Journal of Chromatography A 1778 (2026) 467001

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Retention modeling of oligonucleotides on an amide-based HILIC column: A descriptor-driven approach

Jonathan Maurer [a,b], Balasubrahmanyam Addepalli [c], Fabrice Gritti [c], Matthew A. Lauber [c], Davy Guillarme [a,b,*], Szabolcs Fekete [d]

[a] School of Pharmaceutical Sciences, University of Geneva, CMU-Rue Michel Servet 1, 1211 Geneva, Switzerland
[b] Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, CMU-Rue Michel Servet 1, 1211 Geneva, Switzerland
[c] Waters Corporation, Milford, USA
[d] Waters Corporation, Geneva, Switzerland

ARTICLE INFO

Keywords:
Retention modeling
HILIC
Oligonucleotides
Amide

ABSTRACT

Hydrophilic interaction chromatography (HILIC) has recently gained attention as a powerful tool for the analysis of nucleic acid–based therapeutics, offering high resolving power and excellent compatibility with mass spectrometry. However, the successful development of HILIC methods requires a clear understanding of the underlying retention mechanisms. Despite numerous studies, the respective contributions of hydrophilic partitioning, polar interactions, and ionic interactions to oligonucleotide retention remain poorly understood.

In this work, the role of hydrogen bonding, dipole-dipole interactions, solvophobic effects, and ionic interactions in governing oligonucleotide retention have been systematically investigated. A series of targeted design-of-experiments studies, performed on a representative panel of samples to investigate the impact of oligonucleotide structure on retention, provided a refined framework for interpreting HILIC separations of oligonucleotides. All experiments were performed on an amide-bonded stationary phase, and the proposed descriptor-based retention model applies specifically to this surface chemistry.

Our results demonstrate that solvophobic effects and hydrogen bonding are the dominant drivers of retention, whereas hydrophilic partitioning contributes negligibly under practical conditions. This mechanistic insight has important consequences for method development: introducing protic solvents into the mobile phase, conditions that would disrupt the water-rich layer, substantially increases selectivity and enables the resolution of closely related oligonucleotide species.

Altogether, these findings shift the conceptual basis of HILIC for nucleic acids from a hydrophilic partitioning/interaction model toward a retention mechanism best described as a solvophobic–interaction-dominated HILIC regime, in which protic solvents assist selectivity through H-bond competition. This new perspective provides a more accurate foundation for designing, optimizing, and interpreting HILIC methods for oligonucleotide analysis.

1. Introduction

Therapeutic oligonucleotides have gained a central place in clinical and pre-clinical pipelines, with hundreds of phases I to III products currently under evaluation worldwide [1]. This massive arrival of newly developed drugs has generated a significant demand for fast and efficient techniques capable of assessing critical quality attributes such as sequence identity, length distribution, and integrity, among others [2–4]. Liquid chromatography plays a major role in meeting industrial needs, with its high versatility, resolution power, and proven reproducibility [5].

Ion-pair reversed-phase liquid chromatography (IP-RPLC) is currently the most widely used and best-established chromatographic technique for the analytical separation of oligonucleotides, particularly for impurity profiling and LC–MS-compatible analyses [6–9]. However, ion-pairing reagents suppress signal in mass spectrometry and tend to

* Corresponding author at: Institute of Pharmaceutical Sciences of Western Switzerland (ISPSO), University of Geneva, Group of Analytical Pharmaceutical Chemistry, CMU - Rue Michel Servet, 1, 1206 Geneva, Switzerland.
E-mail address: davy.guillarme@unige.ch (D. Guillarme).

https://doi.org/10.1016/j.chroma.2026.467001
Received 11 March 2026; Received in revised form 15 April 2026; Accepted 16 April 2026
Available online 16 April 2026
0021-9673/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

contaminate chromatographic instruments due to nonspecific adsorption (especially strong ion-pairing reagents). Recently, hydrophilic interaction liquid chromatography (HILIC) has emerged as a promising complementary approach [10]. Without the need for ion-pairing reagents, HILIC offers a promising degree of resolution for large and small nucleic acids [11,12], suitable compatibility with mass spectrometry [13], and versatility regarding the separation of diastereomers [14,15].

Despite its increasing use for nucleic acid analysis, the fundamental retention mechanism of oligonucleotides in HILIC remains poorly understood. The classical HILIC model was largely established on small polar molecules. It is assumed that there is a combination of hydrophilic partitioning between an aprotic organic-rich mobile phase and a water-rich stationary layer, surface adsorption through polar interactions, and ionic interactions when analytes and stationary phase are charged [16,17]. These studies assume a substantial and stable water layer at the stationary phase surface, which governs retention through partitioning. However, growing evidence suggests that this framework does not fully describe the behavior of large, highly charged biomolecules. Recent quantitative work has shown that the relative contribution of hydrophilic partitioning, adsorption, and electrostatic interactions is strongly dependent on analyte structure and mobile-phase composition, and that partitioning can even become negligible under certain conditions [18–20]. Moreover, the majority of HILIC stationary phases have polar ligands. There is increasing evidence that polar compounds can have direct interaction with these groups [21].

For oligonucleotides and mRNA, which are large, flexible, and polyanionic molecules, the relevance of a thick water layer as the dominant retention driver becomes questionable. Typical elution compositions of nucleic acids in HILIC are between 50 and 30% organic (acetonitrile); that is, 50 to 70% water [12,14]. At this composition, it can be reasoned that an isolated water layer is probably non-existent [22]. It has been demonstrated that for bare silica HILIC there is little or no excess water when the water content increases from 50% to 70% [23,24]. Therefore, partitioning of large nucleic acids ranging from 20 Å (20 nt) to 200 Å (5000 nt) seems physically impossible and could not explain the retention. Moreover, retention is still observed in solvent systems where the existence of such a layer is intrinsically limited (e.g., MeOH-rich HILIC) [25]. These discrepancies suggest that additional interactions, including solvophobic effects and hydrogen bonding between nucleobases, protic solvents, and the stationary phase, may play a more central role than previously assumed. The complex retention mechanism of HILIC, which depends on stationary phase chemistry, mobile phase composition and analyte properties, makes the correct identification of the main factor influencing the retention and the overall selectivity of a method challenging [21].

The aim of this study was to deconvolute and quantify the individual contributions of hydrogen bonding, dipole-dipole interactions, solvophobic effects, and electrostatic interactions to oligonucleotide retention on amide-based HILIC column. By building a mechanistic model based on experimental observations, the goal was to demonstrate that oligonucleotide retention can be predicted from molecular descriptors such as sequence, and from mobile phase descriptors such as H-bond acidity and polarizability. This could change the way HILIC is tuned regarding mobile phase composition and thus increase chances for chromatographers to further improve their method performance.

2. Experimental section

2.1. Chemicals and samples

RNase-free water, ammonium acetate, and toluene, were purchased from Sigma-Aldrich (Buchs, Switzerland). LC-MS grade acetonitrile, isopropanol, and methanol were purchased from Thermo Fisher Scientific (Reinach, Switzerland). A Milli-Q™ water purification system from Millipore (Bedford, MA, USA) was used to prepare HPLC mobile phases. DNA LoBind tubes of 0.5 mL purchased from Eppendorf (Hamburg, Germany) were used to incubate the samples in an Incubator Thermomixer Confort from Eppendorf. Polypropylene 100 µL vial inserts were purchased from Macherey-Nagel (Duren, Germany).

Data treatment and statistical analysis were performed using Excel™ spreadsheet (version 16 in Excel 2021, Microsoft), Statistica™ software (version 14.1.0, Tibco) and DryLab™ software (version 4.4.21, Molnar Institute). Oligonucleotide samples described in Table 1 were purchased from Eurogentec (Seraing, Belgium). They were composed of different DNA sequences specifically selected to challenge the analytical method. Their secondary structures were assessed using The Vienna RNA Websuite, available at http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi [26].

2.2. Chromatographic conditions

The chromatographic measurements were performed on an ACQUITY™ UPLC™ H-class System (Waters, Milford, MA, USA), equipped with a 15 µl flow through needle injector, a quaternary solvent manager (QSM) equipped with a 100 µl mixing chamber, a column manager, and an ACQUITY Photodiode Array Detector. UV detection was set at a wavelength of 260 nm. The column used for the DoEs was a Waters ACQUITY Premier BEH™ Amide 130 Å, 1.7 µm, 50 × 2.1 mm Column. For the final optimization, it was a Waters ACQUITY Premier BEH Amide 130 Å, 1.7 µm, 150 × 2.1 mm Column. (The stationary phase materials provided in these columns were not batch tested for oligonucleotide separations, though historical manufacturing data was referenced to assure representative performance was observed. GTxResolve™ BEH Amide Columns are available with certification to produce high recovery, low tailing separations of 10 to 100-mer oligonucleotides.) Data acquisition and instrument control were performed by Empower™ 3 Software (Waters).

The mobile phases were composed of various proportions of water, acetonitrile (MeCN), methanol (MeOH), isopropanol (IPA), and ammonium acetate, combined with different additives. Fixed parameters were injector strong wash 10:90 MeCN:H_2O (v:v), weak wash 90:10 MeCN:H_2O (v:v), injection volume of 2 µL, injection solvent 80:20 MeCN:H_2O (v:v), and gradient from 20 to 60% water. All these conditions were tested in DoE described in later sections and supplementary information, to specifically assess the role of H-bonding, dipole-dipole interactions, solvophobic effect, and ionic interactions in the retention mechanism of oligonucleotides.

3. Methodology and theory

3.1. Identifying the most important method factors (Screening)

To identify the dominant physicochemical interactions driving oligonucleotide retention on amide-type HILIC column, we conducted a structured first-stage screening in which four mechanistic contributors were examined independently: (1) solvent–analyte hydrogen bonding, (2) direct disruption of analyte–surface hydrogen bonds, (3) dipole–dipole interaction competition, and (4) solvophobic rejection of the analyte from organic-rich eluents. Because these mechanisms cannot be isolated in a single solvent experiment, we employed protic modifiers, hydrogen-bond disruptors, dipolar aprotic additives, and changes in organic solvent type to probe each effect selectively.

Hydrogen bonding interactions were studied in two complementary ways. First, we varied the proportion of MeOH replacing MeCN as organic co-solvent. MeOH is a strong H-bond donor, and therefore its presence primarily modifies solvent–analyte H-bond competition. Second, urea was added at two concentration levels (20 and 80 mM) to disrupt analyte–stationary-phase hydrogen bonding. Together, MeOH and urea allow separation of solvent-mediated versus surface-mediated H-bond interactions.

Dipole–dipole interactions were examined by adding cyrene, which is a highly dipolar aprotic solvent that can compete for dipolar

Table 1
List of oligonucleotide samples used in this study, with their sequences and, where applicable, their PARCH descriptors. Special signs are used to indicate modifications highlighted in **bold**: * for phosphorothioate, <> for 2'O-methylation on sugar, *m-* for additional methyl group on base.

First DoE (screening)		
Name	Sequence	
Poly T - 20	5' TTT-TTT-TTT-TTT-TTT-TTT-TT 3'	
Poly A - 20	5' AAA-AAA-AAA-AAA-AAA-AAA-AA 3'	
Poly C - 20	5' CCC-CCC-CCC-CCC-CCC-CCC-CC 3'	
Linear 5A-5T-5G-5C	5' GTA-CAT-ATC-GGC-CCA-TTG-AG 3'	
Hairpin 5A-5T-5G-5C	5' AGT-CAT-GCG-TAC-ATG-CGA-CT 3'	
Linear 2A-8T-6G-4C	5' CCA-GTT-TTT-GAG-CCG-TGT-GT 3'	
Hairpin 2A-8T-6G-4C	5' CCA-GTT-TTT-TTG-GGC-ACT-GG 3'	
Second DoE (H-bond vs solvophobic)		
Name	Sequence	
Poly T - 20	5' TTT-TTT-TTT-TTT-TTT-TTT-TT 3'	
Poly A - 20	5' AAA-AAA-AAA-AAA-AAA-AAA-AA 3'	
Poly C - 20	5' CCC-CCC-CCC-CCC-CCC-CCC-CC 3'	
Linear 5A-5T-5G-5C	5' GTA-CAT-ATC-GGC-CCA-TTG-AG 3'	
Hairpin 5A-5T-5G-5C	5' AGT-CAT-GCG-TAC-ATG-CGA-CT 3'	
Linear 2A-8T-6G-4C	5' CCA-GTT-TTT-GAG-CCG-TGT-GT 3'	
Hairpin 2A-8T-6G-4C	5' CCA-GTT-TTT-TTG-GGC-ACT-GG 3'	
Ds 10A-10T-10G-10C	5' GTA-CAT-ATC-GGC-CCA-TTG-AG 3'	
	3' CAT-GTA-TAG-CCG-GGT-AAC-TC 5'	
Third DoE (test the model)		
Name	Sequence	Descriptor
Poly T - 20	5' TTT-TTT-TTT-TTT-TTT-TTT-TT 3'	76.4
Poly T - 40	5' TTT-TTT-TTT-TTT-TTT-TTT-TTT-TTT-TTT-TTT-TTT-TTT-TTT-T 3'	152.8
Linear 4A-5T-3G-8C	5' ACG-TCA-GCT-CAC-TTC-GTC-AC 3'	70.7
Hairpin 4A-5T-3G-8C	5' ATC-CGA-CCC-CAT-TCT-CGG-AT 3'	35.2
Linear 6A-2T-7G-5C	5' AGT-AGC-AGC-AGC-ATG-ACG-CG 3'	69.4
Hairpin 6A-2T-7G-5C	5' AGT-AGC-AAG-CGC-CGG-CTA-CT 3'	31.9
Volanesorsen PO	5' AGC-TTC-TTG-TCC-AGC-TTT-AT 3'	43.1
Fomivirsen PO	5′ GCG-TTT-GCT-CTT-CTT-CTT-GCG 3′	77.1
Fourth DoE (mixture injection)		
Name	Sequence	
Linear 6A-2T-7G-5C	5' AGT-AGC-AGC-AGC-ATG-ACG-CG 3'	
Hairpin 6A-2T-7G-5C	5' AGT-AGC-AAG-CGC-CGG-CTA-CT 3'	
Hairpin 4A-5T-3G-8C	5' ATC-CGA-CCC-CAT-TCT-CGG-AT 3'	
Poly T - 20	5' TTT-TTT-TTT-TTT-TTT-TTT-TT 3'	
Volanesorsen 3 *m*C	5' AG ***m*C**-TT ***m*C**-TTG-TCC-AG ***m*C**-TTT-AT 3'	

interaction sites without contributing H-bond donation. Cyrene concentration varied between 0.1 and 1.0% (v/v) in both mobile phases. Ammonium acetate concentration (25 and 75 mM) was included in the same block of experiments because ionic strength partially screens dipole interactions through modulation of the electrostatic environment. Note that, in this context, ammonium acetate does not function as a primary retention driver, but rather as a dipole-scrambling agent, making it a natural partner for cyrene within the same design block.

Solvophobic contributions were probed by replacing MeCN with IPA. Compared to MeOH, IPA is considerably more hydrophobic and therefore strengthens solvophobic rejection of highly hydrated oligonucleotides from the organic-rich bulk mobile phase. Although IPA is protic and can participate in hydrogen bonding, its lower H-bond donor acidity and strong solvent–solvent association result in comparatively weak solvation of highly hydrated polyanions; as a consequence, its dominant chromatographic effect for large oligonucleotides is solvophobic in nature, driven by preferential exclusion from the bulk mobile phase [18, 27,28].

For this reason, IPA was used as the main solvophobic factor in a dedicated branch of the screening.

Because MeOH and IPA act in different ways and cannot be mixed simultaneously in the organic line due to the need to isolate their independent contributions, the screening was implemented as two separate two-level fractional factorial designs. The first block ("MeOH-branch") contained MeOH as the co-solvent factor; the second block ("IPA-branch") contained IPA as the co-solvent factor. Each branch used a 2^{4-1} fractional factorial design (8 runs) comprising four factors: co-solvent identity (MeOH or IPA), cyrene concentration, urea concentration, and ammonium acetate concentration. The 2^{4-1} design provides a Resolution IV aliasing, which ensures that all main effects can be estimated even though selected two-factor interactions are confounded. This level of resolution is ideal for a mechanistic screening approach because the primary goal is to rank the effects rather than model curvature or higher-order coupling.

The factor signs in each block were mapped using the classical generator, such that the ammonium acetate level was determined from the product of the co-solvent, cyrene, and urea signs. This strategy places the ionic-strength factor into the role of the least impactful mechanism, which is consistent with our prior expectations and minimizes the total number of experimental runs. The resulting matrices for both the MeOH and IPA branches consisted of 8 feasible mobile-phase compositions each, using the same additive levels in mobile phases A and B, and run under identical chromatographic conditions (10 min linear gradient from 0–100% A, at $F = 0.6$ mL/min and at $T = 60$ °C). This two-branch screening strategy enables unambiguous identification of the strongest mechanistic contributors to retention while ensuring that each experiment corresponds to a physically meaningful mobile-phase composition. (Supplementary information Tables S1 and S2 list the applied conditions for the screening designs.)

3.2. Studying ternary co-solvent mixtures

Because MeOH emerged during screening as the primary mobile-phase component that competes for H-bonds with oligonucleotides, and since IPA produced the strongest solvophobic (but protic) signal, we treated the organic modifier line as a ternary mixture (MeOH/IPA/MeCN). This compositional viewpoint avoids treating MeCN as a passive remainder and instead enables quantitative estimation of the intrinsic contribution of each organic component and solvent interactions. Mixture designs are appropriate here because changing one organic component necessarily changes the other fractions. The Scheffé mixture polynomials produce interpretable intrinsic and interaction coefficients for the components. A simplex-lattice (degree 2) design has been applied, augmented with duplicate vertex and midpoint measurements (and centroid replicates), giving 14 runs.

For each mixture point, the organic line composition (x) was mapped into mobile phase B as follows: mobile phase B is fixed at 20% water and 80% organic. The organic portion was allocated according to the simplex coordinates, so for example the MeOH vertex corresponds to B = 20% water + 80% MeOH, while the centroid corresponds to B = 20% water + (80% × 0.3333) MeOH/IPA/MeCN each. Mobile Phase A was a buffered aqueous phase with the same ammonium acetate concentration as B (this maintains constant ionic strength during the gradient). All runs used an identical linear gradient profile (10 min, 0 → 100% B), at $F = 0.3$ mL/min and at $T = 30$ °C. The low temperature and reduced flowrate were set to avoid the denaturation of double stranded samples and overpressure issues. (Supplementary information Table S3 lists the applied conditions for the mixture design.)

3.3. Decoupling H-bond and solvophobic effects

To decouple retention mechanisms, solvent compositions were expressed not only in volumetric fractions, but also through physicochemical descriptors of the solvents. The following descriptors were considered: (1) H-bond donor acidity (α), via Kamlet–Taft parameter, reflecting solvent ability to donate H-bonds and (2) polarizability (π^*) parameter to estimate the solvophobic nature of the mobile phase. The descriptor values considered were: $\alpha = 0.98$ (MeOH), 0.19 (MeCN), 0.76 (IPA) and 1.17 (water), while $\pi^* = 0.60$ (MeOH), 0.75 (MeCN), 0.48 (IPA) and 1.09 (water) [29]. Alternative descriptors such as refractivity (R) or McGowan volume (Vx) were also considered, however, these primarily describe dispersion interactions and molecular size effects, which are known to play a secondary role in HILIC retention compared to polarity-driven solvation and hydrogen bonding interactions, particularly for large, highly hydrated polyanionic analytes such as oligonucleotides [16,19]. The Kamlet–Taft π^* parameter reflects solvent dipolarity/polarizability rather than molecular size, making it more appropriate for modeling solvophobic contributions.

For a given organic mobile phase composition, the mixture descriptor (S_{MP}) was calculated as a weighted average:

$$S_{MP} = f_w S_w + f_{org} \sum_i x_i S_i \tag{1}$$

where f_w is the water fraction and f_{org} is the organic fraction in the mobile phase, x_i is the fraction of solvent i in the organic mixture (MeOH, IPA, MeCN, with $\sum_i x_i = 1$), S_w is the descriptor value for water and S_i is the descriptor value for organic solvent i. Therefore, for descriptors α and π^*, the following expressions can be written for the organic rich mobile phase which included 20% water:

$$\alpha_{MP} = 0.2 \cdot 1.17 + 0.80(x_{MeOH} \cdot 0.98 + x_{MeCN} \cdot 0.19 + x_{IPA} \cdot 0.76) \tag{2}$$

$$\pi_{MP} = 0.2 \cdot 1.09 + 0.80(x_{MeOH} \cdot 0.60 + x_{MeCN} \cdot 0.75 + x_{IPA} \cdot 0.48) \tag{3}$$

The resulting α_{MP} and π_{MP} values represent effective, composition-dependent solvent descriptors obtained by linear averaging. Although non-additive solvent effects may occur, these parameters are intended to capture relative trends within the constrained solvent space explored here rather than exact microscopic solvation energies.

In addition to solvent (mobile phase) descriptors, solute specific descriptors can be considered too. The ONs can be characterized by various descriptors, such as the H-bonding potential of their bases or their nucleotide hydropathy. In this study we considered the sequence-averaged *PARCH* (Protocol for Assigning a Residue's Character on a Hydropathy index), developed to quantify nucleotide hydropathy, as an analyte-specific descriptor of solvation and solvophobic behavior [30]. *PARCH* values were calculated as weighted averages of backbone and base contributions, corrected for structural accessibility. Considering the resulting $PARCH_{seq}$ descriptors the following multivariate retention model has been proposed:

$$\ln k_{app} = \beta_0 + \beta_H \alpha_{MP} + \beta_S \pi_{MP} + \beta_P PARCH_{seq} + \varepsilon \tag{4}$$

where k_{app} is the apparent retention factor, β_0 represents the system-dependent baseline retention, β_H is regression coefficient for solvent H-bond donor acidity, β_S is coefficient for solvophobic descriptor, and β_p is coefficient for ON hydropathy (sequence), and ε is residual error (accounts for unmodeled variability (instrumental noise, descriptor uncertainty, and neglected minor mechanisms)). This model enables decomposition of the total apparent retention into solvent-dependent and solute-dependent global hydropathy effects. Note that in this PARCH formalism, the weighting factors applied to backbone and base contributions reflect their intrinsic chemical hydropathy, not their accessibility. The accessibility factor independently accounts for structural masking due to folding or duplex formation. Thus, intrinsic polarity (chemical nature) and exposure (geometry) are treated as orthogonal contributions, avoiding double counting of hydration effects.

Eq. (4) follows the logic of Quantitative Structure Retention Relationship (QSRR) models, expressing retention as a linear combination of solvent and solute descriptors. The term 'apparent' retention factor (k_{app}) reflects that measurements were made under gradient conditions, where k values are estimated from elution composition rather than isocratic experiments. Using $\ln k_{app}$ instead of gradient time avoids confounding solvent strength and gradient slope effects. Stationary phase contributions were not parameterized explicitly because all experiments used BEH Amide column, thus, their effect is assumed constant and absorbed into the intercept β_o. Extending the model to other HILIC chemistries would require adding stationary phase descriptors and interaction terms.

Because the absolute values of the three descriptors used in the model differ substantially in magnitude, the regression operates as if the predictors were implicitly centered. In particular, $PARCH_{seq}$ values are one to two orders of magnitude larger than the solvent descriptors, which forces the fitting procedure to adjust the intercept so that the average predicted $\ln k_{app}$ matches the experimental mean. As a result, the intercept β_o does not correspond to a physically meaningful retention value, but it serves as a mathematical offset that anchors the regression surface on the vertical $\ln k$ scale. The solvent-dependent (β_H, β_S) and sequence-dependent (β_P) terms then describe the true mechanistic contributions, while β_o simply compensates for the different numerical scales of the descriptors. This behavior is common in multivariate models combining molecular descriptors with heterogeneous magnitudes and does not affect the interpretability of the mechanistic coefficients.

Each nucleotide has two *PARCH* components: BB (backbone, phosphate and sugar atoms) and NB (base, nitrogenous base atoms). Higher *PARCH* values correspond to greater hydrophilicity (water accessibility). Typical DNA averages BB values ≈ 4.3 ± 1.4, and NB values ≈ 0.6 ± 0.7, while typical RNA averages BB values ≈ 5.0 ± 1.5, and NB ≈ 0.4 ± 0.4. Thus, the backbone dominates the hydrophilic (polar) interactions, while the bases contribute to hydrophobic stacking. Since HILIC retention depends on both solute–solvent polarity and H-bonding,

PARCH provides a quantitative, structure-dependent hydrophilicity descriptor that naturally captures both hydrogen bonding potential and solvophobic accessibility.

For a sequence of N nucleotides, $PARCH_{seq}$ can be determined from:

$$PARCH_{seq} = \sum_{j=1}^{N} a_j \left[w_{BB} PARCH_{BB,j} + w_{NB} PARCH_{NB,j} \right] \quad (5)$$

where f_j is the fraction of base j in the sequence, w_{BB} and w_{NB} are weights for backbone *vs* base (recommended 0.8: 0.2; backbone dominates hydration), and a_j is the accessibility factor (1.0 for fully exposed and stem-loop structures, and 0.1 for duplex structures).

Retention has been measured for various ON samples, at various ternary co-solvent mixtures (MeOH/IPA/MeCN + water). For each chromatographic run, we calculated the effective mobile-phase H-bond donor acidity ($\alpha_{MP,eff}$) and effective dipolarity/polarizability ($\pi^*_{MP,eff}$) corresponding to the solvent composition at the analyte's elution point. These values, together with the sequence-dependent $PARCH_{seq}$ descriptor, were used as fixed predictor variables in a global multivariate regression model of the form given in Eq. (4). The regression coefficients β_o, β_H, β_S, and β_P were estimated simultaneously by ordinary least-squares optimization, by minimizing the sum of squared residuals between experimentally measured ln k_{app} and model-predicted ln k_{app} over the full data set. The minimization was performed using the Excel Solver add-in (GRG non-linear algorithm), with all β coefficients treated as free parameters and no constraints imposed. Convergence was achieved when further changes in the coefficients did not reduce the residual sum of squares. All experiments were performed at low ($T = 30$ °C) and high ($T = 80$ °C) temperatures.

4. Results and discussion

4.1. Identification of the main factors driving ON retention

The two Resolution-IV fractional-factorial screening designs revealed a remarkably consistent picture across all tested oligonucleotides (Supplementary information Figure S1 shows the obtained Pareto charts including factor significance limits). In the MeOH–MeCN branch, the standardized effect estimates showed that the MeOH factor dominated retention by a wide margin, exceeding the significance threshold in all cases. This strong positive effect indicates that replacing MeCN with MeOH substantially reduces ON retention, consistent with competitive solvation and H-bond displacement at the stationary phase. In contrast, urea and cyrene produced no statistically significant effects, while ammonium acetate showed no systematic or reproducible statistically significant contribution across analytes and screening blocks.

The IPA–MeCN block showed a similar pattern. IPA exhibited a significant and reproducible effect on retention, but with a smaller magnitude than MeOH, aligning with its weaker H-bond acidity and stronger solvophobicity. Again, urea, cyrene and ammonium acetate remained below the $p = 0.05$ threshold, confirming that these factors are secondary or negligible under typical elution conditions. The fact that results were comparable for homopolymers and heteropolymers (linear and hairpin 20-mer) suggests that the dominance of the co-solvent factor persists even when secondary structure varies. Overall, the screening DoE clearly demonstrated that organic co-solvent character (a combination of solvophobicity and protic H-bond competition) is the primary driver for ON retention in HILIC.

Although urea, cyrene, and ammonium acetate are often referred to as influential modifiers in HILIC, their effects are strongly concentration- and composition-dependent. In our study, these additives were evaluated under mobile-phase conditions typical for oligonucleotide analysis (30–50% water, moderate salt content), where solvophobicity and protic competition were hypothesized to be the dominant retention drivers and were subsequently assessed experimentally.

Within this constrained and application-relevant design space, the tested ranges of urea (20–80 mM), cyrene (0.1–1% v/v), and ammonium acetate (25–75 mM) are sufficient to detect meaningful effects, yet their impact remains statistically insignificant because they are overshadowed by the much larger retention changes produced by MeOH/IPA substitution. Increasing their ranges further would lead to unrealistic or impractical mobile phases (e.g., excessive urea precipitation, phase instability, or MS incompatibility). Therefore, our results indicate that under realistic analytical conditions, these additives contribute only second-order effects compared to solvent-controlled solvophobic and H-bond-competition mechanisms.

It should also be noted that the present model was established using relatively short oligonucleotides (20-mer), and the influence of mobile-phase additives may become more pronounced for longer sequences or for molecules exhibiting more complex secondary or tertiary structures. For instance, larger oligonucleotides (e.g., ≥100-mers) or structured RNAs may present additional interaction modes with the stationary phase that were not captured in the current screening design. The applicability of the present model to larger and structurally more complex RNAs remains to be systematically evaluated and will be the subject of future investigations.

4.2. Mixture design analysis of MeOH, IPA, and MeCN contributions

To distinguish the individual and combined effects of MeOH, IPA, and MeCN on oligonucleotide retention, a simplex-lattice mixture design was applied, comprising 14 experimental points including vertex conditions, binary midpoints, and centroid replicates. This design allowed the systematic and statistically efficient exploration of solvent composition while maintaining a constant aqueous fraction and buffer concentration. The resulting retention data for six oligonucleotides (three homopolymers, dA20, dT20, dC20 and a heteropolymer in linear, hairpin, and double-stranded forms) showed clear and consistent trends that expose the mechanistic role of each organic modifier.

Fig. 1 shows the mixture-design contour plots illustrating how the combined proportions of MeOH, IPA, and MeCN modulate oligonucleotide retention. For most analytes, MeCN results in the highest retention, whereas MeOH systematically yields the lowest retention. IPA exhibits analyte-dependent behavior, often showing retention intermediate between MeOH and MeCN, but in some cases closely resembling MeCN. These trends reflect the differing contributions of solvophobicity and hydrogen bonding in the mobile phase: MeCN, being aprotic and relatively weakly solvating, enhances solute desolvation, while IPA, although protic, possesses a high hydrophobic surface area and strong cohesive interactions, contributing to a measurable solvophobic effect [27,28]. In contrast, MeOH, with its high H-bond donor activity and strong solvation capability, yielded markedly lower retention. These experimentally observed trends agree with the energetic ordering of donor acidity, cohesive energy density, and solvation properties associated with the three solvents.

The mixture design analysis further revealed strong curvature in the response surfaces, particularly along the MeOH–MeCN and IPA–MeCN edges, confirming the presence of non-linear mixing effects. For most oligonucleotides, the retention increases upon replacing MeOH with MeCN was greater than expected from linear interpolation, illustrating a synergistic enhancement of solvophobicity as the mixture shifts toward a less protic, more weakly solvating environment. The centroid replicates displayed excellent reproducibility, allowing reliable estimation of quadratic terms and validating the statistical robustness of the design. The small residuals observed in the model fits indicate that the chosen mixture model adequately captured the dominant solvent-dependent retention mechanisms.

The mixture-design data did not show a simple, monotonic ordering of retention among the homopolymer analytes (dT20, dA20, dC20). Instead, their relative retention varied with mobile-phase composition, demonstrating that base identity alone is not the primary determinant of their retention under these conditions. This observation is consistent

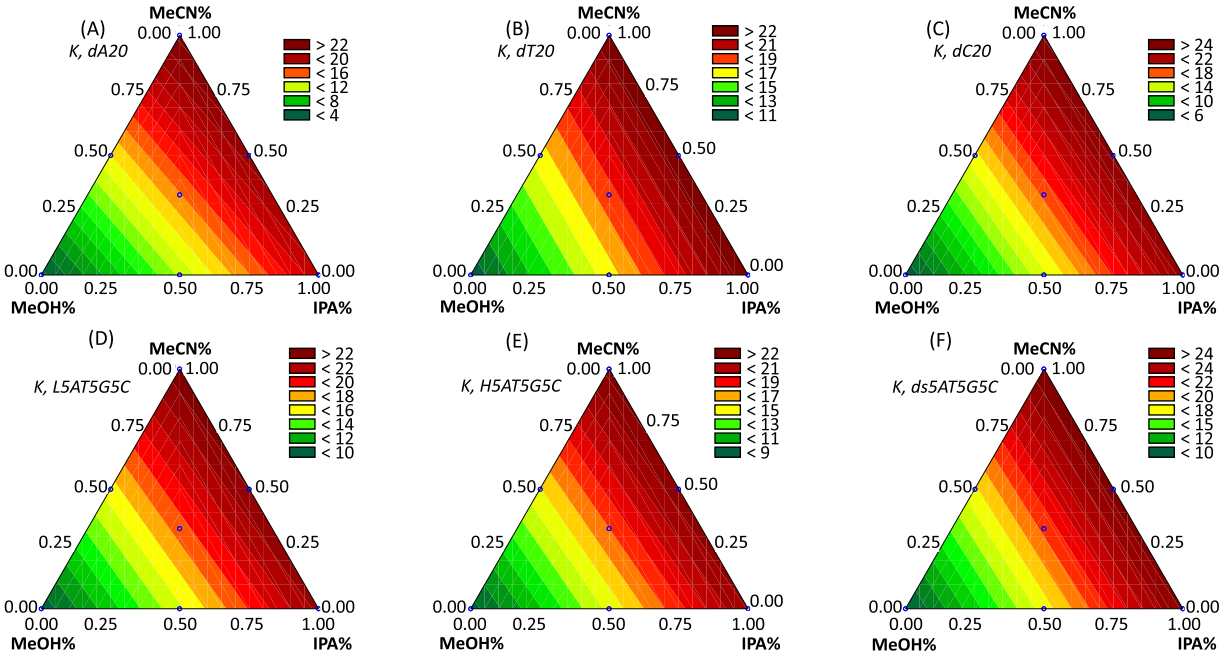


Fig. 1. Contour plots from the ternary mixture design showing the effect of MeOH, IPA, and MeCN composition on k_{app}. Each plot represents the quadratic Scheffé mixture model fitted to the experimental data at the 80:20 (organic:water) composition of mobile phase B. The contour surfaces visualize the balance between H-bond competition (MeOH), solvophobicity (IPA), and weakly solvating aprotic character (MeCN) that governs ON retention in HILIC. Analytes: dA-20 (A), dT-20 (B), dC-20 (C), L5A5T5G5C (D), H5A5T5G5C (E) and ds5A5T5G5C (F).

with our earlier screening results, which showed that MeOH (H-bond competition) and IPA (solvophobic strength) dominate the retention mechanism, while intrinsic differences in nucleobase hydrophilicity contribute less significantly.

Similarly, no universal trend was observed between secondary structure forms of the heteropolymer (i.e. linear, hairpin, duplex). The hairpin and duplex oligonucleotides did not consistently elute earlier than the linear form. Instead, their relative retention changed across the solvent space, indicating that structural compaction and partial shielding of nucleobases do not translate directly into a uniform reduction of retention. In some mixture regions, the hairpin or duplex even exhibited slightly increased retention, likely due to altered desolvation energies and different exposure of the charged phosphate backbone. These observations suggest that oligonucleotide HILIC behavior is governed mostly by solvent-controlled desolvation and competition effects rather than conventional hydrophilic partitioning or base-specific polarity. (Supplementary information Figure S2 and Table S4 show the detailed results of the mixture designs).

4.3. Retention modeling using combined solvent descriptors and PARCH hydropathy

To rationalize how solvent properties and oligonucleotide structural features jointly determine retention, a multivariate model was evaluated incorporating solvent H-bond donor acidity, mobile phase polarizability, and the sequence-dependent *PARCH* hydropathy scale. The resulting regression equation reproduced the experimentally measured ln k_{app} values with high accuracy across all analytes and all mobile-phase compositions. The global model (at $T = 30$ °C) yielded $\beta_o = 2.8609$, $\beta_H = 0.0487$, $\beta_S = 0.1680$, and $\beta_{PARCH} = 0.0272$, with an exceptionally small residual error (<0.012), indicating that the combined descriptors capture the major mechanistic contributions to ON retention. The overall fit was excellent ($R^2 = 0.938$), demonstrating that ~94% of the variance in retention can be explained using only three physically meaningful predictors. The standard error of the regression (0.052, expressed in ln k_{app}) corresponds to a relative prediction uncertainty of approximately 5–6% in k_{app}, based on conversion from logarithmic to fractional error.

Inspection of the regression coefficients shows that solvophobicity, represented by the polarizability term π^*_{MP}, is the dominant contributor to retention. The H-bond donor acidity of the mobile phase also contributes positively, but to a smaller extent, suggesting that protic modifiers influence retention by modulating solute solvation. The *PARCH* descriptor remains statistically significant, indicating that hydropathy captures physicochemical features that go beyond simple H-bond donor/acceptor counting. *PARCH* reflects the solvent-accessible polar surface and hydration propensity of the entire oligonucleotide, thus determining a broader solvation footprint.

To further probe the ability of *PARCH* to predict retention trends, the experimentally measured ln k_{app} values were examined as a function of $PARCH_{seq}$ under three representative mobile-phase compositions (10% MeOH/20% IPA/50% MeCN, 50% MeOH/10% IPA/20% MeCN and 20% MeOH/50% IPA/10% MeCN). In all cases, highly significant correlations were obtained, with R^2 values ranging between 0.76 and 0.85. This confirms that oligonucleotide hydropathy assessed via *PARCH* is an appropriate predictor of retention, independent of solvent polarity or protic character. The strongest dependence was observed in the solvent mixture containing the highest fraction of MeCN, where the eluotropic strength is lower and solvophobic interactions dominate. Under these conditions, oligonucleotides with lower *PARCH* (i.e., less strongly solvated by the mobile phase) exhibited markedly stronger retention. Even in more protic mobile phases containing high MeOH or IPA, the retention *vs PARCH* correlation remained statistically significant, suggesting that the solvophobic component of HILIC retention is closely linked to the solute's intrinsic hydropathy and solvation energy. Fig. 2 shows the regression of the retention model and the correlation between ON retention and *PARCH* values observed at $T = 30$ °C. (Supplementary information Table S5 shows the detailed statistical results obtained at low temperature).

Overall, these results demonstrate that the *PARCH* descriptor provides a valuable sequence- and structure-sensitive measure of analyte hydropathy that can be combined effectively with solvent acidity and polarizability descriptors to obtain a quantitative model of

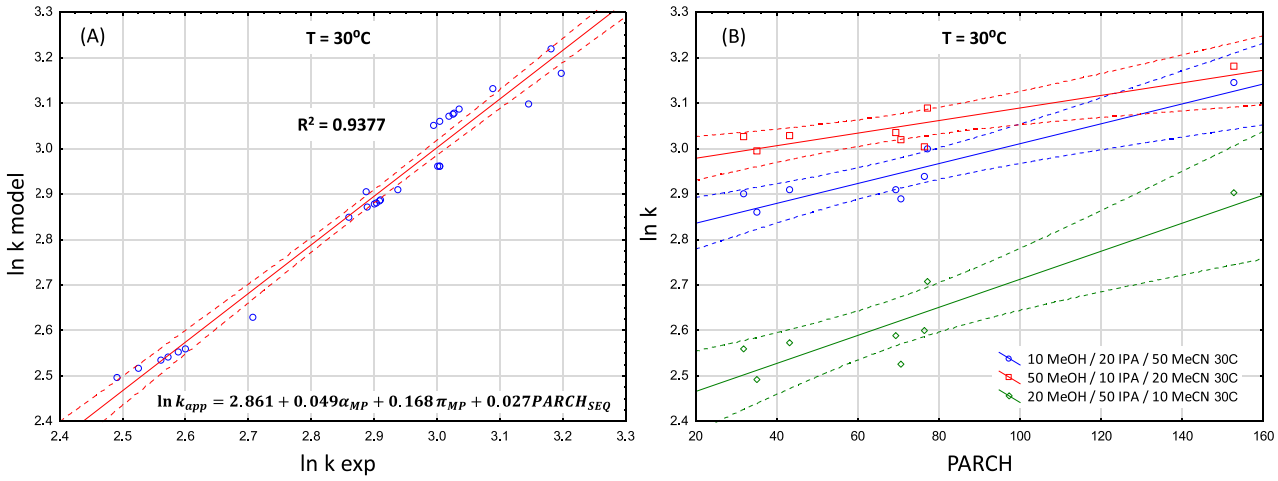


Fig. 2. Evaluation of the combined mechanistic retention model at 30 °C. Plot of predicted versus experimental ln *k* values (A). The solid line shows the ideal 1:1 correlation, and the dashed lines indicate the 95% confidence band. Correlation between ln *k* and *PARCH* for three representative ternary solvent compositions (B).

oligonucleotide retention.

4.4. Effect of temperature

Increasing the column temperature from 30 °C to 80 °C produced systematic and mechanistically meaningful changes in the regression coefficients of the combined retention model. The intercept (β_o) decreased only slightly (from 2.86 to 2.80), indicating that the baseline retention level of the chromatographic system is only weakly temperature-dependent under the applied gradient conditions.

The coefficient β_H remained essentially unchanged between 30 °C (0.0487) and 80 °C (0.0497). This suggests that the contribution of solvent–analyte H-bond interactions is only weakly temperature-dependent within this range, consistent with the relatively low enthalpy of H-bond disruption in mixed MeCN/alcohol eluents. In other words, although temperature accelerates all kinetic processes, the relative ability of MeOH or IPA to compete for oligonucleotide H-bonding sites remains largely invariant at elevated temperature.

In contrast, the β_S increased significantly with temperature (from 0.168 to 0.209). This change indicates that solvophobic rejection from the ACN-rich mobile phase becomes a stronger determinant of retention at high temperature. This is consistent with the expected thermodynamics, solvophobic (desolvation-driven) interactions have a positive heat capacity contribution, and therefore their free energy penalty becomes larger at elevated temperature. Higher temperatures weaken structured solvation around polar analytes and simultaneously enhance differences in solvation between MeCN, MeOH, IPA, and the stationary phase surface. As a result, oligonucleotides are even more strongly "pushed" toward the hydrophilic surface at 80 °C, increasing the apparent contribution of π_{MP} to retention.

The most notable temperature effect was observed for the oligonucleotide sequence-dependent descriptor, which decreased by nearly 50%, from 0.0272 at 30 °C to 0.0147 at 80 °C. This suggests that the structural contribution of the solute, which reflects the accessible hydrophilic surface and hydration penalty of the oligonucleotide, becomes less influential at elevated temperature. Physically, this can be attributed to two effects:

(1) higher temperatures weaken ordered hydration shells surrounding the oligonucleotide backbone and bases, reducing differences between analytes of varying *PARCH* values, and
(2) internal secondary structures (hairpins, duplexes) partly "loosen" at high temperature, increasing conformational averaging and reducing the effective contrast between sequences with different accessibility patterns. Consequently, at 80 °C the retention becomes more strongly governed by solvent-controlled interactions (β_S) and somewhat less by intrinsic solute hydropathy (β_{PARCH}). It is important to note that temperature may influence the conformational stability of structured oligonucleotides, particularly double-stranded species. Duplex melting or partial strand destabilization can occur at elevated temperatures, depending on sequence length, GC content, and ionic strength. Such structural transitions modify the solvent accessibility of nucleobases and the phosphate backbone, which in turn affects the effective hydropathy captured by the PARCH descriptor. Consequently, for duplex oligonucleotides analyzed near or above their melting temperature, the descriptor may represent a temperature-averaged structural state rather than a strictly defined duplex configuration. This effect may contribute to the reduced magnitude of the sequence-dependent coefficient observed at higher temperatures.

Overall, these results show that temperature selectively amplifies solvophobic interactions while reducing sequence-dependent hydration effects, whereas H-bond competition between mobile phase and solute remains nearly unchanged. Fig. 3 shows the retention model and the correlation between ln *k* and *PARCH* values observed at $T = 80$ °C. (Supplementary information Table S6 shows the detailed statistical results obtained at elevated temperature).

4.5. Interpretation of PARCH in relation to HILIC retention

Although *PARCH* was originally introduced as a generalized measure of peptide, protein and nucleic-acid hydropathy [30], in our HILIC system its correlation with oligonucleotide retention must be interpreted without invoking classical hydrophilic partitioning. At the elution conditions used here (30–50% MeCN), an immobilized water layer is not present. In classical HILIC conditions, when MeCN fraction > 65–70% the immobilized water rich layer is about 0.5–2 nm thick corresponding to 2 – 8 molecular layers [29]. A 20-mer oligonucleotide is ≈7 nm long (2–3 nm coiled hydrodynamic thickness) thus cannot penetrate into the immobilized solvent layer. Thus, classical hydrophilic partitioning is unlikely here. Instead, *PARCH* acts as a solute structural descriptor that captures how strongly a given oligonucleotide is repelled from the acetonitrile-rich mobile phase and subsequently attracted to the polar stationary phase. Oligonucleotides with higher *PARCH* values possess a larger accessible polar surface and more H-bond capable functional groups. In a mixed MeOH/IPA/MeCN environment, these highly polar sequences experience particularly poor solvation in the MeCN-rich bulk

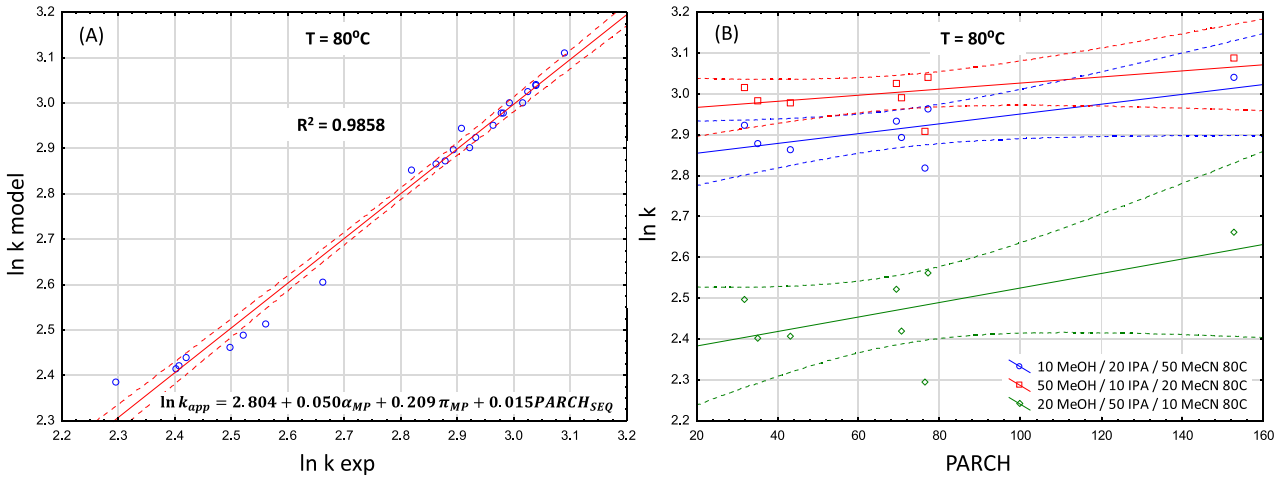


Fig. 3. Evaluation of the combined mechanistic retention model at 80 °C. Plot of predicted versus experimental ln *k* values (A). The solid line shows the ideal 1:1 correlation, and the dashed lines indicate the 95% confidence band. Correlation between ln *k* and *PARCH* for three representative ternary solvent compositions (B).

phase, which enhances the solvophobic driving force that pushes them toward the surface of the stationary phase. Once at the surface, the same polar groups facilitate extensive direct hydrogen bonding and dipolar interactions with the amide ligands. Thus, the observed positive correlation between *PARCH* and retention reflects a two-step mechanism: strong solvophobic expulsion from the MeCN-rich eluent followed by compensating, energetically favorable interactions with the hydrophilic stationary phase through hydrogen bonds. The *PARCH* descriptor therefore does not measure hydrophilic partitioning but rather quantifies the overall polarity-driven tendency of an oligonucleotide to engage in surface interactions once it is rejected from the bulk mobile phase. This interpretation is fully consistent with the absence of a stable water layer at low MeCN content and with the systematically higher retention observed for larger, more polar, and more hydrated oligonucleotide sequences.

Although this study was conducted on BEH Amide stationary phase, the observed ranking of mechanistic contributions (solvophobicity > hydrogen bonding > ionic interactions) is expected to remain qualitatively similar for amide-type and weakly charged HILIC stationary phases operated under buffered conditions typical for oligonucleotide analysis. The absolute magnitude of these effects may vary depending on surface functionality and ligand polarity.

4.6. Proposal for method optimization

Based on the screening results, it became evident that MeOH and MeCN exert markedly different influences on oligonucleotide retention, suggesting that controlled blending of these two co-solvents offers a powerful means for method development. In particular, the strong solvophobic and mild H-bonding characteristics of MeCN contrast sharply with the pronounced H-bond competition and weaker solvophobicity of MeOH, leading to distinct and often opposite retention trends. Consequently, optimizing a ternary mobile phase composed of MeOH/MeCN/water provides a considerably broader and more flexible design space than relying on either co-solvent alone.

To explore this potential, we performed a two-factor optimization involving gradient steepness (*tG*) and a ternary organic composition (*tC*) using a linear-solvent-strength (LSS)–based semi-empirical model (DryLab). Gradient time varied between 10 and 20 min (*tG1* and *tG2*), while the organic modifier fraction was tested at three distinct compositions: 100% MeCN, 50% MeCN/50% MeOH, and 100% MeOH. All experiments were conducted at $F = 0.3$ mL/min and $T = 30$ °C on a short 50 × 2.1 mm column. A representative mixture of five 20-mer oligonucleotides (including linear and hairpin isomers of the same sequence, a second heteropolymer, a homopolymer, and a mimetic ON standard for Volanesorsen with 3 *m*C) was used as the calibration test sample.

It is important to note that the DryLab-based optimization shown here is independent of the descriptor-based statistical model proposed earlier. The DryLab approach uses a semi-empirical framework for gradient prediction and was included only to illustrate the practical benefit of using MeOH-rich mobile phases for selectivity improvement, not to validate or extend the mechanistic interpretation.

Resolution maps generated using DryLab software revealed pronounced elution-order changes as the MeOH/MeCN ratio was varied, highlighting strong solvent-dependent selectivity effects. The map (Fig. 4) indicated that MeOH-rich conditions provided the highest resolution and robust conditions. The solvent strength parameters (*S*) of the fitted LSS models suggested that as MeOH concentration increases, the solute retention becomes less sensitive to mobile phase composition. Lower *S* values were obtained in MeOH compared to MeCN rich mobile phases. (Supplementary information Table S7 lists the obtained model parameters.)

In practice, lower *S* values (i.e. $S < 10$) indicate that a longer effective column length can be utilized to improve the separation efficiency [31]. Therefore, we transferred our optimized method to a longer column format to further improve the separation. Fig. 5 shows the predicted and the experimentally observed chromatograms for both 50 × 2.1 mm and 150 × 2.1 mm columns. High prediction accuracy was observed. This example suggests that despite the common practice of using only MeCN as the preferred solvent in HILIC, it is probably not the best choice. Oligonucleotide separations can benefit from protic co-solvents, especially from MeOH.

5. Conclusion

HILIC has become increasingly popular for the analysis of nucleic acids, yet its retention mechanism is poorly understood. For the first time, this work investigated the various interactions that take place on amide-based HILIC columns, in realistic conditions for the analysis of oligonucleotides. Hydrogen bonding, dipole-dipole interactions, solvophobic effects, and ionic interactions were evaluated through design-of-experiments studies and samples specifically chosen to challenge the conventional HILIC retention mechanism, based on hydrophilic partitioning.

Our results suggest that partitioning is not a dominant mechanism in oligonucleotide HILIC retention. Moreover, retention is still observed when protic-rich mobile phases are used, despite their incompatibility with the formation of the water-rich stationary layer. In contrast, our

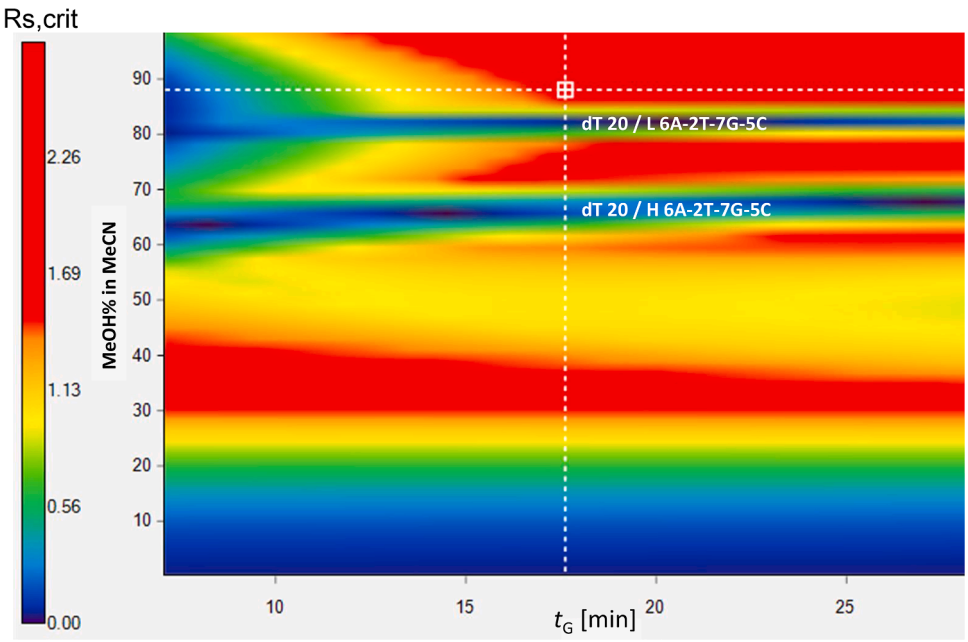


Fig. 4. Resolution map for a 5-compound ON mixture: critical resolution (Rs,crit) as a function of gradient time (t_G) and ternary mobile phase composition (MeOH% in MeCN%). The dark blue areas indicate peak coelution, while the red areas indicate baseline or greater resolution. At 63% MeOH in MeCN, the dT-20 and hairpin 6A-2T-7G-5C ON peaks are co-eluting, while at 80% MeOH in MeCN, the dT-20 and linear 6A-2T-7G-5C ON peaks are co-eluting.

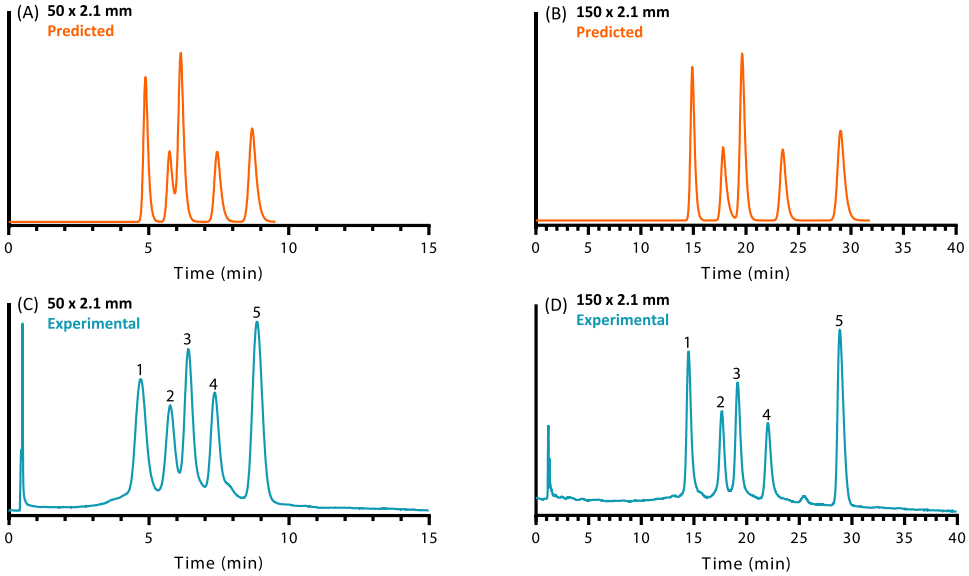


Fig. 5. Optimized separations. Panels A and B show the predicted chromatograms for 50 × 2.1 and 150 × 2.1 mm columns, respectively, while panels C and D show the experimentally measured chromatograms. Peaks: hairpin 4A-5T-3G-8C (1), Volanesorsen 3 *m*C (2), hairpin 6A-2T-7G-5C (3), linear 6A-2T-7G-5C (4), dT-20 (5). Conditions: mobile phase A: 30 mM ammonium acetate in pure water, mobile phase B: 30 mM ammonium acetate in 20% water, 76% MeOH, 4% MeCN, gradient from 100% B to 87% B in 15 min (50 mm column) or 45 min (150 mm column), $F = 0.3$ mL/min, $T = 30$ °C.

data provide evidence for solvophobic effects and H-bonding to be the primary factors in oligonucleotide retention. This retention mechanism is therefore best described as "solvophobic-dominated HILIC".

These findings have important implications for future method development. Indeed, the use of protic co-solvents can tune the selectivity of a HILIC separation and help resolve closely related species. We were able to obtain a baseline separation of five sequence variants of 20-mer oligonucleotides using a 100% to 87% mobile phase B in 15 min that was composed of a ternary mixture of 20% H_2O, 76% MeOH, and 4% MeCN, a result that was not achievable with a traditional pure acetonitrile-based mobile phase. These results highlight that MeOH/MeCN mixtures (or even MeOH rich mobile phases), rather than MeCN alone, provide a more flexible and effective organic mobile phase composition for optimizing HILIC separations of oligonucleotides.

Our findings support the view that modern HILIC should be regarded not as a single hydrophilic partitioning mechanism, but as a spectrum of interaction-dominated retention regimes whose balance depends on analyte size, charge, and solvent environment.

Further exploration of other column chemistries and ligand densities is needed to incorporate descriptors of stationary phase characteristics in the retention model.

Disclosure

Waters, ACQUITY, UPLC, BEH, GTxResolve, BEH and Empower are trademarks of Waters Technologies Corporation. Milli-Q is a trademark of Merck KGaA. Excel is a trademark of Microsoft Corporation, DryLab is a trademark of Imre Molnar. Statistica is a trademark of Cloud Software Group, Inc. All other trademarks are the property of their respective owners.

Szabolcs Fekete, Matthew A. Lauber, Fabrice Gritti and Bala Addepalli are Waters employees, a company that manufactures and sells chromatography columns. Jonathan Maurer and Davy Guillarme declare no competing interests.

CRediT authorship contribution statement

Jonathan Maurer: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Balasubrahmanyam Addepalli:** Writing – review & editing, Visualization, Methodology, Investigation, Conceptualization. **Fabrice Gritti:** Writing – review & editing, Validation, Methodology, Conceptualization. **Matthew A. Lauber:** Writing – review & editing, Validation, Resources, Funding acquisition, Conceptualization. **Davy Guillarme:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Szabolcs Fekete:** Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Waters, ACQUITY, UPLC, BEH, GTxResolve, BEH and Empower are trademarks of Waters Technologies Corporation. Milli-Q is a trademark of Merck KGaA. Excel is a trademark of Microsoft Corporation, DryLab is a trademark of Imre Molnar. Statistica is a trademark of Cloud Software Group, Inc. All other trademarks are the property of their respective owners.

Szabolcs Fekete, Matthew A. Lauber, Fabrice Gritti and Bala Addepalli are Waters employees, a company that manufactures and sells chromatography columns.

Jonathan Maurer and Davy Guillarme declare no competing interests.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2026.467001.

Data availability

Data will be made available on request.

References

[1] M. Liu, et al., Landscape of small nucleic acid therapeutics: moving from the bench to the clinic as next-generation medicines, Signal. Transduct. Target. Ther. 10 (1) (2025) 73.

[2] R. Minkner, et al., Oligonucleotide separation techniques for purification and analysis: what can we learn for today's tasks? Electrophoresis 43 (23-24) (2022) 2402–2427.

[3] S. Pourshahian, Therapeutic oligonucleotides, impurities, degradants, and their characterization by mass spectrometry, Mass Spectrom. Rev. 40 (2) (2021) 75–109.

[4] F. Hannauer, et al., Advancements in the characterisation of oligonucleotides by high performance liquid chromatography-mass spectrometry in 2021: a short review, Anal. Sci. Adv. 3 (3-4) (2022) 90–102.

[5] V. D'Atri, et al., Trends in pharmaceutical analysis: the evolving role of liquid chromatography, Anal. Chem. 97 (9) (2025) 4706–4727.

[6] J. Maurer, et al., Evaluation of ion pairing reversed-phase liquid chromatography for the separation of large RNA molecules, J. Chromatogr. A 1740 (2025).

[7] M. Donegan, J.M. Nguyen, M. Gilar, Effect of ion-pairing reagent hydrophobicity on liquid chromatography and mass spectrometry analysis of oligonucleotides, J. Chromatogr. A 1666 (2022) 462860.

[8] A. Liu, et al., Bioanalysis of oligonucleotide by LC-MS: effects of ion pairing regents and recent advances in ion-pairing-free analytical strategies, Int. J. Mol. Sci. 23 (24) (2022).

[9] M. Gilar, D.R. Stoll, Challenges and solutions in oligonucleotide analysis, part I: an overview of liquid chromatography methods and applications, LCGC Int. 2 (7) (2025).

[10] M.G. Bartlett, Current state of hydrophilic interaction liquid chromatography of oligonucleotides, J. Chromatogr. A 1736 (2024) 465378.

[11] Y. Guo, Separation of nucleobases, nucleosides, nucleotides and oligonucleotides by hydrophilic interaction liquid chromatography (HILIC): A state-of-the-art review, J. Chromatogr. A 1738 (2024).

[12] J. Maurer, et al., Proof of concept application of hydrophilic interaction chromatography for direct online disruption of lipid nanoparticles, intact mRNA analysis, and measure of encapsulation efficiency, Anal. Chem. (2025).

[13] P.A. Lobue, et al., Oligonucleotide analysis by hydrophilic interaction liquid chromatography-mass spectrometry in the absence of ion-pair reagents, J. Chromatogr. A 1595 (2019) 39–48.

[14] H. Lardeux, et al., Unravelling the link between oligonucleotide structure and diastereomer separation in hydrophilic interaction chromatography, Anal. Chem. 96 (24) (2024) 9994–10002.

[15] M. Ntorkou, C.K. Zacharis, Applications of hydrophilic interaction chromatography in pharmaceutical impurity profiling: A comprehensive review of two decades, Molecules 30 (17) (2025).

[16] D.V. McCalley, Understanding and manipulating the separation in hydrophilic interaction liquid chromatography, J. Chromatogr. A 1523 (2017) 49–71.

[17] A.J. Alpert, Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids, and other polar compounds, J. Chromatogr. A 499 (1990).

[18] Y. Guo, D. Baran, L. Ryan, Quantitative assessment of retention mechanisms for ionized compounds in hydrophilic interaction chromatography (HILIC), Anal. Chem. 97 (7) (2025) 4057–4065.

[19] Y. Guo, D. Baran, Hydrophilic partitioning or surface adsorption? A quantitative assessment of retention mechanisms for Hydrophilic interaction chromatography (HILIC), Molecules 28 (18) (2023).

[20] Y. Guo, B. Fattal, Relative significance of hydrophilic partitioning and surface adsorption to the retention of polar compounds in hydrophilic interaction chromatography, Anal. Chim. Acta 1184 (2021) 339025.

[21] Y. Guo, D. Muscatiello, An update on the progress in fundamental understanding of hydrophilic interaction liquid chromatography, J. Chromatogr. A (2026) 1765.

[22] F. Gritti, et al., Retention and mass transfer properties of the series of unbonded, amide-bonded, and alkylsulfobetaine-bonded ethylene bridged hybrid hydrophilic interaction liquid chromatography columns, J. Chromatogr. A 1692 (2023) 463828.

[23] F. Gritti, et al., Comparison of the adsorption mechanisms of pyridine in hydrophilic interaction chromatography and in reversed-phase aqueous liquid chromatography, J. Chromatogr. A 1216 (48) (2009) 8496–8504.

[24] F. Gritti, J. Sehajpal, J. Fairchild, Using the fundamentals of adsorption to understand peak distortion due to strong solvent effect in hydrophilic interaction chromatography, J. Chromatogr. A 1489 (2017) 95–106.

[25] D. Guillarme, S. Fekete, S. Studzinska, Protocol for oligonucleotides characterization using hydrophilic interaction chromatography, J. Sep. Sci. 48 (1) (2025) e70088.

[26] A.R. Gruber, et al., The Vienna RNA websuite, Nucleic. Acids. Res. 36 (Web Server issue) (2008) W70–W74.

[27] W.A. Fouad, et al., Understanding the thermodynamics of hydrogen bonding in alcohol-containing mixtures: cross-association, J. Phys. Chem. B 120 (13) (2016) 3388–3402.

[28] C. Reichardt, T. Welton, Solvents and Solvent Effects in Organic Chemistry, 4th ed., John Wiley & Sons, 2011.

[29] M.J. Kamlet, et al., Linear solvation energy relationships. 23. A comprehensive collection of the solvatochromic parameters, *, a, and ß, and some methods for simplifying the generalized solvatochromic equation, J. Org. Chem (1983).

[30] R. Mandal, et al., Introducing the PARCH scale for quantifying the hydropathy of nucleic acids and nucleic acid-protein complexes, Langmuir 41 (26) (2025) 16874–16886.

[31] S. Fekete, M. Lauber, Studying effective column lengths in liquid chromatography of large biomolecules, J. Chromatogr. A 1692 (2023) 463848.