Enantioselective Sensing of Insect Pheromones in Water

Briana L. Hickey,^{*a*} Junyi Chen,^{*b*} Yunfan Zou,^{*c*} Adam D. Gill,^{*d*} Wenwan Zhong,^{*a,b*} Jocelyn G. Millar^{*a,c*} and Richard J. Hooley^{*a,d*}*

^aDepartment of Chemistry; ^bEnvironmental Toxicology Graduate Program;^cDepartment of Entomology; ^dDepartment of Biochemistry University of California-Riverside, Riverside, CA 92521, U.S.A.

richard.hooley@ucr.edu

Electronic Supplementary Information

Table of Contents

General Information	S-2
New Molecule Synthesis and Characterization	S-3
Spectroscopic Analysis of Guest Properties	S-8
Fluorescence Spectra	<i>S</i> -8
NMR Spectra of Pheromone Interaction with Cavitands	S-9
Fluorescence Data for Pheromone Additions	S-13
Statistical Analysis	S-17
Full Combined Array	<i>S</i> -18
Optimized and Unoptimized Arrays	S-19
Enantioselectivity	S-21
Controls	S-22
References	S-24

General Information

Cavitands TCC¹, AMI², and CHI² as well as fluorescent guest DSMI³ were synthesized and characterized according to literature procedures. For detailed analysis of the fluorescence response of DSMI in hosts TCC, AMI, and CHI² please see the cited references. (*S*)-Fuscumol (*S*-1) and (*R*)-fuscumol (*R*-1) were prepared by kinetic resolution of racemic fuscumol.⁴ (2*R*,3*S*)-2,3-Octanediol (2*R*,3*S*-3) and (2*S*,3*R*)-2,3-Octanediol (2*S*,3*R*-3) were prepared by kinetic resolution of racemic *anti*-2,3-octanediol.⁵ Deuterated NMR solvents were obtained from Cambridge Isotope Laboratories (Andover, MA), and used without further purification. All other materials were purchased from Sigma Aldrich (St. Louis, MO) or Fisher Scientific (Fairlawn, NJ), and were used as received. NMR spectra were recorded on a Bruker Avance Neo 400 MHz NMR spectrometer. All NMR spectra were processed using MestReNova by Mestrelab. Research S.L. Purity of pheromones was determined by GC on a chiral stationary phase Cyclodex B column (30 m × 0.25 mm × 0.25 micron film thickness, J&W Scientific, Folsom CA), programmed from 50 °C/1 min, then 3 °C/min to 220 °C. Fluorescence measurements were performed with a Bio-Tek Synergy HT Multi-Detection Microplate Reader.

Fluorescence measurements. In general, the fluorescence assays were carried out by mixing 10 μ L of the fluorescent guest **DSMI** or **SMITE** (30 μ M in water), 10 μ L of cavitands **TCC**, **AMI**, or **CHI** (200 μ M in water), 10 μ L metal salts (500 μ M in water), and 10 μ L of the pheromone guest at 500 μ M (in water), then adding 60 μ L of the incubation buffer (Tris buffer HCl, pH 7.4, 20 mM) to bring the total volume up to 100 μ L for each well in the 96-well plate, then incubating with mild agitation for 15 min at room temperature. Each experimental condition was repeated in quadruplicate across four separate wells of the 96-well plate using identical sensor components, simultaneously collecting fluorescence signals for each target at one time. The fluorescence signal (F) was recorded with the Ex/Em wavelengths at 400/528 nm for guest **SMITE** and 485/600 nm for guest **DSMI**.

Data analysis. PCA. The quadruplicate raw fluorescence data sets were subjected to prcomp or princomp functions for Principal Component Analysis (PCA) performed with RStudio (Version 1.0.136), an integrated development environment (IDE) for R (version 3.3.2). 2D scores plots and confidence intervals, as well as biplots were graphed in RStudio using the packages ggplot2, ggpubr, ggfortify, devtools, and factoextra. All other fluorescence data charts were created in Microsoft Excel, with values representing the mean of the quadruplicate responses and error bars indicating their standard deviation.

LDA. The discrimination between pheromone enantiomers was achieved by using the supervised Linear Discriminant Analysis (LDA) in Python 3.9. The raw fluorescence data sets of two pheromone

enantiomers were subjected to StandardScaler for standardization, then use LinearDiscriminantAnalysis for classification, resulting in the transformed LD 1 scores. For each class of samples, the probability density of Student's t distribution and 95% confidence intervals of LD 1 values were generated and calculated by scipy.stats.t. The cross validation was performed using RepeatedStratifiedKFold(n_splits=4, n_repeats=10) with LinearDiscriminantAnalysis as the classification estimator.

SVM-RFE. Feature selection was performed with Python 3.9, using StandardScaler for data standardization, Recursive Feature Elimination (RFE) to choose the six most relevant features (n_features_to_select=6), Support Vector Machine (SVM) svm.SVC(kernel='linear') as the classification estimator, and RepeatedStratifiedK-Fold (n_splits=4, n_repeats=3) for cross validation.

New Molecule Synthesis and Characterization

(E)-1-methyl-4-(4-(methylthio)styryl)pyridin-1-ium iodide (SMITE):

1,4-Dimethylpyridinium iodide (235 mg, 1.0 mmol) and 4-(methylthio)benzaldehyde (152 mg, 1.0 mmol) were dissolved in ethanol (5 mL) in a round bottom flask. While stirring, one drop of piperidine was added and the resulting solution was refluxed for 12 hr. The reaction was cooled, then diluted with water (10 mL). The resulting precipitate was filtered, rinsed with water and cold ethanol, then dried under vacuum to yield (*E*)-1-methyl-4-(4-(methylthio)styryl)pyridin-1-ium iodide (340 mg, 92% yield) as a dark yellow powder. ¹H NMR (400 MHz, DMSO) δ 8.83 (d, J = 6.8 Hz, 1H), 8.18 (d, J = 6.9 Hz, 1H), 7.97 (d, J = 16.4 Hz, 1H), 7.69 (d, J = 8.5 Hz, 1H), 7.47 (d, J = 16.4 Hz, 1H), 7.36 (d, J = 8.5 Hz, 1H), 4.24 (s, 3H), 2.54 (s, 3H).¹³C NMR (100 MHz, DMSO) δ 153.03, 145.47, 142.33, 140.67, 131.95, 129.05, 126.20, 123.77, 122.60, 47.34, 14.64. ESI-MS: *m*/*z* C₁₅H₁₆NS⁺ calculated: 242.0998, found: (M)⁺ 242.1001. UV/Vis: Exc. $\lambda_{max} = 385$ nm, Em. $\lambda_{max} = 545$ nm.



Figure S-1. ¹H NMR spectrum of **SMITE** (400 MHz, DMSO-*d*₆, 298K).



Figure S-2. ¹³C NMR of **SMITE** (100 MHz, DMSO-*d*₆, 298K).



Figure S-3. ESI-MS spectrum of SMITE in MeOH.

Synthesis of Pheromone Targets



Scheme S-1. Pheromone targets used in this study.

Octanediols *S*,*S*-4, and *R*,*R*-4:

(2S,3S)-2,3-Octanediol (*S*,*S*-4) and (2R,3R)-2,3-octanediol (*R*,*R*-4) were prepared as shown in Scheme S1, following a similar route to the stereospecific syntheses of (2R,3R)- and (2S,3S)-2,3-hexanediols from (L)- and (D)-threonine.⁶



Scheme S-2. Synthetic route for pheromone targets (2S, 3S)-2,3-octanediol (S,S-4) and (2R, 3R)-2,3-octanediol (R,R-4).

(4*S*,5*S*)-2,2,4-Trimethyl-5-pentyl-[1,3]dioxolane (S-6):

(4S,5S)-4-hydroxymethyl-2,2,5-trimethyl-[1,3]-dioxolane S-4 was prepared in 4 steps from (D)threonine as described in previous literature.⁷ Briefly, (D)-threonine was converted to the corresponding diol S-1 by treatment with aqueous NaNO₂ and sulfuric acid, followed by sequential methylation of the carboxylic acid to give ester S-2, and ketalization of the vicinal diol with dimethoxypropane, yielding ester S-3. Reduction of ester S-3 then gave (4S,5S)-4-hydroxymethyl-2,2,5-trimethyl-[1,3]-dioxolane S-4. Pyridine (1.6 mL, 20 mmol) and triflic anhydride (4.0 mL, 24 mmol) were added sequentially to a cold (-15 °C), stirred solution of (4S,5S)-4-hydroxymethyl-2,2,5-trimethyl-[1,3]-dioxolane S-4 (2.92 g, 20 mmol) in dry CH2Cl2 (80 mL). The reaction was warmed to 0 °C and stirred for 1 h, then diluted with hexane (160 mL) and filtered through a pad of Celite. The filtrate was concentrated in vacuo to afford crude triflate S-5, which was used immediately in the next step without further purification. n-BuMgCl (20 mL, 2 M in THF, 40 mmol) was added to a suspension of CuBr Me₂S (0.82 g, 4.0 mmol) in Et₂O (80 mL) at 0 °C followed by triflate S-5 in Et₂O (40 mL). The mixture was stirred for 2.5 h while warming to room temperature, then poured into a solution of saturated aqueous NH₄Cl and NH₃'H₂O (9:1) and extracted with Et₂O. The ether extract was washed with water and brine, then dried over anhydrous Na₂SO₄. The crude product was purified by flash chromatography (hexane/EtOAc = 40/1) to give S-6 as a colorless liquid (2.43 g, 65%). The ¹H NMR spectrum matched that previously reported.^{8,9}

(2*S*,3*S*)-2,3-Octanediol (*S*,*S*-4): Ketal S-6 (2.43 g, 13 mmol) was heated in a mixture of aqueous HCl (6 M, 10 mL) and MeOH (10 mL) at 60 °C overnight. The reaction mixture was poured into aqueous K₂CO₃ and extracted with EtOAc. The combined organic layers were washed with brine and dried over anhydrous Na₂SO₄. The crude product was purified by vacuum flash chromatography (hexane/EtOAc = 2/1, then EtOAc) to give *S*,*S*-4 as a colorless liquid (1.67 g, 87%). The ¹H NMR spectrum matched that previously reported.^{10,11}

(2*R*,3*R*)-2,3-Octanediol (*R*,*R*-4): This compound was prepared from (L)-threonine using the same series of reactions as described above for synthesis of its enantiomer, diol *S*,*S*-4.

Enantiopurity Determination:



Figure S-4. GC trace of *S*,*S*-4 on a chiral stationary phase Cyclodex B column. *S*,*S*-4 eluted at 29.61 min with an e.e. of 98.5%.



Figure S-5. GC trace of *R*,*R*-4 on a chiral stationary phase Cyclodex B column. *R*,*R*-4 eluted at 29.88 min with an e.e. of 95.9%.

Spectroscopic Analysis of Guest Properties



Dye Fluorescence Spectra

Figure S-6. Fluorescence spectra of **DSMI** and **SMITE** with **TCC/AMI/CHI**. a) [**DSMI**] = 0.625 μ M, [**Host**] = 0.625 μ M¹²; b) [**SMITE**] = 0.625 μ M, [**Host**] = 16 μ M, in 10mM K₂HPO₄/KH₂PO₄, 1mM EDTA buffer, pH 7.4.

Fluorescence Titrations of Dyes and Hosts



Figure S-7. Affinity measurement of Dye: **DSMI/SMITE** with Host: **TCC/AMI/CHI** via fluorescence. a) **DSMI** + Host **TCC/AMI/CHI**, [**DSMI**] = 1.5 μ M, in 20 mM Tris buffer, pH 7.4, Ex/Em = 485nm/605nm; b) **SMITE** + Host **TCC/AMI/CHI**, [**SMITE**] = 1.5 μ M, in 20 mM Tris buffer, pH 7.4, Ex/Em = 400nm/528nm.

NMR Analysis of Pheromone Interaction with Cavitands

AMI + 2-heptanol:



Figure S-8. ¹H NMR spectra (400 MHz, D₂O, 298K) showing rapid in and out exchange of guest **2** with **AMI** where a) an upfield shift of the methyl group of guest **2** can be seen in the aliphatic region with addition of **2** to **AMI**; b) full spectra for the addition of **2** to **AMI**.





Figure S-9. ¹H NMR spectra (400 MHz, $D_2O @ 298K$) showing rapid in and out exchange of guest 2 with **CHI** where a) the aromatic peaks of the host **CHI** become sharper with addition of guest 2 and b) an upfield shift of the methyl group of guest 2 can be seen in the aliphatic region; c) full spectra for the addition of 2 to **CHI**.

AMI + octanediol 4



Figure S-10. ¹H NMR spectra (400 MHz, D₂O @ 298K) showing rapid in and out exchange of guest **4** with **AMI** where a) an upfield shift of the methine proton of guest **4** can be seen in the aliphatic region with addition of **4** to **AMI**; b) full spectra for the addition of **4** to **AMI**.

CHI + octanediol 4



Figure S-11. ¹H NMR spectra (400 MHz, D₂O @ 298K) showing rapid in and out exchange of guest **4** with **CHI** where a) the aromatic peaks of the host **CHI** become sharper with addition of guest **4** and b) an upfield shift of the methine proton of guest **4** can be seen in the aliphatic region; c) full spectra for the addition of **4** to **CHI**.

Fluorescence Data for Pheromone Additions



Target additions to TCC•DSMI•Metal Combinations:

Target additions to AMI•DSMI•Metal Combinations:





Target additions to CHI•DSMI•Metal Combinations:

Figure S-12. Relative fluorescence responses of the **Host-DSMI-M⁺-Pheromone** complex in 20 mM Tris buffer, pH 7.4. [Host] = 20 μ M, [DSMI] = 3.0 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M. F₀ = fluorescence response of the Host-DSMI-M⁺ complex, F = fluorescence response of the Host-DSMI-M⁺ complex, F = fluorescence response of the Host-DSMI-M⁺.



Target additions to TCC•SMITE•Metal Combinations:

Target additions to AMI•SMITE•Metal Combinations:





Target additions to CHI•SMITE•Metal Combinations:

Figure S-13. Relative fluorescence responses of the Host•SMITE•M⁺•Pheromone complex in 20 mM Tris buffer, pH 7.4. [Host] = 20 μ M, [SMITE] = 3.0 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M. F₀ = fluorescence response of the Host•SMITE•M⁺ complex, F = fluorescence response of the Host•SMITE•M⁺ complex.

Statistical Analysis

Machine Learning Output Tables

Table S-1. Tables showing the SVM-RFE ranking of all 24-components used for statistical analysis with the top six components highlighted (top); Performance metrics of 3 repeated 4-fold cross validation with SVM as the estimator by using the 6 best features selected by SVM-RFE (bottom).

Component	Rank	Component	Rank
TCC*DSMI*noM+	5	CHI*SMITE*noM ⁺	1
TCC*DSMI*La ³⁺	10	CHI*SMITE*La ³⁺	2
TCC*DSMI*UO2+	6	CHI*SMITE*UO2+	11
TCC*DSMI*Ce ³⁺	12	CHI*SMITE*Ce ³⁺	8
TCC*SMITE*noM+	13	AMI*DSMI*noM⁺	1
TCC*SMITE*La ³⁺	19 AMI*DSMI*La ³⁺		1
TCC*SMITE*UO2+	14	AMI*DSMI*UO2+	17
TCC*SMITE*Ce ³⁺	18	AMI*DSMI*Ce ³⁺	1
CHI*DSMI*noM⁺	9	AMI*SMITE*noM+	4
CHI*DSMI*La ³⁺	15	AMI*SMITE*La ³⁺	3
CHI*DSMI*UO2+	16	AMI*SMITE*UO2+	7
CHI*DSMI*Ce ³⁺	1	AMI*SMITE*Ce ³⁺	1

Evaluation Metrics	Score (standard deviation from 3 repeated runnings of the 4-fold cross validation)			
Accuracy	1.0000 (0.0000)			
Sensitivity	1.0000 (0.0000)			
Specificity	1.0000 (0.0000)			
Precision	1.0000 (0.0000)			
F1 Score	1.0000 (0.0000)			
AUC	1.0000 (0.0000)			

Principal Component Analysis Plots to Distinguish Pheromones 1-5

PCA Scores Plot using the Full 24-component Combined Array



Figure S-14. PCA scores plots with 95% confidence intervals for the full 24-factor **Host•DSMI•M**⁺ and **Host•SMITE•M**⁺ arrays in Tris buffer (obtained from statistical analysis of data in Figures S-10 and S-11). [Host] = 20 μ M, [DSMI] and [SMITE] = 3 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M, [Tris] = 20 mM (pH 7.4).



Figure S-15. PCA biplot (combining both PCA scores plot and loading plot) using the full 24-component array system from Figure S-14 with prcomp (x,center = TRUE, scale. = TRUE) as the PCA function. Loadings are gradient-colored according to the contribution of each variable. Ellipses indicate 95% confidence intervals.





Figure S-16. PCA scores plots with 95% confidence intervals for an optimized a) 4-factor array with **Host-DSMI** or **Host-SMITE** and either no metal or Ce³⁺ in Tris buffer and b) a 6-factor array with **Host-DSMI** or **Host-SMITE** and either no metal, La³⁺, or Ce³⁺ (obtained from statistical analysis of selected data from Figures S-12 and S-13). [**Host**] = 20 μ M, [**DSMI**] and [**SMITE**]= 3 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M, [Tris] = 20 mM (pH 7.4).



Figure S-17. PCA biplot (combining both PCA scores plot and loading plot) using the 6-factor optimized array system from Figure S-16b with princomp(x,cor=TRUE, scores = TRUE) as the PCA function. Loadings are gradient-colored according to the contribution of each variable. Ellipses indicate 95% confidence intervals.



PCA Scores Plot using the Unoptimized Arrays

Figure S-18. PCA scores plots with 95% confidence intervals for unoptimized arrays with **Host•DSMI** or **Host•SMITE** and either La³⁺, UO₂²⁺, or Ce³⁺ in Tris buffer using a) a 12-factor array and b) a 6-factor array (obtained from statistical analysis of selected data from Figures S-12 and S-13). [**Host**] = 20 μ M, [**DSMI**] and [**SMITE**]= 3 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M, [Tris] = 20 mM (pH 7.4).



Figure S-19. PCA biplot (combining both PCA scores plot and loading plot) using the 6-factor unoptimized array system from Figure S-18b with princomp(x,cor=TRUE, scores = TRUE) as the PCA function. Loadings are gradient-colored according to the contribution of each variable. Ellipses indicate 95% confidence intervals.

Linear Discriminant Analysis Plots to Distinguish Enantiomers

Cross Validation Scores of Pheromone Enantiomer Classification

Sample	R/S-2-heptanol	2R,3R-/2S,3S- octanediol	R/S-1- phenylethanol	R/S-fuscumol
Accuracy	1.0000 (0.0000)	0.8750 (0.2905)	0.9125 (0.1900)	0.9875 (0.0781)
Sensitivity	1.0000 (0.0000)	0.8750 (0.2905)	0.9125 (0.1900)	0.9875 (0.0781)
Specificity	1.0000 (0.0000)	0.8750 (0.2905)	0.9125 (0.1900)	0.9875 (0.0781)
Precision	1.0000 (0.0000)	0.8500 (0.3298)	0.8688 (0.2850)	0.9812 (0.1171)
F1 Score	1.0000 (0.0000)	0.8583 (0.3152)	0.8833 (0.2533)	0.9833 (0.1041)
AUC	1.0000 (0.0000)	1.0000 (0.0000)	1.0000 (0.0000)	1.0000 (0.0000)

Table S-2. Performance metrics of pheromone enantiomers classification in Fig. 4 calculated by 10 repeated 4-fold cross validation with LDA as the estimator.

1D LDA Plots using other Chiral Additives:

a) (+)-β-methylglucopyranoside additive:







c) L-(+)-tartaric acid additive:



Figure S-20. 1D LDA (x-axis=LD 1) plots for the 6-factor **Host-DSMI-Additive** sensor arrays in Tris buffer with a) chiral additive (+)- β -methylglucopyranoside, b) chiral additive Eu(hfc)₃, and c) chiral additive L-(+)tartaric acid were each separately tested. [**Host**] = 20 μ M, [**DSMI**] = 3 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M, [Additive] = 50 μ M, [Tris] = 20 mM (pH 7.4). Red/blue dots = datapoints, curve = probability density of student's t-distribution, vertical markers = 95% confidence intervals.

Discriminant Analysis Plots with Non-cavitand Control Arrays:



Figure S-21. PCA scores plots with 95% confidence intervals for the control **DSMI**•**M**⁺ **Pheromone** and **SMITE**•**M**⁺ **Pheromone** sensor array in Tris buffer. [**DSMI**] and [**SMITE**] = 3 μ M, [Metal] = 50 μ M, [Pheromone] = 50 μ M, [Tris] = 20 mM (pH 7.4).

a) 2-Heptanol 2:

b) 2,3-octanediol 4:



Figure S-22. 1D LDA (x-axis=LD 1) plots for the control **DSMI-2-Additive** and **DSMI-4-Additive** sensor array in Tris buffer. [**DSMI**] = 3 mM, [Metal] = 50 mM, [Pheromone]= 50 mM, [Additive] = 50 μ M, [Tris] = 20 mM (pH 7.4). Red/blue dots = datapoints, curve = probability density of student's t-distribution, vertical markers = 95% confidence intervals.

References

- 1. S. M. Biros, E. C. Ullrich, F. Hof, L. Trembleau and J. Rebek Jr., J. Am. Chem. Soc. 2004, **126**, 2870-2876.
- A. D. Gill, B. L. Hickey, S. Wang, M. Xue, W. Zhong and R. J. Hooley, *Chem. Commun.* 2019, 55, 13259-13263.
- 3. Y. Liu, M. Mettry, A. D. Gill, L. Perez, W. Zhong and R. J. Hooley, *Anal. Chem.* 2017, **89**, 11113-11121.
- 4. G. P. Hughes, Y. Zou, J. G. Millar and M. D. Ginzel, *Can. Entomol.* 2013, **145**, 327-332.
- 5. J. D. Wickham, J. G. Millar, L. M. Hanks, Y. Zou, J. C. H. Wong, R. D. Harrison and Y. Chen, *Environ. Entomol.* 2016, **45**, 223-228.
- 6. R. F. Mitchell, D. T. Hughes, C. W. Luetje, J. G. Millar, F. Soriano-Agatón, L. M. Hanks and H. M. Robertson, *Insect Biochem. Mol. Biol.* 2012, **42**, 499-505.
- 7. P. Bianchi, G. Roda, S. Riva, B. Danieli, A. Zabelinskaja-Mackova and H. Griengl, *Tetrahedron* 2001, **57**, 2213-2220.
- 8. S. K. Kang, D. H. Lee and J. M. Lee, *Bull. Korean Chem. Soc.* 1990, **11**, 274-275.
- 9. C. Paolucci, C. Mazzini and A. Fava, J. Org. Chem. 1995, 60, 169-175.
- 10. R. Bel-Rhlid, A. Fauve and H. Veschambre, J. Org. Chem. 1989, 54, 3221-3223.
- D. R. Hall, A. Cork, S. J. Phythian, S. Chittamuru, B. K. Jayarama, M. G. Venkatesha, K. Sreedharan, P. K. Vinod Kumar, H. G. Seetharama and R. Naidu, *J. Chem. Ecol.* 2006, **32**, 195-219.
- 12. J. Chen, B. L. Hickey, L. Wang, J. Lee, A. D. Gill, A. Favero, R. Pinalli, E. Dalcanale, R. J. Hooley and W. Zhong, *Nat. Chem.* 2021, **13**, 488-495.