

UCLA

UCLA Previously Published Works

Title

Evolving neural network optimization of cholesteryl ester separation by reversed-phase HPLC

Permalink

<https://escholarship.org/uc/item/6th6x86v>

Journal

Analytical and Bioanalytical Chemistry, 397(6)

ISSN

1618-2650

Authors

Jansen, Michael A.
Kiwata, Jacqueline
Arceo, Jennifer
[et al.](#)

Publication Date

2010-07-01

DOI

10.1007/s00216-010-3778-5

Peer reviewed

Evolving neural network optimization of cholesteryl ester separation by reversed-phase HPLC

Michael A. Jansen · Jacqueline Kiwata ·
Jennifer Arceo · Kym F. Faull · Grady Hanrahan ·
Edith Porter

Received: 28 February 2010 / Revised: 11 April 2010 / Accepted: 22 April 2010 / Published online: 21 May 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Cholesteryl esters have antimicrobial activity and likely contribute to the innate immunity system. Improved separation techniques are needed to characterize these compounds. In this study, optimization of the reversed-phase high-performance liquid chromatography separation of six analyte standards (four cholesteryl esters plus cholesterol and tri-palmitin) was accomplished by modeling with an artificial neural network–genetic algorithm (ANN-GA) approach. A fractional factorial design was employed to examine the significance of four experimental factors: organic component in the mobile phase (ethanol and methanol), column temperature, and flow rate. Three separation parameters were then merged into geometric

means using Derringer’s desirability function and used as input sources for model training and testing. The use of genetic operators proved valuable for the determination of an effective neural network structure. Implementation of the optimized method resulted in complete separation of all six analytes, including the resolution of two previously co-eluting peaks. Model validation was performed with experimental responses in good agreement with model-predicted responses. Improved separation was also realized in a complex biological fluid, human milk. Thus, the first known use of ANN-GA modeling for improving the chromatographic separation of cholesteryl esters in biological fluids is presented and will likely prove valuable for future investigators involved in studying complex biological samples.

Electronic supplementary material The online version of this article (doi:10.1007/s00216-010-3778-5) contains supplementary material, which is available to authorized users.

M. A. Jansen · J. Kiwata · E. Porter (✉)
Department of Biological Sciences,
California State University Los Angeles,
Los Angeles, CA 90032, USA
e-mail: eporter@calstatela.edu

J. Arceo · G. Hanrahan (✉)
Department of Chemistry, California Lutheran University,
Thousand Oaks, CA 91360, USA
e-mail: ghanraha@clunet.edu

K. F. Faull
The Pasarow Mass Spectrometry Laboratory,
The NPI-Semel Institute for Neuroscience Human Behavior,
University of California Los Angeles,
Los Angeles, CA 90095, USA

E. Porter
School of Medicine,
University of California Los Angeles,
10833 LeConte Ave.,
Los Angeles, CA 90095, USA

Keywords Lipids · Cholesteryl linoleate · Innate immunity · Biological fluids · Artificial neural networks · Genetic algorithms

Introduction

All internal body surfaces are bathed with fluids that contain a variety of molecules, some of which are involved in innate immunity defense. Innate immunity forms the first line of defense against microbial invasion, and well-established mediators of this system include antimicrobial peptides [1]. Lipids are also ubiquitously present in body fluids [2], and there is increasing evidence that selected lipids have antimicrobial activity [3, 4] and form a lipid-mediated arm of the innate immunity response. We have shown that cholesteryl esters have antimicrobial properties and contribute to the innate immunity of secretions of the airway mucosa [5, 6]. Developing an optimized separation method in order to

characterize the role of individual lipids in innate immunity would be a significant advance to this field. Methods requiring minimal sample handling with high recovery for subsequent functional assays are essential. Current methods for cholesteryl ester analysis include HPTLC [7], TLC-GC [8], and ES-MS/MS [9, 10]. A high-performance liquid chromatography (HPLC) protocol for cholesteryl ester analysis has been described by Cullen et al. [11]. However, this protocol is suboptimal because triglycerides are first removed, preventing complete lipid analysis and thus adding additional steps. To meet our interest in characterizing antimicrobial lipids, we developed a reversed-phase HPLC (rpHPLC) one-step protocol for total lipid extracts that is suitable for both analytical and preparative scale work, but in this method, cholesteryl esters were incompletely resolved [5]. Because of the evidence of an antimicrobial role for cholesteryl esters [5, 6], optimization of the analytical protocol became of paramount importance.

The use of information processing techniques, in particular artificial neural networks (ANN), has proved valuable for a variety of separation methods [12–14]. When combined with experimental design techniques, ANN quickly optimizes the separation conditions and shortens analysis time, and does so without knowledge of the physical or chemical properties of the analytes [15]. In terms of applications, a limited number of studies have utilized ANN for optimizing HPLC experimental conditions [15–17]. Others have used a hybrid artificial neural network–genetic algorithm (ANN-GA) approach to improve separation methods for isolated compounds [18] and for routine pesticide analysis [19]. However, the incorporation of ANN-based methods for investigating the separation of complex human biological fluids is lagging, particularly with regards to lipids. Herein, we describe the first use of an ANN-GA approach for characterizing cholesteryl esters in human body secretions.

Experimental

Chemicals and reagents

HPLC grade acetonitrile, reagent grade alcohol, water, and chloroform were obtained from Fisher Scientific (New Jersey, USA). HPLC grade dichloromethane was obtained from EMD Chemicals Inc. (Darmstadt, Germany). Cholesterol, cholesteryl arachidonate, cholesteryl linoleate, cholesteryl palmitate, tri-palmitin (a highly hydrophobic triglyceride with similar elution times as cholesteryl esters in rpHPLC), cholesteryl stearate, and heptadecanoic acid (internal standard for milk lipid extraction) were obtained from Sigma-Aldrich (Missouri, USA).

Lipid standards

A mixture of cholesterol, cholesteryl arachidonate, cholesteryl linoleate, cholesteryl palmitate, tri-palmitin, and cholesteryl stearate, each at 1 mg/mL, was prepared in dichloromethane.

Lipid extraction from human milk

Human milk from three different donors was purchased from Mother's Milk Bank, Denver, CO. Lipids were extracted according to Bligh and Dyer [20] using 10 μ L aliquots for rpHPLC analysis and 20 μ L aliquots for ESI/MS/MS.

Chromatography

Under the control of Chromeleon[®] software (version 6.60 SP2), solutions of standards and extracts re-dissolved in dichloromethane were manually injected (1–3 μ L/injection, 20 μ L injection loop) onto a reversed-phase column (Dionex Acclaim PolarAdvantage II, a silica-based column with a proprietary amide-embedded ligand, 150 mm \times 2.1 mm ID, 3 μ m particle size) preceded by a guard column (Dionex Acclaim PolarAdvantage II, 2.1 \times 10 mm, 5- μ m particle size) in a temperature-controlled compartment (Dionex model TCC-100 column oven). The column had been previously equilibrated in the selected eluant and was eluted (Dionex P680 low-pressure quaternary pump with degasser) at specified flow rates. The eluant was passed through an evaporative light-scattering detector (ELSD, Alltech model 800). For fraction collection, the ELSD was bypassed and collected fractions were dried in a stream of nitrogen and stored at -20 °C for further analysis.

Mass spectral analysis

Cholesteryl esters in collected fractions were identified by ES/MS and MS/MS according to Duffin et al. [9] and as previously reported [5].

Experimental design

The influence of flow rate, ethanol and methanol content in the mobile phase (acetonitrile), and column temperature on cholesteryl ester separation was investigated with a fractional factorial design where the number of experiments of a full factorial design is reduced by a number p according to 2^{k-p} . The choice of mobile phase components and column temperature were based on preliminary observations. Moreover, the selection of flow rate was based on studies showing its effect on peak shape [21] and resolution [22]. The eight randomized experiments (run in triplicate for a

total of 24 runs) and acquired data are highlighted in Table 1. Data were analyzed in JMP (SAS Institute, USA) statistical software.

Derringer’s desirability function

Cholesteryl ester separation was evaluated for the total number of peaks, the retention time of the last eluting peak, and the sum of all peak resolutions, as described in Novotna et al. [15]. Each parameter was transformed into a desirability variable (*d*) using Derringer’s two-sided transformation [23]:

$$d_i = \begin{cases} \left[\frac{Y_i - Y_i^*}{c_i - Y_i^*} \right]^s & Y_i^* \leq Y_i \leq c_i \\ \left[\frac{Y_i - Y_i^*}{c_i - Y_i^*} \right]^t & c_i < Y_i \leq Y_i^* \\ 0 & Y_i < Y_i^* \text{ or } Y_i > Y_i^* \end{cases} \quad (1)$$

where Y_i is the measured response, Y_i^* the minimum acceptable value, Y_i^* the maximum acceptable value, and c_i is the target value of the measured response. Values of d_i range from 0 to 1, with 1 being the most desirable. A measured response below the minimum acceptable value or above the maximum acceptable value produces the unacceptable result $d_i=0$. Furthermore, the exponents s and t adjust the importance of the target value relative to the minimum and maximum values, respectively.

For the transformation of the total number of peaks, the target value was set at 6 to match the number of analytes in the standard mixture. The minimum acceptable value was set at 1, while the maximum acceptable value was set at 8 to accommodate lipid oxidation or degradation that might result in additional peaks. The exponents s and t were both set at 3 to assign more weight to values closer to the target value of 6. For the transformation of the retention time of the last eluting peak, a target value of 25 min was chosen to

accommodate future studies of biological fluids and high throughput analysis. The maximum value was set at 100 min, and s was set to 1 and t to 3 to match the preference for a shorter run.

Before the sum of all peak resolutions was transformed into a desirability value, resolution between consecutive peaks was first calculated as:

$$R = \frac{2(t_2 - t_1)}{w_1 + w_2} \quad (2)$$

where t and w indicate the retention time and associated baseline peak widths of the consecutive peaks, respectively. To arrive at a target value, resolutions between consecutive peaks were evaluated to allow for the adequate spacing of additional cholesteryl esters or other non-polar compounds like triglycerides [11]. The resolution between the first and second peaks, cholesterol and cholesteryl arachidonate, was set at 9 to avoid interference from diglycerides and phospholipids, which may elute in the same region [5]. The resolution between each of the remaining consecutive peaks was set at 4. All peak resolutions were summed, resulting in an overall target value of 25. The minimum acceptable value was set at 1, the maximum acceptable value was set at 28, s was set at 1, and t was set at 3 so that the desirability would increase moderately toward 25, but decline more rapidly thereafter. The individual desirabilities of the three parameters were subsequently merged into one single numeric value D (response) as follows:

$$D = (d_p \times d_t \times d_r)^{1/3} \quad (3)$$

where d_p is the transformation of number of peaks, d_t the transformation of the retention time of the last peak, and d_r the transformation of the sum of all peak resolutions.

Table 1 Summary of factors and responses for screening runs

Experiment	Factors ^a				Response (<i>D</i>)				
	Column T (°C)	Ethanol (%)	Methanol (%)	Flow rate (mL/min)	Run A	Run B	Run C	Means ± SD	Relative SD (%)
1	-1 (25°C)	-1 (25)	-1 (0)	-1 (0.15)	0.418	0.503	0.203	0.375±0.154	41.2
2	+1 (35°C)	-1 (25)	-1 (0)	+1 (0.45)	0.866	0.837	0.817	0.840±0.025	3.0
3	-1 (25°C)	+1 (75)	-1 (0)	+1 (0.45)	0.148	0.156	0.150	0.151±0.004	2.8
4	+1 (35°C)	+1 (75)	-1 (0)	-1 (0.15)	0.232	0.211	0.208	0.217±0.013	5.9
5	-1 (25°C)	-1 (25)	+1 (20)	+1 (0.45)	0.908	0.922	0.924	0.918±0.009	0.9
6	+1 (35°C)	-1 (25)	+1 (20)	-1 (0.15)	0.585	0.544	0.685	0.605±0.072	11.9
7	-1 (25°C)	+1 (75)	+1 (20)	-1 (0.15)	0.247	0.260	0.237	0.248±0.011	4.6
8	+1 (35°C)	+1 (75)	+1 (20)	+1 (0.45)	0.127	0.111	0.122	0.120±0.008	6.6

The response values from the individual runs (24 total) were used for ANN-GA analysis

^a-1 and +1 indicate low and high value for the respective factors

Artificial neural network–genetic algorithm approach

Figure 1 shows the ANN-GA strategy employing Levenberg–Marquardt backpropagation with a sigmoid transfer function. ANNs are computational modeling tools defined by structures comprising interconnected adaptive processing elements that perform parallel computations for data processing and interpretation [24]. A collection of processing units communicates by activation of neurons over a number of weighted connections [25]:

$$\xi_j^l = \sum_{i=1}^{N_{l-1}} w_{ij}^l x_i^{l-1} \tag{4}$$

where ξ_j is the net effect and w_{ij}^l is a connection weight. In a given layer l , a standard neuron, j , integrates the signals, x_j , impinging upon it and producing the net effect. Input and output layers are present and symbolize data inputs into the neural network and response of the network to the inputs, respectively.

Genetic optimization has proven beneficial in the determination of efficient neural network structure through a population of individuals, which evolves toward optimum solutions through defined genetic operators (selection, crossover, and mutation) [26, 27]. The GA used here operated according to a general two-step process: (1) initialization of the population and evaluation of each member of the initial population and (2) reproduction until a stopping condition was met. A population of random binary strings, each of

which represented a specific network architecture and set of training parameters, were defined. The fitness of each trained network was calculated according to:

$$\text{Fitness} = W_1 f_e(\hat{e}) + W_2 f_t(\hat{t}) \tag{5}$$

where $f_e(\hat{e})$ is the error between real output and ANN output, $f_t(\hat{t})$ is the training time of ANN, and W_1 and W_2 are appropriate values of weight. According to the observed fitness, the GA method selected a new group of strings, ostensibly representing the parents of the next generation with an assigned probability of reproduction.

For this study, a two-point crossover approach (one that calls for two points to be selected on the parent organism string) was employed. Mutation ensured that the probability of searching regions in the experimental space was never zero and prevented loss of genetic material through reproduction and crossover. The above process was repeated until the best string that gave the maximum fitness or minimum mean square error (MSE) was chosen. The MSE was calculated as:

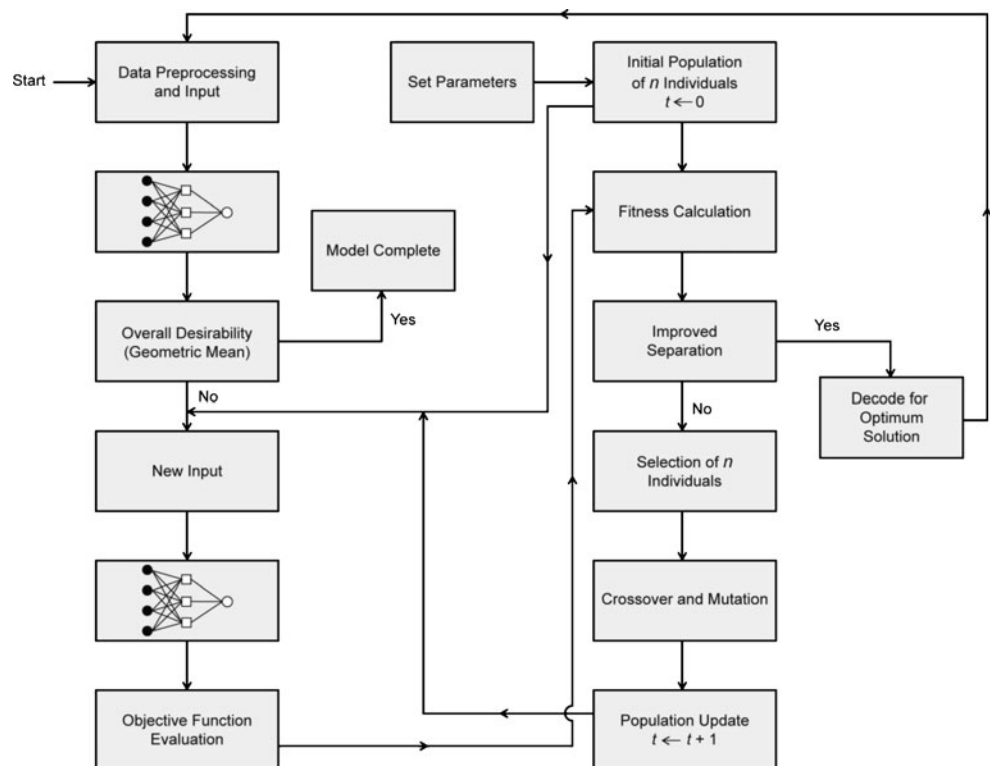
$$\text{MSE} = \frac{\text{SSE}}{n} \tag{6}$$

where SSE is the sum of squares error given by:

$$\text{SSE} = \text{Sum}_{(i=1 \text{ to } n)} \{w_i (y_i - f_i)^2\} \tag{7}$$

and y_i is the observed data value, f_i the predicted value from the fit, and w_i the weighting applied to each data point,

Fig. 1 Schematic of the hybrid artificial neural network–genetic algorithm (ANN-GA) method



usually $w_i=1$. For objective appreciation, synaptic weights and bias were set at 0.01 and 0.1, respectively. GA configuration values included a maximum generation of 100, a crossover probability of 0.5, and a mutation probability of 0.01. All data were cross-checked to ensure accuracy and validity and analyzed via JMP (SAS Institute, USA) and MATLAB 7.8 (The Math Works, USA).

Results and discussion

Experimental design

From the experimental design matrix (Table 1), eight experiments (performed in triplicate) were analyzed in randomized order with the mean response and standard error reported. Factorial ANOVA results confirmed good agreement between measured and predicted values, with the summary of fit reported at $r^2=0.96$. The general ANOVA calculations for the linear model are presented in Table 2. $\text{Prob} > F$ represents the significance probability for the F ratio, which states that if the null hypothesis is true, a larger F statistic would only occur due to random error. Significance probabilities of ≤ 0.05 are evidence of at least one significant regression factor in the model. Close examination of $\text{Prob} > F$ revealed that percent ethanol has the greatest significant single effect (<0.001) on the model response, D . Ethanol also exhibited a significant interactive effect when combined with flow rate (<0.001).

ANN-GA prediction and validation

Transformed data obtained from the fractional factorial design was used as the study data set for ANN-GA modeling. This data set was randomly divided with a 90% split between the training (70%) and test sets (30%), and the additional 10% set aside for validation. The training data set was used for model fitting in computing the gradient and updating network weights and biases. The optimal ANN architecture (Fig. 2), realized in 18 generations, included three hidden neurons and a MSE of 0.112. Typically, the smaller the MSE value, the smaller is the difference between the predicted results and the actual data, and thus the better the prediction performance was realized

Table 2 ANOVA calculations for the linear model

Source	<i>df</i>	Sum of squares	Mean square	<i>F</i> ratio
Model	7	2.082	0.2974	60.94
Error	16	0.078	0.0048	$\text{Prob} > F$
C. Total	23	2.160		<0.0001

df degrees of freedom, *C.Total* corrected total

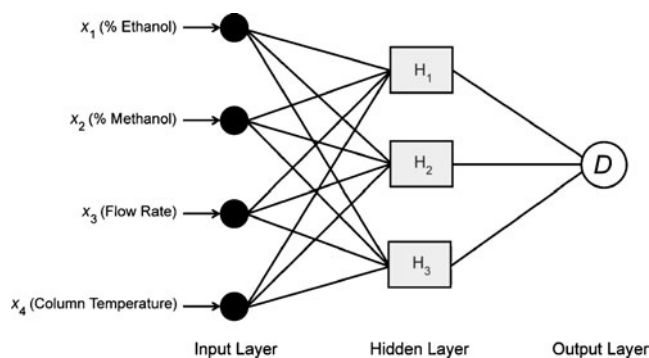


Fig. 2 Schematic of the optimized ANN architecture realized in this study

for the neural model. At this point, the network had the best generalization ability. The test set was then employed to check and verify the trained neural network. A plot of experimental responses versus peak responses predicted by ANN-GA for training and verification of the optimum architecture, which exhibited excellent model fitting ($r^2=0.97$, $n=10$) is presented in Fig. 3. Poor agreement would be indicative of overlearning. However, the level of correlation exhibited in this study confirms that the optimal neural network selected adequately modeled the response (D) for the six analyte standards.

From the data patterned by the ANN-GA model, a response surface was generated for the two interactive factors (percent ethanol and flow rate), with the largest value on the response surface corresponding to the greatest distribution of associated peaks. Software-generated ANN simulation runs were obtained as a result of examining the response surface, with five validation conditions that resulted in maximum predicted response values chosen for experimental verification. A summary of the modeled factors and responses are presented in Table 3. As shown, successful validation was achieved, with experimental responses in good agreement

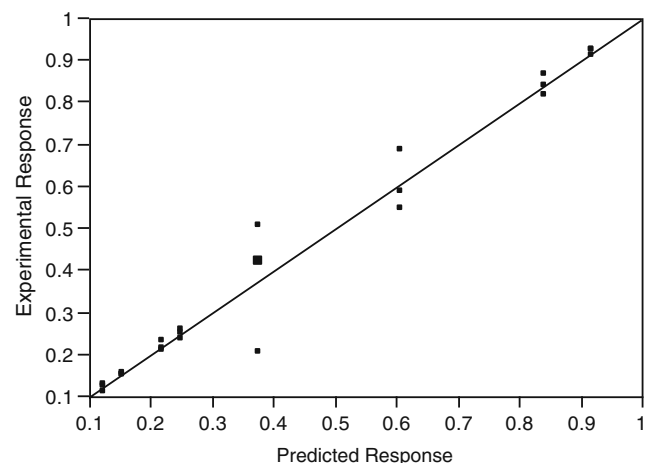


Fig. 3 Plot of experimental response versus the response predicted by the ANN-GA model for training and verification

Table 3 Summary of factors and responses for validation runs

Experiment	Factors ^a				Response (<i>D</i>)			
	Column T (°C)	Ethanol (%)	Methanol (%)	Flow rate (mL/min)	Mean ± SD (<i>n</i> =3)	Relative SD (%)	Predicted	Deviation ^b (%)
1	25.7	19.1	7.1	0.59	0.924±0.025	2.7	0.948	-2.4
2	31.7	19.0	17.1	0.45	0.889±0.036	4.1	0.947	-6.2
3	27.3	22.0	12.9	0.43	0.879±0.012	1.4	0.920	-4.4
4	30.2	22.2	18.0	0.42	0.912±0.020	2.1	0.908	0.4
5	29.9	23.9	2.0	0.56	0.915±0.030	3.3	0.901	1.5
6	30.0	20.0	10.0	0.55	0.896±0.007	0.8	0.937	-4.4

^a-1 and +1 indicate low and high value for the respective factors

^b $[(\text{measured response value} - \text{predicted response value})/\text{predicted response value}] \times 100$

(<6.2% discrepancy difference) with model-predicted responses. Modeled conditions from validation run 1 were used to compare peak separation before and after ANN-GA optimization as shown in the representative chromatographs presented in Fig. 4a, b, respectively. In Fig. 4b, six discernible and fully resolved peaks are noticeable, including the two previously co-eluting peaks, cholesteryl palmitate and tripalmitin. The average retention time of the last peak for this separation was 38.0 min. As this criteria was considered less critical, the trade-off between retention time and peak resolution was deemed acceptable.

Considering that a satisfactory peak separation was achieved with a limited number of HPLC runs, this study confirms the suitability of the ANN-GA model in method

optimization for rpHPLC. In contrast, manipulation of parameters via trial and error did not lead to a substantially improved separation of cholesteryl esters after more than 50 runs, and hence, this approach was abandoned (data not shown). Previously, ANN has been used successfully to separate structurally similar compounds, namely, indinavir and lactone, using HPLC [28], and complex mixtures of neuropeptides [15].

To assess whether the improved method (validation run 1) would be applicable to complex biological fluids, we subjected extracted human milk lipids to rpHPLC using the same experimental conditions dictated by the ANN-GA model. Choosing a lipid-rich fluid provided a robust test of the ANN-GA model. Since fatty acids and other lipids in human milk may interfere with cholesteryl ester separation [29], we included a water gradient (15% to 0% H₂O, 16.5% to 19.4% reagent alcohol with acetonitrile in the first 5 min of the separation) before the isocratic gradient used for validation run 1. This water gradient had previously been successful in separating free fatty acids in our laboratory (unpublished data).

Figure 5 shows representative chromatograms of three different lipid analyses from human milk before (Fig. 5a, isocratic gradient with 50% alcohol/50% acetonitrile) and after method optimization (Fig. 5b, gradient according to validation run 1). The numbers in the chromatogram indicate fractions that were collected in a subsequent run for further ESI/MS and MS/MS analysis. Even with this lipid-rich complex fluid, the optimized method designed for cholesteryl ester analysis produced an improved separation of non-polar lipids as reflected in peak shape and baseline separation. Specifically, 14 discernible peaks in the region of less polar and non-polar lipids of which only two returned to the baseline were recorded using the pre-optimization method, while 17 distinct peaks in this region of which seven returned to the baseline were recorded after ANN-GA optimization.

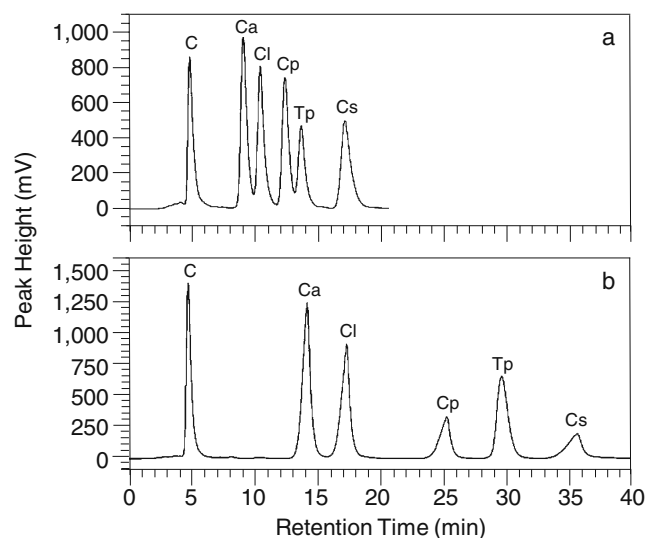


Fig. 4 Reversed-phase HPLC chromatograms of cholesteryl ester analysis before (a) and after (b) ANN-GA optimization employing validation run 1. Three micrograms each of cholesterol (C), cholesteryl arachidonate (Ca), cholesteryl linoleate (Cl), cholesteryl palmitate (Cp), tri-palmitin (Tp), and cholesteryl stearate (Cs) were injected

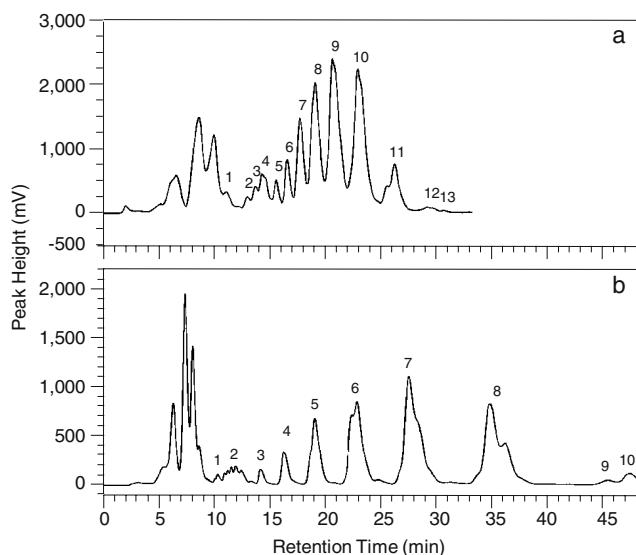


Fig. 5 Reversed-phase HPLC chromatograms of lipid extract from human milk before (a) and after (b) ANN-GA optimization employing validation run 1. Lipid extract from 10 μ L of human milk, spiked with 40 μ g heptadecanoic acid as internal standard, was injected. Fraction numbers that were further analyzed by ES/MS and MS/MS are indicated (see Table 4 and ESM Table S1)

ESI/MS and MS/MS analysis of the collected fractions revealed an overall improved separation of the four cholesteryl esters that had been employed for modeling (Table 4). A complete listing of masses and cholesteryl esters found in the collected fraction is given in Electronic supplementary materials (ESM) Table S1. The complexity of the identified cholesteryl esters is in agreement with previous reports on cholesteryl ester content in human milk [30], though we were able to identify additional cholesteryl esters not previously described, possibly due to our one-step separation approach. The former study employed preparative TLC prior to methylation and gas liquid chromatography to identify the constituent fatty acids of

milk cholesteryl esters. Carroll and Rudel [31] established an HPLC method in which cholesterol and ten cholesteryl esters were separated, of which seven lipids were partially and four were fully resolved. However, this method required two identical columns connected in series [32]. This effectively increased peak resolution by increasing column length. Further peak separation of complex biological fluids with our current one protocol would be also expected with increasing the column length. Others have achieved a combination of partial and complete separation of up to 14 cholesteryl esters and cholesterol, but pre-purification to remove triglycerides by hydrolysis or silica gel-60 column chromatography was required to remove interfering components [11, 33]. In contrast, here, we have shown an rpHPLC method that improves separation of cholesteryl esters in a complex biological fluid without a pre-purification step on a single column, which will accelerate studies on the biological role of individual cholesteryl esters in mucosal secretions.

Conclusion

This study provides guidance for the development and application of experimental design methodology and ANN-GA modeling tools for rapid optimization of the HPLC method used to separate cholesteryl esters. This approach optimally determined a set of conditions in which cholesteryl esters were fully separated in standard mixtures and better resolved in a complex biological fluid in a timely fashion. This method improvement will facilitate studies to unveil the biological function of cholesteryl esters in innate immunity. Moreover, as a general-purpose optimization approach, ANN-GA modeling will likely prove useful for a wide range of method optimizations concerning complex biological samples.

Table 4 Elution profile of selected cholesteryl esters extracted from human milk in rpHPLC before and after ANN-GA modeling

Cholesteryl ester	Before ANN-GA		After ANN-GA	
	Fraction	Relative intensity (%)	Fraction	Relative intensity (%)
Cholesteryl palmitate	2	1.7	2	1.25
	4	5.4	4	7.3
	8	12	7	5
	9	10	8	15
	11	12		
Cholesteryl stearate	9	12	9	11
	10	9		
Cholesteryl linoleate	7	100	2	1.4
			6	100
			10	1.2
Cholesteryl arachidonate	6	12.5	6	20
	7	10		

Individual fractions collected as indicated in Fig. 5 were subjected to ESI/MS and MS/MS analysis and the masses determined for NH_4^+ adducts were compared with reported masses for cholesteryl esters. Reported are the fraction(s) in which cholesteryl esters used for ANN-GA modeling eluted and their relative intensity among all masses found in this fraction. See ESM Table S1 data file for full cholesteryl ester profile

Acknowledgments This research was funded by 1 P20 MD001824 (EP) from the National Institutes of Health. GH acknowledges support from the John Stauffer Charitable Trust and the Swenson Summer Research Internship Program. The authors also acknowledge Katina Landon and Christine Markowitz for their help with initial model development exercises.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Mookherjee N, Rehaume LM, Hancock RE (2007) Cathelicidins and functional analogues as antiseptic molecules. *Expert Opin Ther Targets* 11:993–1004
- Strandvik B (2004) Fatty acid metabolism in cystic fibrosis. *N Engl J Med* 350:605–607
- Georgel P, Crozat K, Lauth X, Makrantonaki E, Seltmann H, Sovath S, Hoebe K, Du X, Rutschmann S, Jiang Z, Bigby T, Nizet V, Zouboulis CC, Beutler B (2005) A Toll-like receptor 2-responsive lipid effector pathway protects mammals against skin infections with Gram-positive bacteria. *Infect Immun* 73:4512–4521
- Tollin M, Bergsson G, Kai-Larsen Y, Lengqvist J, Sjovalld J, Griffiths W, Skuladottir GV, Haraldsson A, Jorvall H, Gudmundsson GH, Agerberth B (2005) Vernix caseosa as a multi-component defence system based on polypeptides, lipids and their interactions. *Cell Mol Life Sci* 62:2390–2399
- Do TQ, Moshkani S, Castillo P, Anunta S, Pogosyan A, Cheung A, Marbois B, Faull KF, Ernst W, Chiang SM, Fujii G, Clarke CF, Foster K, Porter E (2008) Lipids including cholesteryl linoleate and cholesteryl arachidonate contribute to the inherent antibacterial activity of human nasal fluid. *J Immunol* 181:4177–4187
- Lee JT, Jansen MA, Yilma AN, Nguyen A, Desharnais R, Porter E (2010) Antimicrobial lipids: novel innate defense molecules are elevated in sinus secretions of patients with chronic rhinosinusitis. *Am J Rhinol Allergy* 24:99–104
- Stahl J, Niedorf F, Kietzmann M (2009) Characterisation of epidermal lipid composition and skin morphology of animal skin *ex vivo*. *Eur J Pharm Biopharm* 72:310–316
- Sattler W, Reicher H, Ramos P, Panzenboeck U, Hayn M, Esterbauer H, Malle E, Kostner GM (1996) Preparation of fatty acid methyl esters from lipoprotein and macrophage lipid subclasses on thin-layer plates. *Lipids* 31:1302–1310
- Duffin K, Obukowicz M, Raz A, Shieh JJ (2000) Electrospray/tandem mass spectrometry for quantitative analysis of lipid remodeling in essential fatty acid deficient mice. *Anal Biochem* 279:179–188
- Liebisch G, Binder M, Schifferer R, Langmann T, Schulz B, Schmitz G (2006) High throughput quantification of cholesterol and cholesteryl ester by electrospray ionization tandem mass spectrometry (ESI-MS/MS). *Biochim Biophys Acta* 1761:121–128
- Cullen P, Fobker M, Tegelkamp K, Meyer K, Kannenberg F, Cignarella A, Benninghoven A, Assmann G (1997) An improved method for quantification of cholesterol and cholesteryl esters in human monocyte-derived macrophages by high performance liquid chromatography with identification of unassigned cholesteryl ester species by means of secondary ion mass spectrometry. *J Lipid Res* 38:401–409
- Jurado JM, Ballesteros O, Alcazar A, Pablos F, Martin MJ, Vilchez JL, Navalon A (2008) Differentiation of certified brands of origins of Spanish white wines by HS-SPME-GC and chemometrics. *Anal Bioanal Chem* 390:961–970
- Jinno K, Quiming NS, Denola NL, Saito Y (2009) Modeling of retention of adrenoreceptor agonists and antagonists on polar stationary phases in hydrophilic interaction chromatography: a review. *Anal Bioanal Chem* 393:137–153
- Riveros TA, Porcasi L, Muliadi S, Hanrahan G, Gomez FA (2009) Application of artificial neural networks in the prediction of product distribution in electrophoretically mediated microanalysis. *Electrophoresis* 30:2385–2389
- Novotna K, Havlis J, Havel J (2005) Optimisation of high performance liquid chromatography separation of neuroprotective peptides. Fractional experimental designs combined with artificial neural networks. *J Chromatogr A* 1096:50–57
- Webb R, Doble P, Dawson M (2009) Optimisation of HPLC gradient separations using artificial neural networks (ANNs): application to benzodiazepines in post-mortem samples. *J Chromatogr B Analyt Technol Biomed Life Sci* 877:615–620
- Petritis K, Kangas LJ, Yan B, Monroe ME, Strittmatter EF, Qian WJ, Adkins JN, Moore RJ, Xu Y, Lipton MS, Camp DG, Smith RD (2006) Improved peptide elution time prediction for reversed-phase liquid chromatography–MS by incorporating peptide sequence information. *Anal Chem* 78:5026–5039
- Tham SY, gatonic-Kustrin S (2002) Application of the artificial neural network in quantitative structure–gradient elution retention relationship of phenylthiocarbonyl amino acids derivatives. *J Pharm Biomed Anal* 28:581–590
- Carneiro RL, Braga JW, Bottoli CB, Poppi RJ (2007) Application of genetic algorithm for selection of variables for the BLLS method applied to determination of pesticides and metabolites in wine. *Anal Chim Acta* 595:51–58
- Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 37:911–917
- McCalley DV (2000) Effect of temperature and flow-rate on analysis of basic compounds in high-performance liquid chromatography using a reversed-phase column. *J Chromatogr A* 902:311–321
- Aboul-Enein HY, Ali I (2002) Optimization strategies for HPLC enantioseparation of racemic drugs using polysaccharides and macrocyclic glycopeptide antibiotic chiral stationary phases. *Farmaco* 57:513–529
- Derringer G, Suich R (1980) Simultaneous optimization of several response variables. *J Qual Technol* 12:214–219
- Massart DL, Despagne F (1998) Neural networks in multivariate calibration. *Analyst* 123:157R–178R
- Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. *J Microbiol Methods* 43:3–31
- Mitchell M (1999) An introduction to genetic algorithms. MIT Press, Cambridge
- Jalali-Heravi M, Kyani A (2007) Application of genetic algorithm-kernel partial least square as a novel nonlinear feature selection method: activity of carbonic anhydrase II inhibitors. *Eur J Med Chem* 42:649–659
- Jancic-Stojanovic B, Ivanovic D, Malenovic A, Medenica M (2009) Artificial neural networks in analysis of indinavir and its degradation products retention. *Talanta* 78:107–112
- Isaacs CE (2005) Human milk inactivates pathogens individually, additively, and synergistically. *J Nutr* 135:1286–1288
- Bitman J, Wood DL, Mehta NR, Hamosh P, Hamosh M (1986) Comparison of the cholesteryl ester composition of human milk from preterm and term mothers. *J Pediatr Gastroenterol Nutr* 5:780–786
- Carroll RM, Rudel LL (1981) Evaluation of a high-performance liquid chromatography method for isolation and quantitation of cholesterol and cholesteryl esters. *J Lipid Res* 22:359–363
- Dong MW (2006) Modern HPLC for practicing scientists. Wiley, Hoboken
- Cheng B, Kowal J (1994) Analysis of adrenal cholesteryl esters by reversed phase high performance liquid chromatography. *J Lipid Res* 35:1115–1121