NPR

A Comprehensive Classification of Solvent Systems used for Natural Product Purifications in Countercurrent and Centrifugal Partition Chromatography'

SCHOLARONE™
Manuscripts

ROYAL SOCIETY
OF CHEMISTRY

NPR

REVIEW

A comprehensive classification of solvent systems used for natural product purifications in countercurrent and centrifugal partition chromatography

Krystyna Skalicka-Woźniak,*[a] and Ian Garrard[b]

Using both library paper copies and modern electronic copies, every known, published, English-language journal paper that employs either countercurrent or centrifugal partition chromatography solvent systems for natural product purifications has been studied and the solvent systems classified in a comprehensive database. Papers were studied from the earliest found examples containing natural product separations in 1984 until the end of 2014. In total, 2594 solvent systems have been classified of which 272 are gradient systems. To observe any trends or patterns in the data, the natural product solutes were divided into 21 classes and the solvent systems into 7 different types. The complete database, sorted according to natural product class, is available for download to assist separation scientists in future liquid-liquid chromatography purifications.

1 Introduction

The purification of natural products is a complex process, due in part to the vast array of diverse molecules available for discovery in our natural world. The development of a purification protocol to pull a single, active, target molecule from the many hundreds that may be present in a mixture is a task that can take months or even years to achieve. To succeed with this rather daunting challenge, the natural product chemist requires, at their disposal, a diverse range of purification techniques.

Countercurrent chromatography (CCC), together with its sister technique, centrifugal partition chromatography (CPC), is one such technique which has shown considerable promise when applied to the purification of crude natural product mixtures. The technique utilises two immiscible phases. Frequently one is water-based and the other an organic solvent, but it is also possible to have solvent-solvent phase systems and aqueous two-phase systems. In CCC, one of these phases (the stationary phase) is retained in a spinning coil by centrifugal forces while the other phase (the mobile phase) is pumped through to effect the separation.[1] In CPC, the stationary phase retention is assisted by the use of tiny oblong or rectangular chambers connected by capillary ducts.[2]

The technique has a number of significant advantages over traditional solid-phase chromatography systems and many of these advantages particularly come into their own with natural product separations. For example, the solvent usage is generally far lower than that of solid phase chromatography systems operating at the same scale.[3] Furthermore, since the process is frequently an isocratic one, a simple analysis of solvent composition allows the recycling of the solvents, reducing the usage still further.[4] The technique also allows for 100% recovery of the sample components, since there is no loss arising from irreversible adsorption onto the solid matrix. Particulates, such as cell debris, are tolerated in CCC.[5] Thus filtering a sample is not always necessary and a direct extraction of compounds from a crude mixture is possible. With processing times similar to that of other purification methods[3], scale up is also possible with modern instruments, with examples existing running from milligram to kilogram levels.[6,7] A wide range of polarities can be processed due to the range of possible solvent systems. The literature reports examples with a logP range of at least −4.7 (colistin peptide antibiotic)[8] to +17.6 (lycopene)[9]. Running and maintenance costs are low since a set of coils would be expected to last the lifetime of the centrifuge and there are no columns to replace. Unlike solid phase chromatography, there is no change to component retention over time (no column aging effects) as a freshly-filled coil of solvents is used each run. This makes it easier to consistently satisfy current regulatory requirements when performing purifications under a GLP or GMP

a. Department of Pharmacognosy with Medicinal Plant Unit, Medical University of Lublin, 1 Chodzki Str, 20-093 Lublin, Poland. E-mail: kskalicka@pharmacognosy.org Tel/Fax.: +48 81448 7080.
b. Advanced Bioprocessing Centre, Brunel Institute for Bioengineering, Brunel University, Uxbridge UB8 3PH, UK. E-mail: ian.garrard@brunel.ac.uk

environment. All of these advantages have resulted in CCC and CPC showing an exponential increase in popularity over the past 15 years, as demonstrated by the number of publications utilising the technique.

However the biggest drawback to the technique, and the one that most inhibits its general take-up, is the selection of an appropriate two-phase solvent system. With the possible combinations of solvents almost infinite and the modification of one phase necessarily altering the composition of the other phase, this is by far the most time-consuming aspect of the technique. Many authors have tried to develop a systematic approach to the problem. For example, Foucault & Chevolot[10] use a "best solvent" approach, whilst Margraff[11] and Oka[12] developed systematic selection tables. Even liquid-handling robots have been employed to assist with this selection process.[13] However, when faced with a new separation of a natural product mixture, the technique that is frequently initially employed by countercurrent chromatographers is to search the literature to see whether similar compounds have been separated before. The solvent system thus employed for a previous, similar separation can be used as an excellent starting point for the purification protocol.

In this publication, the authors have searched every known, published, English-language journal paper (including conference papers) that employs CCC or CPC solvent systems and thus produced a comprehensive classification of solvent systems for natural product purifications. The authors hope that this reference will greatly assist future chromatographers in finding the best solvent system to employ.

2 Methodology

At Brunel University, a full-time librarian was employed for a number of years to collect and file all English-language, peer-reviewed, journal articles relating to CCC or CPC. This collection, kept at the Brunel Institute for Bioengineering, consists of 1487 papers spanning from 1966[14] until late 2008, although the first papers quoting aqueous-organic solvent systems for CCC or CPC applications were published in 1984. Beyond 2008, it was assumed that all publications could be found electronically. To cover a period of overlap, the electronic searching commenced with the year 2005. The broad search terms "current + chromatography" and "centrifugal + partition + chromatography" were used in the academic search program Scopus to call up all papers on CCC and CPC. It was also noticed that some journals allowed the use of abbreviations in the title and abstract, so the additional search terms, "CCC", "HPCCC", "HSCCC", and "CPC" were searched. Each hit was then individually inspected before entering the solvent system and data presented. This was done for each year from 2005 until December 2014. In addition, for the year 2005, these search terms were used in the literature search programs PubMed and SciFinder. This called up many entries, including lots of foreign language papers. Each was inspected individually but not a single extra English language paper on CCC or CPC purifications was found

that had not been picked up in Scopus. For the remaining years therefore, only Scopus was used for the search.

In total, over 4000 papers related to CCC or CPC were studied over the period of a year. Papers that contained an application example of natural product, i.e. a purification performed by CCC or CPC, were noted in an Excel spreadsheet, together with the compounds purified and the solvent system used for the purification. Papers which gave examples of separations reported elsewhere were ignored, as were papers that described the aqueous-aqueous polymer systems used to purify many peptides and proteins. Only papers which gave specific details of the solvent system and solute were recorded and only natural product secondary metabolites were noted e.g no synthetic compounds, dyes or chemicals.

In total therefore, 2594 solvent system entries were listed together with the corresponding solutes that they separated. Some of the solvent systems corresponded to more than one solute, and some of the solutes corresponded to more than one solvent system, but if the same solute and the same solvent system were listed, this was simply a duplicate entry and was therefore removed. Overall, this data was entered into a large spreadsheet with 32 different make-up solvents on one axis and 2594 solutes separated on the other axis.

In order to study trends in the data table, the solvent systems were then further divided into six hierarchical groups, each based on the primary, water-immiscible solvent in the system.

1. Alkane-based systems (hexane, heptane, isooctane, including cyclohexane)
2. Chlorinated solvents (chloroform, dichloromethane, dichloroethane, tetrachloromethane, trichloroethane)
3. Ethers (t-buthylmethyl ether, light petroleum, diethyl ether)
4. Water-immiscible alcohols (butanol, methanol, ethanol)
5. Water-immiscible esters (ethyl acetate, methyl acetate)
6. Miscellaneous (remaining solvent systems that do not fit any of the above classes e.g. toluene, butanone)

The classification was established as a hierarchy to avoid duplicating entries. So solvent systems containing, for example, hexane were placed into group one, even if they also contained ethyl acetate.

In addition, the solutes separated were classified into 21 natural product classes as follows:

1. Alkaloids
2. Amides & Alkylamides
3. Anthocyanins
4. Anthraquinones & Naphthoquinones
5. Aromatic Derivatives (including acetophenones, alkylbenzene derivatives and bibenzyls, phloroglucinol derivatives, thiophene)
6. Coumarins & Chromones
7. Flavonoids & Flavonolignins
8. Glucosinolates
9. Iridoids & Secoiridoids
10. Lactones
11. Lignans

12. Lipids (including fatty acids and fatty acid esters, galacto-phospholipids)

13. Phenol Derivatives (including stilbenes, gingerols, bitter alpha acids & phenylpropanoids, phenylethanoids)

14. Phenolic Acids & Derivatives (including benzyl esters)

15. Phthalides

16. Polyphenols (including catechins, tannins and theaflavins)

17. Saponins

18. Steroids (including cardenolides and sterols)

19. Terpenoids (including carotenoids)

20. Xanthones

21. Miscellaneous (anything that does not fall into the other categories e.g. chlorophyll, squamocin, maleic acid, sulforaphene, sucrose, glucuronic acid and falcarinol, betalain pigments, decarboxylic acids, acetogenins)

All gradient systems were classified in a separate table. Both stepwise and linear gradients of mobile phase were included but it should be noted that rare examples exist in the literature that include a temperature gradient or a stepwise flow gradient and these could not easily be included, though a note was placed in a comments column. It should also be noted that any change in one phase composition necessarily induces a change in the other phase composition.[15] Linear gradients were entered as a range under the relevant solvent e.g. a gradient from 10 to 25% methanol was entered under the methanol column as 10-25. Stepwise gradients had their steps expressed e.g. a stepwise methanol gradient using 10, 15, 20 and 25% methanol was entered as 10-15-20-25.

A major exclusion was any system using pH zone refining for elution of components. However, systems that had acids, alkalis or salts added as modifiers had this noted in a comments column, unless the modifier was a volume part of the solvent system and this was only the case for acetic acid. Thus the system butanol-water-acetic acid (5:4:1) had 1 part (or 10%) listed under the Acetic Acid column, whereas the system ethyl acetate-2% aqueous acetic acid solution (1:1) had 1 part (or 50%) entered under the Water column and "2% acetic acid solution" added to the Comments column.

3 Observations on the Data

In total, 2322 isocratic solvent systems were classified made up from 29 different solvents of which water (or water containing a salt or pH modifier) was one. The solvents butyronitrile, pentane and trichlorethane (1,1,1) were only found in gradient solvent systems. Figure 1 shows the relative proportions of solvents in all 2322 of these solvent systems by way of the number of examples utilising that particular solvent component. The top four solvents used are water, ethyl acetate, methanol and hexane, showing the popularity of the range of solvent systems made from these four components, known as HEMWat systems. The next two most popular solvents are butanol (for polar systems) and chloroform, which despite having serious environmental and health concerns, is growing in worldwide popularity (Fig. 2) with 295 isocratic entries and 9 gradient entries being solvent systems based on this solvent.

Water was used in 2236 out of 2322 isocratic solvent systems, meaning that 86 solvent systems recorded (3.7%) were non-aqueous, organic-organic phase systems. Of these 86 systems, 78 contain acetonitrile and 67 hexane (with a further 8 containing the similar heptane). The next most popular non-aqueous phase system solvents were methanol (20 systems) and tert-butylmethyl ether tBME (14 systems) (Fig. 3). Of the gradient solvent systems, 18 systems (6.6%) used no aqueous component.
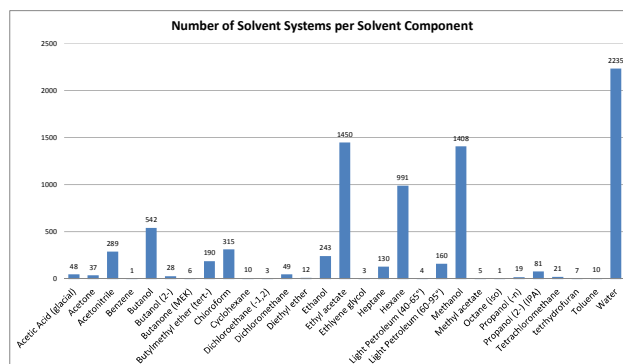

Fig. 1 The total number of isocratic CCC and CPC solvent systems containing the solvent component labelled as part of the two-phase solvent system.
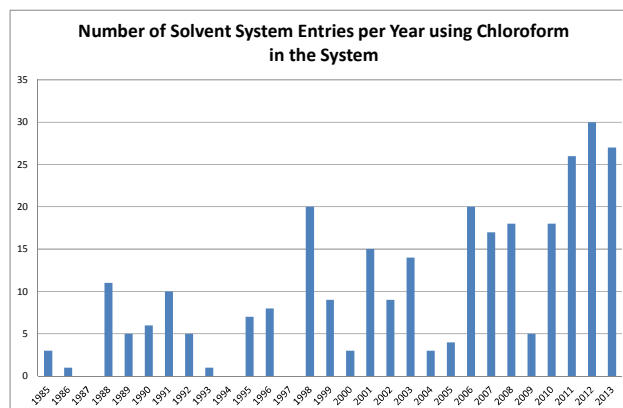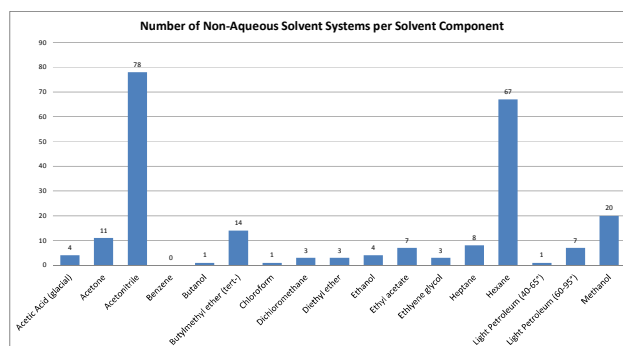

Fig. 2 The number of CCC and CPC solvent system entries published each year using chloroform as one of the components of the solvent system.

Fig. 3 The number of non-aqueous CCC and CPC solvent system entries using the labelled solvents as one of the components of the solvent system.

In this study, only papers containing details of a solvent system and a natural product solute were recorded. Papers looking at the engineering or modelling of the technique, for example, were not classified, likewise papers using aqueous two-phase systems, pH zone refining or synthetic compounds. However, despite this restricted sub-set of papers counted in the topic area, the growth of the technique is apparent (Fig. 4). The number of natural product purifications published each year has been growing at a steady rate since 1995.
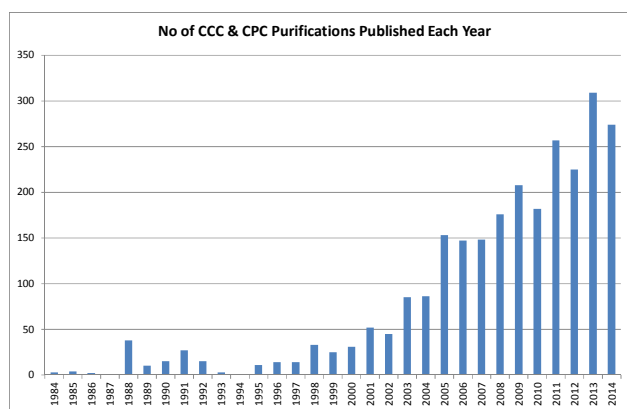

Fig. 4 The total number of CCC and CPC purifications published each year only including those describing solvent systems for natural product separations.

The organisation of natural products into compound classes is a task that will always be challenging. With many hundreds of thousands of compounds to classify, and the almost continuous spectrum of compounds spanning from class to class, no attempt is likely to be fully satisfactory. However, a broad classification of compound class helps identify the type of solute and thus the authors have placed the compounds separated into 21 categories as listed above. The number of entries discovered for each type of solvent system in each of these compound classes is shown in Table 1. By far the most popular class of natural product separated by CCC is the class of flavonoids and flavonolignans (a total of 695 published entries). With less than half that amount, terpenoids and alkaloids come second and third   (312 and 297 published entries each respectively), and with almost half that amount again come phenol derivatives (188 entries) and saponins (155 entries). It therefore seems that CCC particular suits certain classes of natural products, especially the flavonoids and flavonolignans. The solvent system most commonly employed for the separation of flavonoids is an alkane-based one, although this is true for the majority of other compound classes too. The saponins class is one exception to this trend, with chlorinated solvent systems proving the most popular and water-immiscible alcohols also just showing more published entries than the alkane-based solvent systems. Similarly, the most popular solvent systems for the purification of Iridoids, Secoiridoids and polar Glucosinolates are the ones based on

water-immiscible alcohols, while ether-based systems proved most popular for the purification of anthocyanins. Gradient systems were most commonly used to purifiy Flavonoids & Flavonolignans and Terpenoids.

Within the isocratic separations, 457 systems (19.7%) contain an additive of sorts with 85 of these having an additive in both the mobile and the stationary phase. The most common additive to a solvent system is an acid, with 389 systems (16.8%) containing one of just four acids, either formic acid (18 systems, 0.8%), hydrochloric acid (114 systems, 4.9%), trifluoroacetic acid (147 systems, 6.3%) or acetic acid (110 systems, 4.7%). Glacial acetic acid was also frequently used as an integral part of the solvent system and a further 48 systems adopted this approach, making this acid the most popular with 158 systems (6.8%) utilising it. For basic additives, ammonia (28 systems, 1.2%) and trimethylamine (66 systems, 2.8%) are the most popular.

The remaining systems utilising additives contain a varied range of chemicals from bases such as sodium hydroxide and diethylamine to acids such as methanesulfonic acid and trifluoric acid to salts such as sodium chloride, ammonium formate, ammonium sulphate, copper nitrate, sodium carbonate, sodium iodide, sodium phosphate and silver nitrate. No clear pattern of usage was identified for these additives.

The majority of published CCC purifications used isocratic solvent systems but 10.5% of systems (272 out of 2594 systems) used a gradient. As with the isocratic systems, gradients were most frequently employed to separate flavonoids and flavonolignans, however terpenoids, phenol derivatives, coumarins and chromones, and saponins also feature strongly with gradient separations. For the gradient solvent systems, 187 (68.8%) are methanol gradients with a further 15 (5.5%) being ethanol gradients and a further 40 (14.7%) being butanol-based gradients. It is clear that alcohol-based gradients are by far the most popular type in CCC purifications.

Of the remaining 30 gradient solvent systems that are not alcohol driven, ethyl acetate is by far the most commonly used gradient solvent (25 systems, 9.2%). The earliest discovered paper mentioning gradient elution was published in 1973 by Ito and Bowman[16]. However this used a primitive I-type centrifuge and the separation was performed on an artificially-created mixture of dipeptides. The earliest natural product gradient separation that was found was published by Vanhaelen and Vanhaelen-Fastre in 1988 on the separation of flavonol glycosides from *Ginko biloba* leaves[17].

4 Conclusions

Although initially dismissed by many chemists and purification scientists as slow, unreliable and temperamental, steady development of the technique of CCC on both the engineering and the application side has transformed it into a technique worthy of inclusion in the natural product scientist's arsenal. On one side, engineering developments have produced machines that are robust, capable of fast, efficient separations and able to accept high injection loadings. On the other side, developments in the application protocols have produced modes of operations and solvent systems to purify out

compounds from the full polarity spectrum. Combined, these have produced a technique that is wonderfully suited to natural product purifications, particularly on a large preparative scale, with advantages over solid phase techniques such as the ability to accept particulates and to always recover all components, and advantages over the old-style liquid-liquid techniques such as high speed, high loading and high resolution. With the selection of an appropriate two-phase solvent system being the greatest barrier to the acceptance of this interesting purification technique, it is hoped that the creation of a comprehensive data base of previous solvent systems will encourage more separation scientists to consider liquid-liquid approaches in their work. The full tables, together with the complete list of references used in their creation, are available as Supplementary Material. Tables are presented as follow:

- Table S1 – Isocratic Solvent Systems table
- Table S2 – Gradient Solvent Systems table
- Doc. S1 – Complete List of References.

5 Acknowledgements

6 References

1   A. Berthod, Comp. Anal. Chem., 2002, 38, 1-20.
2   K. D. Yoon, Y. W. Chin and J. Kim, J. Liq. Chromatogr. Rel. Technol. 2010, 33, 1208-1254.
3   I. Sutherland, C. Thickitt, N. Douillet, K. Freebairn, D. Johns, C. Mountain, P. Wood, N. Edwards, D. Rooke, G. Harris, D. Keay, B. Mathews, R. Brown, I. Garrard, P. Hewitson and S. Ignatova, J. Chromatogr. A, 2013, 1282, 84-94.
4   I. Garrard, L. Janaway and D. Fisher, J. Liq. Chromatogr. Rel. Technol., 2007, 30, 151-163.
5   Y. Ito, J. Chromatogr. A, 2005, 1065, 145-168.
6   P. Wood, S. Ignatova, L. Janaway, D. Keay, D. Hawes, I. Garrard and I. A. Sutherland, J. Chromatogr. A, 2007, 1151, 25-30.
7   I. Garrard, D. Fisher and I. A. Sutherland, LCGC North America, 2008, 26, 424-438.
8   Y. Y, H. Oka, J. Hayakawa, N. Kawamura, K. Harada, M. Suzuki, H. Nakazawa and Y. Ito, J. Liq. Chromatogr. Rel. Technol., 1998, 21, 143-155.
9   Y. Wei, T. Zhang, G. Xu and Y. Ito, J. Chromatogr. A, 2001, 929, 169-173.
10 A. P. Foucault and L. Chevolot, J. Chromatogr. A, 1998, 808, 3-22.
11 E. Camacho-Frias and A. Foucault, Analusis, 1996, 24, 159-167.
12 F. Oka, H. Oka and Y. Ito, J. Chromatogr. A, 1991, 538, 99-108.
13 I.J. Garrard, J. Liq. Chromatorg. Rel. Technol., 2005, 28, 1923-1935.
14 Y. Ito, M. Weinstein, I. Aoki, R. Harada, E. Kimura and K. Nunogaki, Nature, 1966, 212, 985-987.
15 S. Ignatova, N. Sumner, N. Colclough and I. Sutherland, J. Chromatogr. A, 2011, 1218, 6053-6060.
16 Y. Ito and R.L. Bowman. J.Chromatogr.Science, 1973, 11, 284-291.
17 M. Vanhaelen and R. Vanhaelen-Fastre. J.Liq.Chrom., 1988, 11(14), 2969-2975.

	Type of Solvent System							
Natural Product Class	**Alkane based**	**Chlorinated**	**Ethers**	**Water-immiscible alcohols**	**Water-immiscible esters**	**Miscellaneous**	**Gradient Solvent Systems**	**TOTAL**
Alkaloids	105	76	79	26	1	0	10	297
Amides & Alkylamides	26	6	0	0	0	0	0	32
Anthocyanins	2	0	56	9	2	0	7	76
Anthraquinones & Naphthoquinones	40	8	25	1	3	0	16	93
Aromatic derivatives	17	2	9	10	0	0	6	44
Coumarins & Chromones	51	8	19	11	1	0	27	117
Flavonoids & Flavonolignans	279	97	67	96	89	6	61	695
Glucosinolates	1	0	0	7	0	3	0	11
Iridoids & Secoiridoids	3	8	2	20	1	0	0	34
Lactones	13	0	0	0	0	0	0	13
Lignans	36	8	7	8	1	0	19	79
Lipids	15	2	0	5	0	0	2	24
Miscellaneous	25	4	9	11	0	0	0	49

Phenol derivatives	59	22	17	38	17	6	29	188
Phenolic acids and derivatives	64	18	23	17	13	0	8	143
Phtalides	26	0	1	0	0	0	0	27
Polyphenols	82	0	4	15	3	0	1	105
Saponins	31	52	5	37	4	0	26	155
Steroids	26	12	1	0	0	0	18	57
Terpenoids	210	24	28	5	3	0	42	312
Xanthones	20	6	8	4	5	0	0	43
TOTAL	**1131**	**353**	**360**	**320**	**143**	**15**	**272**	**2594**

Table 1 The number of published purifications classified by class of compound and type of solvent system employed