8.52	37225	683.4371	-0.5177	-0.76	+HCOO	confirmed
39	Notoginsenoside Fe	C47H80O17	8.54	15952	975.5528	-0.6039	-0.62	+CH3COO	confirmed
40	Ginsenoside Rd	C48H82O18	8.68	880119	991.5489	0.5661	0.57	+HCOO, -H	confirmed
41	Ginsenoside Rb3	C53H90O22	9.11	1715	1123.5896	-0.9565	-0.85	+HCOO	confirmed
42	Ginsenoside Rd	C48H82O18	9.21	105207	991.5476	-0.7623	-0.77	+HCOO, -H	confirmed
43	20-O-Glucopyranosyl ginsenoside Rf	C48H82O19	9.39	3311	961.5371	-0.6979	-0.73	-H	confirmed
44	Ginsenoside Rh4	C36H60O8	10.23	65829	665.4266	-0.4154	-0.62	+HCOO	confirmed
45	Danshenxinkun A	C18H16O4	10.27	22334	295.0972	-0.4093	-1.39	-H	confirmed
46	Ginsenoside Rh4	C36H60O8	10.39	97784	665.4261	-0.9420	-1.42	+HCOO	confirmed
47	20(S)-Ginsenoside Rg3 (Ginsenoside Rg3)	C42H72O13	10.70	68724	829.4949	-0.6142	-0.74	+HCOO, -H	confirmed
48	Ginsenoside F2	C42H72O13	10.79	8835	829.4944	-1.1291	-1.36	+HCOO	confirmed
49	Methylenedihydrotan-shinquinone	C18H16O3	11.12	54015	279.1022	-0.5008	-1.79	-H, +CH3COO	confirmed
50	Salviolone	C18H20O2	11.25	28599	313.1442	-0.3307	-1.06	+HCOO	confirmed
51	Dihydrotanshinone I	C18H14O3	11.57	171319	277.0868	-0.2240	-0.81	-H	confirmed
52	Sugiol	C20H28O2	11.85	53468	299.2016	-0.0323	-0.11	-H	confirmed
53	Tanshinone II B	C19H20O3	12.31	95703	295.1340	0.0329	0.11	-H	confirmed
54	Miltirone	C19H22O2	12.48	47717	281.1547	-0.0260	-0.09	-H	confirmed
55	Salvianen	C21H21NO2	12.66	3930	378.1703	-0.7997	-2.11	+CH3COO	confirmed
56	Miltiodiol	C19H22O3	13.37	5256	297.1492	-0.4174	-1.40	-H	confirmed
57	Ursolic acid	C30H48O3	13.52	53957	455.3529	-0.1933	-0.42	-H	confirmed
58	Linolic acid	C18H32O2	13.73	80398	279.2330	0.0428	0.15	-H	confirmed
59	Hexadecanoic acid	C16H32O2	14.19	145228	255.2332	0.2836	1.11	-H	confirmed

Table 1. Summary table of identified components for the unknown tablet. The table was automatically obtained by importing the Component Summary Reporting Template in UNIFI.

CONCLUSIONS

This application note has described the overall workflow obtained by applying the Natural Products Application Solution with UNIFI to identify and deduce chemical and herbal composition for unknown samples. The workflow progressed from an initial non-targeted screening into a targeted screening process.

Sample analysis by UPLC/QTof MS required just 14 minutes. The initial non-targeted screening identified 37 major chemical ingredients, which clearly showed association with DanShen (*Salvia miltiorrhiza*) and SanQi (*Panax notoginseng*). By searching the known TCM recipes from the Internet, ingredients related to relevant herbs (*Salvia miltiorrhiza, Panax notoginseng, Panax quinquefolius,* and *Crataegi fructus*) were matched against components detected from experiment data for the second time. As a result, among the 103 chemical ingredients associated with *Salvia miltiorrhiza* and *Panax notoginseng* within the Traditional Medicine Library, 59 were identified and confirmed. No match was found to match any of the major chemical ingredients related to the other two potential herbs, *Panax quinquefolius* and *Crataegi fructus*. This led to our final conclusion that the unknown product could be either a Sanqi Danshen or Compound DanShen tablet.

The Natural Products Application Solution with UNIFI is based on UPLC/QTof MS^E data acquisition, accompanied by the Traditional Medicine Scientific Library, which are integrated with an automatic identification process. This is a novel approach for ingredient analysis of total unknown samples. The result is the reduction of the blindness of such a research and significant enhancement of productivity.

Reference

1. Using Natural Products Application Solution with UNIFI for the Identification of Chemical Ingredients of Green Tea Extract. Waters Application Note, November 2013; 720004837en.

Waters

THE SCIENCE OF WHAT'S POSSIBLE.®

Waters Corporation
34 Maple Street
Milford, MA 01757 U.S.A.
T: 1 508 478 2000
F: 1 508 872 1990
www.waters.com