ChemComm

Optical Discrimination of Terpenes in Citrus Peels with a Host:Guest Sensing Array

Journal:	*ChemComm*
Manuscript ID	CC-COM-03-2024-001309.R1
Article Type:	Communication

ChemComm

COMMUNICATION

Optical Discrimination of Terpenes in Citrus Peels with a Host:Guest Sensing Array

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Junyi Chen,[a] Jose L. Moreno Jr.,[a] Wen Zhang,[a] Lucas J. Gibson-Elias,[a] Ria Lian,[a] Saba Najafi,[a] Haofei Zhang,[a] Wenwan Zhong[b]* and Richard J. Hooley[a]*

A simple aqueous host:guest sensing array can selectively discriminate between different types of citrus varietal from peel extract samples. It can also distinguish between identical citrus samples at varying stages of ripening. The discrimination effects stem from detection of changes in the terpenoid composition of the peel extracts by the host:guest array, despite the overwhelming excess of a single component, limonene, in each sample. The hosts are insensitive to limonene but bind other monoterpenes strongly, even though they are similar in structure to the major limonene component. This work demonstrates the capability of host:guest arrays in sensing target molecules in environments with the competing agents present at high abundances in the sample matrix.

Recognition and sensing of biorelevant molecules in biological environments[1] is complicated by the presence of competing targets in the sample matrix. This is most obviously seen when using macrocyclic cavity-containing hosts for recognition: there are myriad examples of target binding in organic solvents[2] or aqueous solution,[1,3] but fewer examples are seen in more complex systems such as in saliva, urine, serum, or living cells.[4] Moreover, the targets that are most accessible in complex biological environments tend to be those with uncommon structures that allow good selectivity: the R-NMe_3^+ in choline is a good example,[5] as are unique anions such as thiolates.[6]

One class of molecules almost exclusively absent from optical detection in biomedia are neutral hydrocarbons, notably terpenes. These species are the perfect targets for macrocyclic detection in water, as hydrophobic effects favor their binding inside lipophilic cavities.[7] Deep cavitands and toroidal macrocycles are excellent hosts for cyclic and polycyclic hydrocarbons in water,[8] but they are also promiscuous, so show little selectivity for different hydrocarbons unless there are large size and shape differences: essentially, if it fits, it sits. This promiscuity has many advantages in molecular recognition, catalysis and sensing,[9] but makes selective detection of hydrocarbon elements in natural systems very difficult. This is where targeted differential sensors have an advantage: for example, peptide-derived sensors can differentiate wine varietals with chelating metal-coordination motifs.[10] However, as synthetic host molecules are among the most successful receptors for neutral hydrocarbons, this type of recognition and sensing should be possible in a host system.

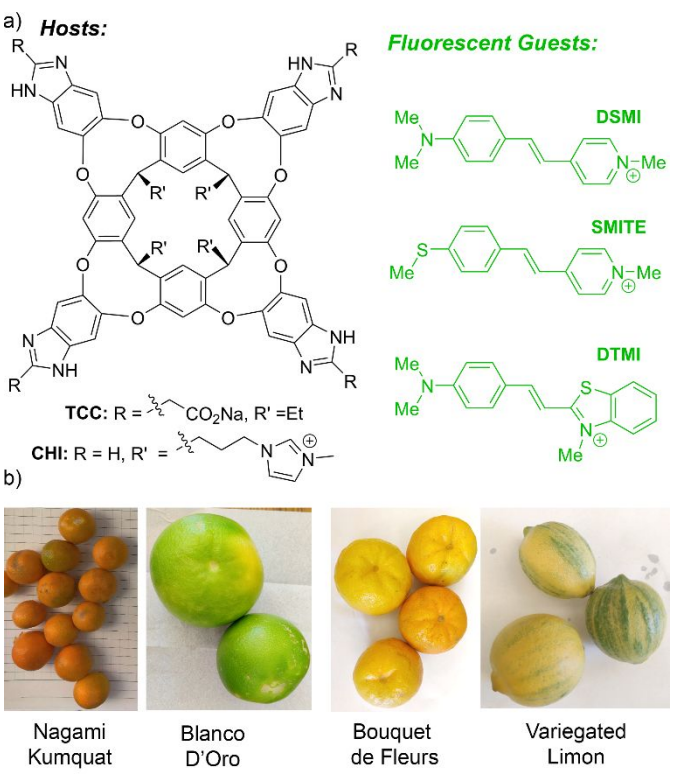

Figure 1. a) Cavitands and indicator dyes used for optical sensing of b) different citrus varietals.

[a]*Department of Chemistry; University of California–Riverside, Riverside, CA 92521, U.S.A. *E-mail: richard.hooley@ucr.edu*
[b] *Key Laboratory of Precision and Intelligent Chemistry; Department of Chemistry, School of Chemistry and Materials Science; University of Science and Technology of China, Hefei, Anhui, 230026, P.R.China. *E-mail: wenwanz@ustc.edu.cn*
† Electronic Supplementary Information (ESI) available: Experimental methods, fluorescence, UV/Vis absorbance, GC-MS and NMR data not shown in the text. See DOI: 10.1039/x0xx00000x

An excellent example can be found in the discrimination and identification of citrus varietals. The peels of citrus such as oranges, lemons, limes and others contain a wide variety of terpenoid species, which have been used in perfumes, scents and flavorings.[11] However, the dominant component is limonene – the distillation and isolation of limonene from orange peel is a well-known undergraduate laboratory experiment.[12] The differences lie in the minor terpenoid components, which are often similarly structured to limonene and are present at far lower concentrations.

These small differences are well-suited for differential sensing, however, as small differences in structure can be teased out by the application of multiple different hosts and indicators in an array-based format.[13] We have shown that host-based differential sensing can discriminate small molecule targets such as steroids, drugs of abuse, insect pheromones and other biological targets in biomedia such as saliva, urine and cells.[8,14] This led to the question: can a host-based sensor array discriminate citrus varietals, based on extracts from their peel, despite the overwhelming dominance of limonene?

We tested this possibility using four different citrus varietals (japonica nagami kumquat, Blanco D'Oro grapefruit, "Bouquet de Fleurs" sour orange, and variegated limon). They were directly picked from trees on the UC Riverside campus at the same time of year (October 2023). The fruits were peeled, and the peel components isolated *via* a simple CH_2Cl_2 extraction, filtered, evaporated, and reconstituted in 1,2-dimethoxyethane (DME) to form stock solutions. The samples were analyzed by GC-MS to determine the overall composition of the extracts. As expected, the dominant component in each case is limonene (Figures S-1 – S-10, Table S-1 – S-3). The proportion of limonene varied from 95% of the total constituents (sour orange) to ~79% (limon). The other terpenes and terpenoids all showed abundances ~100-fold lower than limonene, and most were <1% of the sample. Furthermore, many of the minor constituents are isomers of limonene. This illustrates the challenge: while the four citrus species exhibit different compositions, they contain a large mixture of different, yet structurally similar species, and a single component is dominant in all cases.

We have previously shown that deep cavitands such as **TCC** and **CHI** are capable of binding small molecule hydrocarbons in water,[15] as well as hydrophobic species such as tetrahydrocannabinol and insect pheromones.[14] Both **TCC** and **CHI** (Figure 1) form kinetically stable, folded conformations in water, and are soluble up to 1 mM, so these hosts were chosen for the detection. They pair well with styrylpyridinium dyes for optical detection of hydrophobic species, so three of these dyes were chosen, **DSMI**, **DTMI** and **SMITE** (Figure 1).[14] The selected elements were also chosen (in part) to minimize their overlap with absorptive compounds in the extracts: the maximal absorbance of the extracts themselves is λ_{max} = 320 nm, with minimal absorbance at 390 nm, while the selected dyes and their host:dye complexes all show excitation maxima of 390 nm or above. Notably, host:guest complexation causes a red shift in dye absorbance – for example, λ_{max} (**SMITE**) = 390 nm, whereas λ_{max} (**SMITE•TCC**) = 421 nm. Multiple excitation wavelengths were used, for free and bound dye in each case.

The citrus extracts were added to each of the host:dye combinations, as well as dye alone, and the emission responses recorded at multiple λ_{ex} values for free and bound dye. The fluorescence profiles were collected of 0.2 mg/ml citrus sample by using 0.5 µM dye with 4 µM **TCC/CHI** cavitand. The responses (see Figure 2a-c and S-23 – S-25) were quite variable, due to the differing behavior of each of the dyes in the different hosts. For example, **DSMI** binds strongly in both **TCC** and **CHI**, and the binding effects a fluorescence increase of 30-fold (in **TCC**) or 10-fold (**CHI**). The emission of the dyes themselves is only slightly affected by the extracts. Addition of the various citrus extracts caused a significant drop (~5-fold) in the **DSMI•CHI** emission, but only a small one in that of **DSMI•TCC**. This suggests that indicator displacement is one of the contributing mechanisms, and the stronger dye binding in anionic **TCC** vs cationic **CHI** is a cause of the differing response. However, a second mechanism can also occur: both **TCC** and **CHI** are prone to triggered aggregation in the presence of lipophilic species,[14b] and this also causes a change in dye emission - this is seen in the titration data (Figures S-35 – S-38), where both enhancement and decrease can be seen with different dye:host pairings. Nonetheless, it is clear that the different extracts effect differential responses in each of the host:dye combinations.

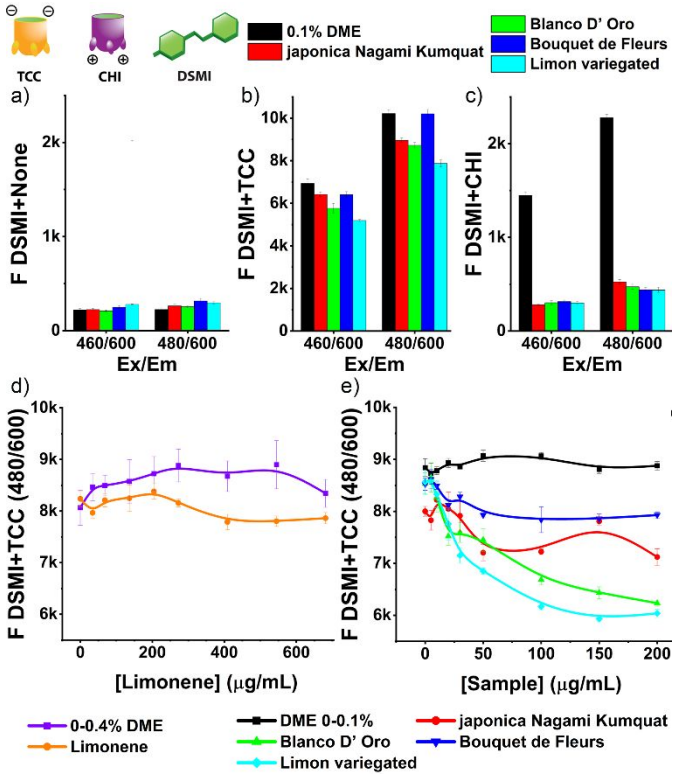

Figure 2. Fluorescence emission plots of 4 citrus varietals (200 µg/mL) sensed by **Dye•Host** complexes at the excitation frequencies of free dye and host-bound dye. a) **DSMI**; b) **DSMI•TCC**; c) **DSMI•CHI**. See Figure S-23 for further details and bar plots at all wavelengths. Fluorescence titration of **DSMI•TCC** at Ex/Em 480/600 with increasing concentrations of d) limonene and e) citrus samples. See Figure S-38 for further details.

The variability of the response in initial tests was encouraging, but introduces a question: if the response is simply due to indicator displacement and target recognition inside the hosts,

then why are there variations in signal, and why does the excess of limonene not dominate the response? The answer is that limonene is, quite surprisingly, a poor guest for both **TCC** and **CHI**. Commercial samples of limonene and seven other terpenes identified as minor components in the citrus peels (α-, β-pinene, sabinene, β-caryophyllene, δ-cadinene, α-terpineol and linalool) were sonicated with **TCC** or **CHI** ([host] = 2 or 1 mM) in D_2O, and the binding analyzed by ^{1}H NMR spectroscopy (See ESI Figures S-11 – S-14, Table S-4). Limonene did not form a kinetically stable host:guest complex with **TCC**, and showed no binding at all with **CHI.** Interestingly, despite the fact they are constitutional isomers of limonene ($C_{10}H_{16}$), α-pinene, β-pinene and sabinene in both cavitands significantly more strongly (K_a(**TCC**) = 7500, 15000, <100 M^{-1}, respectively, see Table S-4 for all affinities). The larger sesquiterpenes did not bind, nor did linalool, while terpineol bound, but shows a rapid in/out exchange profile. The affinity of limonene is many orders of magnitude lower than those between the dyes and the hosts (K_a >10^5 M^{-1}),[14c] but minor components such as α- and β-pinene bind in a similar range to the dyes. The reason for the low affinity of limonene for the cavitands is not clear, but it has been shown that alkenyl substrates bind more weakly in **TCC** than saturated hydrocarbons,[7] and the ^{1}H NMR spectra show evidence of multiple carceroisomers, suggesting that there is not a single favorable conformation for limonene when bound.

Similarly (Figure 2d, ESI), adding pure limonene to the **DSMI•TCC** complex has a minimal effect on the emission, suggesting that limonene does not cause indicator displacement. In contrast, substantial and highly variable responses are seen with the natural extracts (Figure 2e). This effect is similar (although not identical) for the other host•dye combinations (see ESI). The sensor elements are only minimally affected by the large limonene excess in the citrus samples, and are instead responsive to the minor components.

From this data, it is evident that host•dye complexes can undergo multiple different response mechanisms to the minor components of the citrus extracts, and the responses of the different dyes and hosts all vary. The host•dye complexes can either undergo indicator displacement upon treatment with the citrus extracts, or the hydrophobic species can trigger self-aggregation of the host•dye complexes.[14b] These two events can cause either enhancement of the dye emissions, or reduction. In addition, analysis at two excitation wavelengths (dye alone, host•dye) can also provide variables, allowing monitoring of indicator displacement efficiency. This complexity is perfect for differential sensing.

Of course, merely having variables in array-based sensing does not ensure that the samples can be properly discriminated. As such, we further analyzed the fluorescence profiles of the varietal extracts exposed to the host•dye array consisting of the three dyes and two hosts described above. As the responses of the extracts with the dyes alone were small, they were not included in the array. The F/F_0 values (Figures 3 and S-23 – S-25) of the host•dye elements were subjected to Principal Component Analysis (PCA). The full 16-element array can fully discriminate all the four citrus varietals (Figure 3a), but this is overkill for a 4-target sample, so we focused on creating a minimal array. Machine learning optimization (SVM-RFECV[16]) was employed, and from this, two elements were chosen as most effective (with 3-repeat 4-fold cross-validation scores equal to 1.000, see Table S-6): **DTMI•CHI(λ_{510})** and **SMITE•CHI(λ_{390})**. The discrimination was repeated with these two elements alone, and as can be seen in Figure 3b, full discrimination of the different varietals is possible with a simple 2-element host:guest sensor combination.

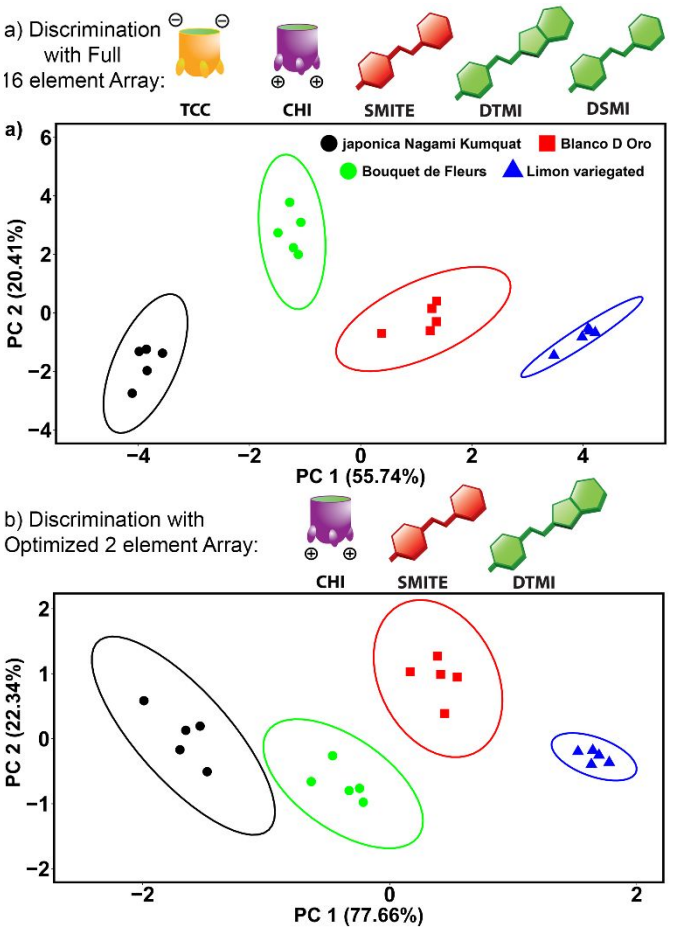

Figure 3. The differentiation of 4 citrus varietals by using a) the full 16-element array employing **DSMI, SMITE, DTMI** and **TCC, CHI;** b) the SVM-RFECV selected 2-element array **DTMI•CHI** and **SMITE•CHI.** See ESI for details and specific Ex/Em values for the arrays.

Of course, an important test of the sensor is whether it can be reproduced – discrimination is one thing, but is the sensor actually detecting a varietal, or just a sample difference? Additionally, can the sensor detect changes in peel composition from ripening over time? To test these two questions, a series of extracts from Blanco D'Oro grapefruits were obtained. Six grapefruits were harvested from the same tree on two different dates. Three (labeled as Oct-A, Oct-B, and Oct-C) were harvested on October 27, 2023, and three (Dec-D, Dec-E, and Dec-F) were harvested on December 11, 2023. The peel extract samples from the six different fruits were obtained *via* identical methods as described above, and were applied to the optimized host:guest sensing arrays (either the full 16-element or the minimal 2-element arrays, 5 repeats each). In both cases, (Figure 4a and 4b), the three different extracts from 3 different

fruits obtained at the same time almost fully overlap each other, indicating that the sensor is not detecting random differences in sample, but is identifying a specific fruit type (in this case, 3 different Blanco D'Oro fruits). In addition, the sensors were fully able to discriminate between fruits of different ripeness: in each case the October batches were widely separated on the scores plot from the December batches.

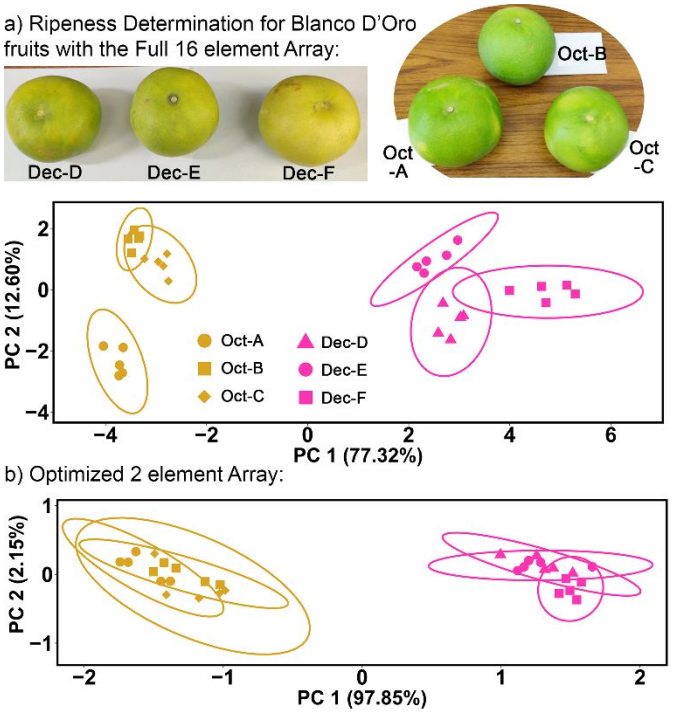

Figure 4. The differentiation of six Blanco D'Oro grapefruit samples by using a) the full 16-element array employing **DSMI, SMITE, DTMI** and **TCC, CHI;** b) the SVM-RFECV selected 2-element array **DTMI•CHI** and **SMITE•CHI.** See ESI for details and specific Ex/Em values for the arrays.

These results show that a simple 2 element host:guest array is capable of recognizing the chemical composition differences in a variety of citrus peel extracts, despite the presence of a single overwhelmingly concentrated component in each peel, limonene. The sensor can ignore this dominant component and shows sensitivity to other hydrophobic small molecules in the peel, despite their structural similarity to the major component and lack of any definable "recognition handles" other than simple aliphatic and hydroxy functional groups. In addition, the sensor is so effective that it can not only identify individual fruits using extracts taken from different fruit samples, but can also detect changes in peel composition based on ripeness of individual fruits. The array responses reflect the chemical differences in the peel extracts, with little influence from different biological batches and processing.

The authors thank the National Science Foundation (CHE-2305089 to R.J.H., CHE-2002413 to H.Z.) and the University of Science and Technology of China (to W.Z.) for support, as well as Mark Jones and Toshio G. Ishida for aid in varietal identification, and Prof. Min Xue for use of a microplate reader.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 *(a)* F. Hof, S. L. Craig, C. Nuckolls and J. Rebek, Jr. *Angew. Chem., Int. Ed.* 2002, **41**, 1488; *(b)* J. Rebek Jr, *Acc. Chem. Res.,* 2009, **42**, 1660.

2 (*a*) R. Pinalli, A. Pedrini and E. Dalcanale, *Chem. Soc. Rev.* 2018, **47**, 7006; *(b)* A. C. Sedgwick, J. T. Brewster, T. Wu, X. Feng, S. D. Bull, X. Qian, J. L. Sessler, T. D. James, E. V. Anslyn and X. Sun, *Chem. Soc. Rev.*, 2021, **50**, 9; *(c)* S. van Dun, C. Ottmann, L.-G. Milroy and L. Brunsveld, *J. Am. Chem. Soc.*, 2017, **139**, 13960; *(d)* J. Krämer, R. Kang, L. M. Grimm, L. De Cola, P. Picchetti and F. Biedermann, *Chem. Rev.*, 2022, **122**, 3459.

3 *(a)* J. H. Jordan and B. C. Gibb, *Chem. Soc. Rev.* 2015, **44**, 547; (*b*) Z. Laughrey and B. C. Gibb, *Chem. Soc. Rev.*, 2011, **40**, 363; (*c*) K. D. Zhang, D. Ajami, J. V. Gavette and J. Rebek Jr., *J. Am. Chem. Soc.*, 2014, **136**, 5264.

4 (*a*) M. A. Beatty and F. Hof, *Chem. Soc. Rev.*, 2021, **50**, 4812; (*b*) M. A. Beatty, A. J. Selinger, Y. Li, F. Hof, *J. Am. Chem. Soc.*, 2019, **141**, 16763; (*c*) C. Hu, L. Grimm, A. Prabodh, A. Baksi, A. Siennicka, P. A. Levkin, M. M. Kappes and F. Biedermann, *Chem. Sci.*, 2020, **11**, 11142; (*d*) Y. Liu, T. Minami, R. Nishiyabu, Z. Wang and P. Anzenbacher, *J. Am. Chem. Soc.*, 2013, **135**, 7705.

5 *(a)* E. E. Harrison, B. A. Carpenter, L. E. St. Louis, A. G. Mullins and M. L. Waters, *J. Am. Chem. Soc.*, 2021, **143**, 14845; (*b*) E. E. Harrison and M. L. Waters, *Chem. Sci.*, 2023, **14**, 928; (*c*) M. A. Beatty, J. Borges-González, N. J. Sinclair, A. T. Pye and F. Hof, *J. Am. Chem. Soc.* 2018, **140**, 3500.

6 *(a)* H. A. Henthorn and M. D. Pluth, *J. Am. Chem. Soc.* 2015, **137**, 15330; *(b)* T. S. Bailey, T. and M. D. Pluth, *J. Am. Chem. Soc.* 2013, **135**, 16697.

7 R. J. Hooley, H. J. Van Anda and J. Rebek, Jr., *J. Am. Chem. Soc.*, 2007, **129**, 13464.

8 *(a)* S. Liu, H. Gan, A. T. Hermann, S. W. Rick and B. C. Gibb, *Nat. Chem.*, 2010, **2**, 847; (*b*) M. R. Sullivan and B. C. Gibb, *Org. Biomol. Chem.*, 2015, **13**, 1869; (*c*) S. Liu, D. H. Russell, N. F. Zinnel and B. C. Gibb, *J. Am. Chem. Soc.*, 2013, **135**, 4314; (*d*) C. L. Gibb and B. C. Gibb, *Chem. Commun.*, 2007, **41**, 1635.

9 W. Zhong and R. J. Hooley, *Acc. Chem. Res.* 2022, **55**, 1035.

10 A. P. Umali, S. E. LeBoeuf, R. W. Newberry, S. Kim, L. Tran, W. A. Rome, T. Tian, D. Taing, J. Hong, M. Kwan, H. Heymann, E. V. Anslyn, *Chem. Sci.*, 2011, **2**, 439.

11 M. C. Gonzalez-Mas, J. L. Rambla, M. P. Lopez-Gresa, M. A. Blazquez, A. Granell, *Front. Plant Sci.* 2019, **10**, 12.

12 a) S. A. Siddiqui, M. J. Pahmeyer, E. Assadpourd, S. M. Jafari, *Ind. Crops Prod.* 2022, **177**, 114484; b) https://edu.rsc.org/experiments/extracting-limonene-from-oranges/692.

13 L. You, D. Zha and E. V. Anslyn, *Chem. Rev.* 2015, **115**, 7840.

14 *(a)* A. D. Gill, B. L. Hickey, W. Zhong and R. J. Hooley, *Chem. Commun.* 2020, **56**, 4352; *(b)* A. D. Gill, L. Perez, I. N. Q. Salinas, S. R. Byers, Y. Liu, B. L. Hickey, W. Zhong and R. J. Hooley, *Chem. – Eur. J.*, 2019, **25**, 1740; *(c)* B. L. Hickey, J. Chen, Y. Zou, A. D. Gill, W. Zhong, J. G. Millar and R. J. Hooley, *Chem. Commun.* 2021, **57**, 13341.

15 *(a)* R. J. Hooley, S. M. Biros and J. Rebek, Jr. *Chem. Commun.* 2006, **42**, 509; *(b)* R. J. Hooley, J. V. Gavette, D. Ajami and J. Rebek, Jr. *Chem. Sci.* 2014, **5**, 4382–4387; *(c)* L. Perez, M. Mettry, B. G. Caulkins, L. J. Mueller and R. J. Hooley, *Chem. Sci.* 2018, **9**, 1836. *(d)* C. J. Easley, M. Mettry, E. M. Moses, R. J. Hooley and C. J. Bardeen, *J. Phys. Chem. A.* 2018, **122**, 6578.

16 (*a*) C. Cortes and V. Vapnik, *Machine Learning* 1995, **20**, 273; (*b*) O. Ivanciuc, *Rev. Comput. Chem.* 2007, **23**, 291.