



Automated flow synthesis of algorithmically designed ferroelectric nematogens

Richard J. Mandle

To cite this article: Richard J. Mandle (2023) Automated flow synthesis of algorithmically designed ferroelectric nematogens, *Liquid Crystals*, 50:3, 534-542, DOI: 10.1080/02678292.2023.2208550

To link to this article: <https://doi.org/10.1080/02678292.2023.2208550>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 08 May 2023.



Submit your article to this journal [↗](#)



Article views: 697



View related articles [↗](#)



View Crossmark data [↗](#)

Automated flow synthesis of algorithmically designed ferroelectric nematogens

Richard J. Mandle^{a,b}

^aSchool of Physics and Astronomy, University of Leeds, Leeds, UK; ^bSchool of Chemistry, University of Leeds, Leeds, UK

ABSTRACT

The synthesis of the ferroelectric nematic material RM734 is telescoped into a single continuous flow process, utilising inline liquid–liquid separation to remove reactants and by-products from the reaction stream. Following the final synthetic step, we subject the semi-crude material to offline automated chromatography. Materials can be prepared on a gram scale (12 minutes) and chromatographically purified (8 minutes) in around 20 minutes. Over a given time interval the system described herein can prepare a far larger number of materials (on gram scale) than an experienced chemist. Secondly, we implement an established process for generation of new molecular structures via a genetic algorithm, which generates many new candidate structures per iteration. A filter removes undesirable structures, and a fast semi-empirical QM calculation is performed on those that remain. A neural network trained against a library of nematic and ferroelectric nematic materials is used to rank the candidate structures based on the QM output and a fingerprinting process, with the highest scoring candidates then used to seed a new round of mutations. We combine these two approaches; we generated and ranked several tens of thousands of variants of RM734 and synthesised a small library of materials using our continuous flow protocol.

ARTICLE HISTORY

Received 23 February 2023
Accepted 26 April 2023

KEYWORDS

Liquid crystals; flow chemistry; RM734; ferroelectric nematics



Introduction

The ferroelectric nematic (N_F) phase is the most recently discovered nematic polymorph, in which the orientational order of the ‘classical’ nematic phase is augmented with polar order [1–5], endowing this phase and the materials that exhibit it with remarkable properties not easily replicated by other systems [6–9].

The design of new molecules and materials that exhibit the N_F phase is in its infancy [10–22]; we know little about how molecular structure generates this phase, and the majority of materials known to exhibit it are derivatives of either RM734 [11] or DIO [23]. The pool of known materials remains small and grows slowly, limiting the pull towards applications.

CONTACT Richard J. Mandle  r.mandle@leeds.ac.uk

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

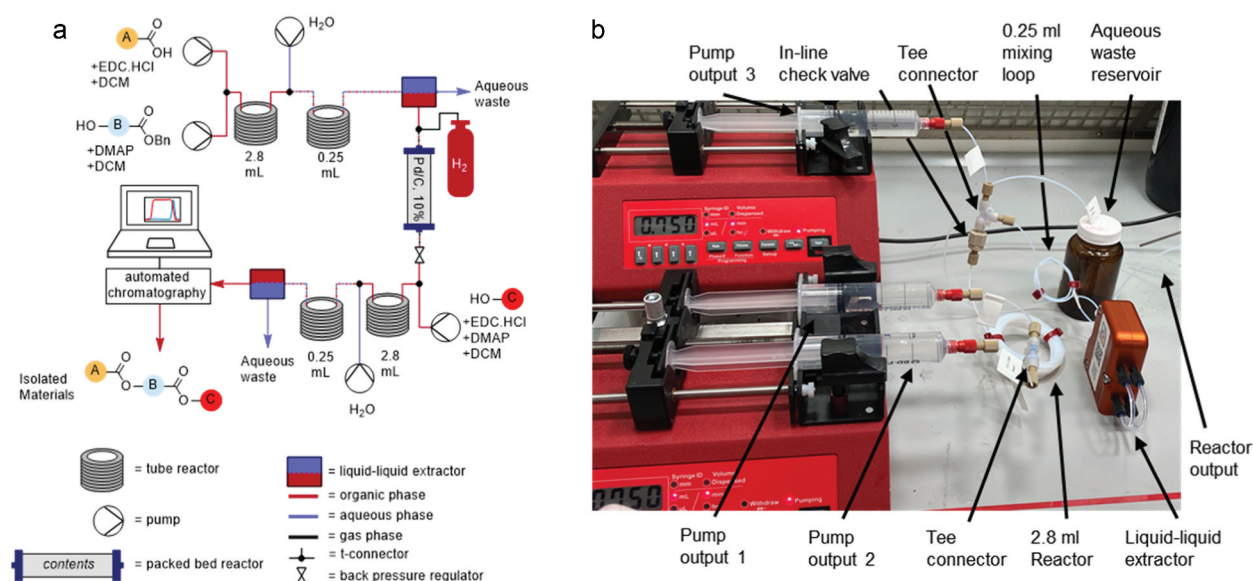


Figure 1. (Color online) (a) Schematic overview of the continuous flow system used in this work. The output of the first reaction step is fed directly into the H-cube for the second reaction step, which is in turn fed directly into the third reaction step. (b) Photograph of an assembled GRM with key parts highlighted.

We considered that simply increasing the production rate of new materials is a logical way to increase the rate of discovery of N_F materials (accelerated serendipity). The productivity of the synthetic chemist is constrained not by their creativity, nor their capacity to design materials, but rather the sheer time taken to set-up, execute, and purify a chemical reaction (typically 1–48 h). This time can be extended greatly by troublesome emulsive liquid–liquid extractions, complex chromatographic separations, and the need to maintain inert (e.g. argon, nitrogen) or reactive (e.g. hydrogen, oxygen) atmospheres, among many other variables. Most (if not all) of these reactions are conducted in batch, with multiple synthetic steps each being subjected to isolation and some form of purification. Together, these factors conspire to present many challenges in scaling up the synthesis of larger quantities of materials.

Continuous flow systems offer a simplified route to scale up as well as simplifying the synthesis of libraries of materials, and as such, and are widely employed in academic and industrial chemistry, yet are largely absent from the liquid crystal literature save for a few examples [24–27].

Then, we come to the topic of the chemical libraries themselves. For a multitude of reasons, some benign and some not, the synthesis of libraries of liquid crystals often resorts to the same ‘tricks’; a fluorine here, a nitrogen there, methyl to ethyl to butyl to *futile*. These variations are often logical progression based on prior art and are sometimes borne out of habit. Algorithmic methods for generating new chemical

structures are an intensive area of research [28–32], and offer a route to generate new candidate structures from some initial starting point. Algorithmic generation of molecular structure could potentially reduce or even avoid human biases, suggesting new or underexplored chemical spaces for experimental study. However, as with continuous flow synthesis, the deployment of such methods is as yet essentially absent from the liquid crystal literature.

Herein we combine these two approaches. We first report a series of continuous flow reactors configured in such a way as to produce useful quantities of liquid crystalline materials in a short timeframe, as well as deploying an algorithmic method to design new materials which can then be synthesised in flow.

Methods

Flow-chemistry

RM734 and related materials were synthesised here by using a continuous flow protocol which performs sequential esterification/debenzylation/esterification (Figure 1(a)).

Our flow system consists of what we term ‘general reaction modules’ (GRM, Figure 1(b)) and an off-the-shelf H-cube system for high-pressure continuous flow hydrogenation. A GRM consists of two programmable push/pull dual syringe pumps (New Era), i.e. a total of four outputs. Outputs 1 and 2 of pump 1 are dedicated to reagents; they are combined by a tee connector which

flows into a PTFE reaction loop (vol = 2800 μ l). The outputs of pump 2 deliver the aqueous quench/washing solution to the output of the reaction loops *via* a tee connector; a small PTFE reactor loop (vol = 250 μ l) ensures thorough mixing. In-line liquid–liquid extraction is performed by a membrane-based separator (Zaiput) with the aqueous waste being discarded and the organic portion being passed directly to the next reactor (or to automated chromatography).

The H-cube system employs a HPLC pump to deliver reagents into the system and performs hydrogenolysis of the benzyl protecting unit using a 10% Pd/C catalyst bed and hydrogen gas at a pressure of up to 100 bar and a temperature of up to 100°C. The H-cube system contains an inbuilt back pressure regulator.

The post-reaction solution is subjected to automated off-line chromatographic purification by a Combiflash NextGen 300+ system (Teledyne Isco); the chromatographed material is then filtered through a 0.2 micron PTFE filter and finally recrystallised ethyl acetate/hexanes.

Algorithmic molecular design

Algorithmic generation of candidate structures was performed *via* random mutations to the structure of RM734. The input molecular structure is first converted to a text string under the group-SELFIES (group self-referencing embedded strings) representation protocol [33]. We use a custom fragment library which contains only ‘liquid crystal-like’ units and combinations thereof; the generation of this library is detailed in the ESI. Looking forward, the ability to construct most liquid crystalline materials from a relatively small library of fragments suggests the possibility of creating a semantically correct language, with clear advantages for machine-lead material discovery.

This first representation undergoes a random mutation in which a fragment (e.g. ring, functional group) is inserted and an existing fragment replaced or deleted, giving a library of variant structures. We generate an additional structural diversity by using the STONED algorithm [34] to generate additional structural variants through random mutation again, this time with the molecule represented as a SELFIES string [35].

We perform several sequential rounds of mutations, each generating many thousands of candidate structures. After each round of mutation, we remove molecular structures that: are duplicates, contain rings which have less than five or more than seven atoms; have more than 3 contiguous rotatable bonds; have no rings. A second filter removes molecules that do not have two ester groups; while there is no *a priori* reason we

cannot synthesise these, they would require a different configuration of the flow-reactor setup described above, and fall outside the scope of this paper as a result. After the final iteration, a visual inspection removes structures which are likely to be unstable, is probably unsynthesisable (e.g. the phosphaalkyne in Figure 2) and/or is unstable under atmospheric conditions (e.g. silabenzynes, benzyne and such).

For these plausible structures, we then generate starting 3D coordinates for each structure with RDKit and performs an initial optimisation with the MMFF force field. The MMFF optimised geometry was then passed to Gaussian G09 [36] for further optimisation with the semi-empirical AM1 method. We then calculated a series of molecular properties at the AM1 level (dipole tensor, polarisability tensor, hyperpolarizability tensor, molecular dimensions, molecular volume, and shape anisotropy parameter). Additionally, we use RDKit to compute a ‘fingerprint’ for each molecule which counts the number of instances of various structural features which are found in one or more N_F materials (e.g. number of nitro groups, aryl fluorines, contiguous rotatable bonds, pi electrons, and so on; over 100 properties in total).

Scoring is then made by a trained a multilayer perceptron feedforward neural network (using Pytorch). This input to the neural network is the matrix of molecular properties (from AM1 calculations) and fingerprint data (from RDkit), which is used to predict the scaled N_F -N(Iso) transition temperature which we define here as:

$$T_{scaled} = T_{N_F-N(Iso)} / T_{Cr}$$

The input used to train this network is all N_F materials that have either: been publicly disclosed; are unpublished N_F materials synthesised in our group; are closely related molecules which do not exhibit the N_F phase.

The top scoring molecules from each round were then each used to seed another round of mutations. We performed eight rounds of mutation all in all, and we selected the top 100 molecules from all rounds based on their scores. A visual inspection was then performed, and we selected the highest ranking molecules that could be synthesised using the continuous flow platform described in this paper using readily available building blocks.

The GitHub repository (<https://github.com/RichardMandle/AlgMol1>) contains Python code for structure modification *via* SELFIES/STONED [34] and group-Selfies. A Jupyter Notebook is provided in the GitHub repository that provides a ‘walk-through’ for generation of 5CB variants using SELFIES/STONED,

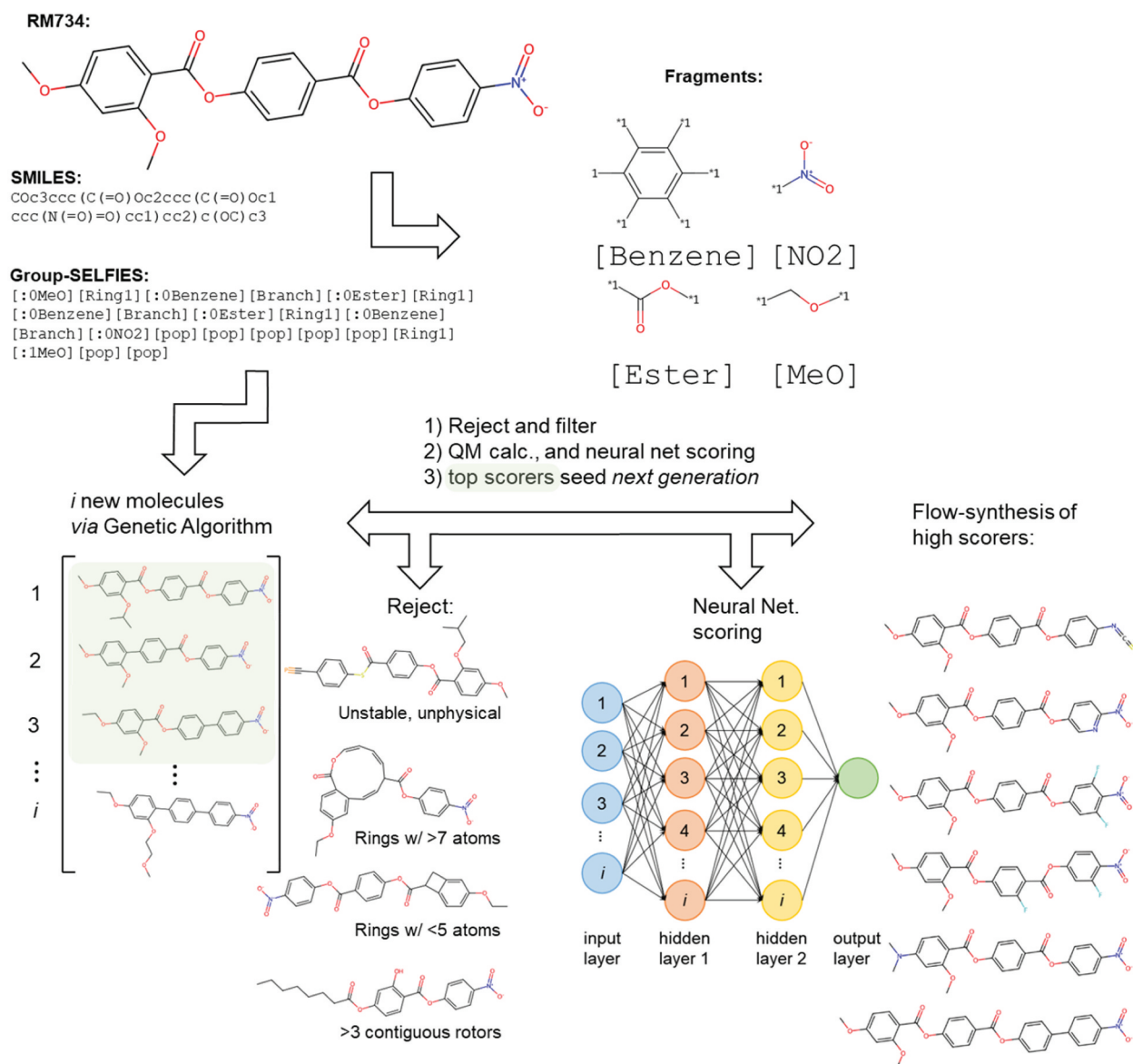


Figure 2. (Color online) Top: Example of user input (RM734) and corresponding representation as SMILES and group-SELFIES, with ring fragments indicated. Bottom: Exemplar molecules from a single round of mutation to RM734; we filter molecules using the rules shown, and score using a neural network. Materials which score highly and were later synthesised using our continuous flow apparatus (vide infra).

fragment-based generation of RM734-variants *via* group-selfies and genetic algorithms which utilise fingerprint-based scoring regimes. Functions are provided for filtering, substructure searching, scoring, initial geometry creation and optimisation using MMFF/RDKit, creating input files for Gaussian (and Mopac, although this is not used here).

Results

Flow-reactor optimisation

We considered that, in the present case, a continuous flow approach is only beneficial if it can offer

a significant advantage in throughput. The most person-time consuming portion of organic synthesis is chromatographic purification; we therefore purify synthetic intermediate steps with an in-line liquid-liquid extraction only, with chromatographic separation reserved for the final synthetic step. We sought to benchmark our system on a model synthesis in this case, the well-known ferroelectric material RM734, targeting the highest possible throughput that gives >95% conversion to the target material at each step.

We utilise a single general reaction module (GRM) for the initial esterification between 2,4-dimethoxybenzoic acid and benzyl 4-hydroxybenzoate. The time for the reaction to complete depends upon the flow rate and

Table 1. Conversion (^1H NMR) for the esterification of 2,4-dimethoxybenzoic acid (1.05 mol eqv.), benzyl 4-hydroxybenzoate (1.00 mol eqv.), EDC.HCl (1.2 mol eqv.), and DMAP (0.05 mol eqv.) in dichloromethane at the indicated flow rate (Q ; ml min^{-1}) and concentration (mmol ml^{-1}). Reactions were conducted at ambient temperature (measured as 18°C), and after each run the flow path was washed with 5 volumes of DCM (~ 15 ml) and 5 volumes of water (quench zone only, ~ 1.5 ml). Dashes (-) indicate experiments not performed due to low conversion at lower concentrations/flow-rates.

Q C	≤ 0.50	0.75	1.00	1.50
≤ 0.10	>95%	>95%	>95%	>95%
0.25	>95%	>95%	>95%	91%
0.50	>95%	>95%	93%	91%
0.75	>95%	92%	88%	85%
1.00	76%	73%	-	-

also on the size of the reaction loop; we fixed the reaction loop size at $2800 \mu\text{l}$ as an arbitrary first choice (giving a residence time of 2.8 minutes at a flow rate of 1 ml min^{-1}). We then explore simultaneous variations in both reagent concentration and flow rate, the latter being the sum of the flows from both reagent pumps. For each run, we adjust the flow rate of the quench pump to match the flow rate of the organic stream.

We explored flow rates in the range of 0.05 – 1.00 ml min^{-1} , with concentrations ranging from 0.1 M to 1.50 M . In Table 1 we focus only on the limiting cases: at low flow rates ($\leq 0.75 \text{ ml min}^{-1}$) we observe quantitative (e.g. >95%) conversion at all but the highest concentration. Similarly, at low concentration ($\leq 0.10 \text{ M}$) we observe quantitative conversion at all but the highest flow rates. At high concentrations (1.00 M) we encounter problems with the solubility of the acid-EDC complex, which has a deleterious effect upon conversion, and so subsequent experiments with high flow rates were not performed.

Concentrations in the range of 0.25 – 0.75 M give quantitative conversion at certain flow rates (e.g. 0.25 M at 1.00 ml min^{-1}), and incomplete conversion at higher rates (e.g. 0.75 M at 1.50 ml min^{-1}). Our objective is to obtain the highest throughput possible with near-quantitative conversion, enabling us to pass the reactor output straight into the next reaction without purification and so eliminating the need for time-consuming chromatography and solvent-removal. A flow rate of 0.75 ml min^{-1} at a concentration of 0.50 M is the highest throughput that gives quantitative conversion [37].

We next sought to explore the hydrogenolysis step. Here, the concentration is determined by the (optimised) output of the preceding step, while the flow rate is required to be equal to the outgoing flow rate of the preceding step (i.e. 0.75 ml min^{-1}). Our variables to optimise here are therefore temperature and pressure; the former increasing the rate of reaction (at the cost of reduced hydrogen gas solubility) and the latter – partially – dictating the solubility of hydrogen in the solvent stream (Henry's law). Initial investigations employed the flow rate (0.75 ml min^{-1}) and concentration ($0.50 \text{ mmol ml}^{-1}$) which gave the highest throughput in the preceding step, however we were unable to attain quantitative debenzoylation under any of the explored conditions. Moreover, the liberated benzoic acid is not entirely soluble in DCM at these concentrations which presented additional challenges (e.g. precipitation in the high pressure hydrogen cell, blocked output lines, fouling of the catalyst cartridge (etc.), Table 2). We therefore selected the second-highest throughput conditions from Table 1, using a flow rate of 0.75 ml min^{-1} and a concentration of 0.25 M .

DCM is hardly a widely encountered solvent for hydrogenation, yet it performs adequately in here. At ambient temperature (internal temperature set at

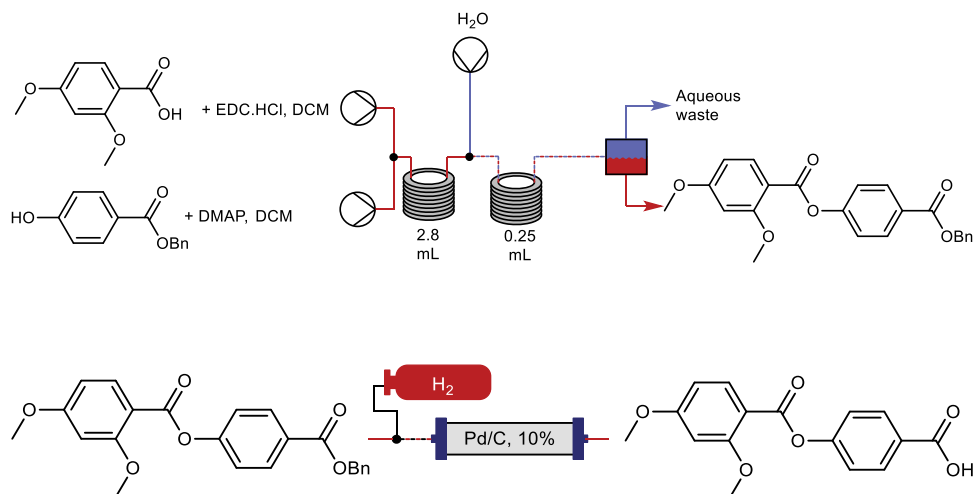
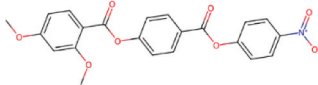
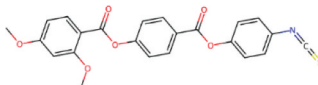
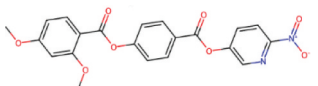
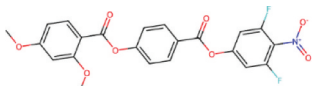
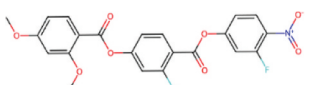
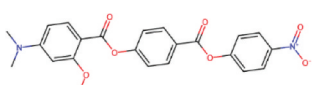
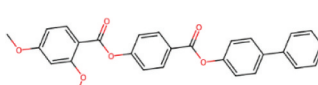


Table 2. Conversion (^1H NMR) for the debenzyla-tion via hydrogenolysis of benzyl 4-(2,4-dimethoxybenzoyloxy)benzoate (aka RM733), obtained in continuous flow (Table 1), in dichloromethane at a concentration of ~ 0.25 M. Reactions were performed with a flow rate of 0.75 ml min^{-1} . Reactions were conducted at the indicated temperature (T , $^\circ\text{C}$) under an atmosphere of hydrogen gas generated by electrolysis, at the indicated pressure (P , Bar). After each run, the flow path was washed with 5 volumes of DCM then 5 volumes of ethanol.

P		
T	1	100
20	$\sim 85\%$	$\sim 85\%$
40	-	$>95\%$
60	-	$>95\%$
80	-	$>95\%$
100	-	$>95\%$

20°C) there is incomplete conversion at atmospheric (1 Bar) and elevated (100 Bar) pressure. We considered that the rate of reaction is probably not limited by the solubility of hydrogen (which is higher in DCM than in common hydrogenation solvents such as THF or ethanol) [38] but rather the rate of adsorption/desorption onto the catalyst itself, which will be temperature dependent. Taking advantage of the backpressure regulation within the H-cube system, we deployed elevated pressure (up to 100 Bar) enabling us to work at temperatures far beyond the boiling point of DCM at atmospheric pressure (up to 100°C here). A temperature of 40°C is sufficient to attain quantitative debenzylation at a concentration of 0.25 M and a flow rate of 0.75 ml min^{-1} .

Table 3. Transition temperatures (T , $^\circ\text{C}$) and associated enthalpies (ΔH , kJ mol^{-1}) for a family of RM734 analogues. # The clearing point was above 250°C , and we were unable to measure the enthalpy or temperature for this phase transition owing to the sample decomposing. * An additional phase transition is seen at $\sim 160^\circ\text{C}$ by microscopy which is presumably an antiferroelectric/splay nematic; we omit this from the table for clarity. Transition temperatures for several compounds are in keeping with prior reports [11,16,22].

Name		Melt	N _F -N	N-Iso
RM734	T ΔH	139.8 29.9	132.7 0.2	187.9 0.5
				
RM734-S	T ΔH	153.0 51.2	- -	162.7 0.3
				
RM734-3Py	T ΔH	164.9 66.2	154.0* 0.5	192.6 0.4
				
RM734-U	T ΔH	173.7 55.4	147.6 2.1	- -
				
RM734-GG	T ΔH	150.7 50.9	140.6 3.5	- -
				
RM734.DMA	T ΔH	186.9 44.7	171.7 1.5	- -
				
RM734.P	T ΔH	191.4 38.9	163.0 1.7	>250 # -#
				

For the final esterification, the flow-rate and concentration of carboxylic acid is determined by the output of the hydrogenolysis step (0.75 ml min^{-1} , $\sim 0.25 \text{ mmol ml}^{-1}$). The output of the H-cube is fed into a T-mixer along with a DCM solution of EDC.HCl (0.5 mmol ml^{-1}), 4-nitrophenol ($0.28 \text{ mmol ml}^{-1}$), and DMAP ($0.01 \text{ mmol ml}^{-1}$) at a flow rate of 0.75 ml^{-1} . The streams are then fed through a 2.8 ml reaction loop as described elsewhere. Following in-line liquid-liquid extraction this provides $>95\%$ conversion to RM734, and so there are no parameters to optimise here.

Next, a confirmatory experiment was performed in which the optimised reaction conditions were applied to each step in the sequence. After achieving a steady state, we collected the output of the final flow reactor for 12 minutes, and subjected this to automated off-line chromatography (~ 8 minute run time). The chromatographed material was concentrated to ~ 10 ml volume, filtered through a 0.2 micron PTFE filter and triturated with hexanes to induce crystallisation. The solid material was then collected and dried *via* suction filtration. This afforded 0.85 grams of isolated material (1.94 mmol , $\sim 85\%$ yield), giving a throughput of $11.3 \text{ mmol hr}^{-1}$ ($\sim 5 \text{ g hr}^{-1}$).

Mini-library synthesis

With optimised conditions at hand, we now apply our continuous flow platform to the synthesis of some RM734-like materials. We algorithmically generated many variants of RM734 through systematic replacement of molecular fragments under the group-selfies representation. Variants were scored using a neural network, and the top 100 were visually inspected to select those that could be synthesised using the flow platform, i.e. those constructed by sequential esterification-hydrogenolysis/debenzylation-esterification. Clearly, the system could be modified to permit the synthesis of many other things, but this goes beyond the scope of this paper. We also restricted our synthetic endeavours to targets whose building blocks could either be purchased, gifted to us or were available in-house.

Some high-scoring materials have been reported previously and were resynthesised here, with excellent agreement between the transition temperatures in this work and the original reports (Table 3). One notable consequence of using the neural network is that it appears to overestimate the importance of polarisability, biasing our selection towards highly polarisable materials by scoring these highly. Some suggested functional groups lack the polarity to generate the N_F phase (e.g. RM734-S), yet others (RM734.dma, RM734.pyr) exhibit enhancements to the stability of the N_F phase relative to RM734, albeit with increased melting points. Relative to

the parent material, the biphenyl containing RM734.P has large increases at the melting point, the onset of N_F phase, and also at the clearing point. Presumably the increased aspect ratio helps to arrest head-to-tail flipping.

Our use of a neural network to score the algorithmically designed molecules, especially one trained on such a small set of materials, is that molecules that are very similar to RM734 score strongly, whereas more radical deviations (e.g. Figure 2) generally score poorly or are filtered out. The neural network is probably significantly less adept at discerning the likelihood of any given structure being worth synthesising (or not!) than one skilled in the art; however, the advantage is that many thousands of molecules can be evaluated simultaneously. It is difficult to see how (or why) a human should go about sifting through the huge number of structures that can be generated per minute. It is to be expected that our ability to score candidate molecules will improve significantly over time, both as more sophisticated networks are trained and deployed and also as we can start to draw on an ever-growing pool of experimental data. The possibility stands that the generated molecules, which we currently filter out – either *via* the neural network or by human intervention – may in fact be the roughest diamonds in the mine.

Conclusions

Here, we outline the use of a continuous flow reactor to telescope multiple chemical transformations into a single process, which vastly increases the number of materials we can synthesise in a given time. By telescoping multiple optimised reactions into a single continuous flow process, the bottleneck is no longer synthesis but characterisation, chemical analysis and so on. We also deploy an algorithmic method for designing new chemical matter from an initial starting molecule. This method reduces human bias, and produces output structures which range from the obvious to the bizarre, from trivial to unphysical. The output of this algorithmic method is scored using a neural network which is trained to predict the scaled onset temperature of the N_F phase. We then synthesise several of the high-scoring candidate structures using our continuous flow system, providing a mini-library of RM734-like materials.

Acknowledgments

RJM gratefully acknowledges funding from UKRI via a Future Leaders Fellowship, grant no. MR/W006391/1, and funding from the University of Leeds *via* a University Academic Fellowship.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the UK Research and Innovation [MR/W006391/1].

Data availability statement

The data associated with this paper are openly available from the University of Leeds Data Repository at <https://doi.org/10.5518/1343>. Annotated code is available from GitHub (<https://github.com/RichardMandle/AlgMoll>).

References

- [1] Mertelj A, Cmok L, Sebastian N, et al. Splay nematic phase. *Phys Rev X*. 2018;8(4). DOI:10.1103/PhysRevX.8.041025.
- [2] Chen X, Korblova E, Dong D, et al. First-principles experimental demonstration of ferroelectricity in a thermotropic nematic liquid crystal: polar domains and striking electro-optics. *PNAS*. 2020;117(25):14021–14031.
- [3] Mandle RJ, Sebastián N, Martínez-Perdiguero J, et al. On the molecular origins of the ferroelectric splay nematic phase. *Nat Commun*. 2021;12(1):4962.
- [4] Kikuchi H, Matsukizono H, Iwamatsu K, et al. Fluid layered ferroelectrics with global ∞v symmetry. *Adv Sci*. 2022;9(26):2202048.
- [5] Chen X, Martínez V, Nacke P, et al. Observation of a uniaxial ferroelectric smectic a phase. *arXiv preprint arXiv:12965*. 2022.
- [6] Sebastian N, Cmok L, Mandle RJ, et al. Ferroelectric-ferroelastic phase transition in a nematic liquid crystal. *Phys Rev Lett*. 2020;124(3):037801.
- [7] Folcia CL, Ortega J, Vidal R, et al. The ferroelectric nematic phase: an optimum liquid crystal candidate for nonlinear optics. *Liq Cryst*. 2022;49(6):1–8.
- [8] Sebastian N, Mandle RJ, Petelin A, et al. Electrooptics of mm-scale polar domains in ferroelectric nematic phase. *Liq Cryst [Internet]*. 2021;48(14):48.
- [9] Tibor Máthé M, Buka Á, Jáklí A, et al. Ferroelectric nematic liquid crystal thermo-motor. *ArXiv e-prints*. 2022:arXiv:2201.07556.
- [10] Mandle RJ, Cowling SJ, Goodby JW. A nematic to nematic transformation exhibited by a rod-like liquid crystal. *Phys Chem Chem Phys*. 2017;19(18):11429–11435.
- [11] Mandle RJ, Cowling SJ, Goodby JW. Rational design of rod-like liquid crystals exhibiting two nematic phases. *Chem: Eur J*. 2017;23(58):14554–14562.
- [12] Mandle RJ, Cowling SJ, Goodby JW. Structural variants of RM734 in the design of splay nematic materials. *Liq Cryst*. 2021;48(12):1780–1790.
- [13] Li J, Nishikawa H, Kougo J, et al. Development of ferroelectric nematic fluids with giant- ϵ dielectricity and nonlinear optical properties. *Sci Adv*. 2021;7(17):eabf5047.
- [14] Zhao X, Zhou J, Li J, et al. Spontaneous helielectric nematic liquid crystals: electric analog to helimagnets. *PNAS*. 2021;118(42):e2111101118.
- [15] Pocięcha D, Walker R, Cruickshank E, et al. Intrinsically chiral ferronematic liquid crystals: an inversion of the helical twist sense at the chiral nematic - chiral ferronematic phase transition. *J Molec Liq*. 2022;361:119532.
- [16] Saha R, Nepal P, Feng C, et al. Multiple ferroelectric nematic phases of a highly polar liquid crystal compound. *arXiv e-prints*. 2021:arXiv:2104.06520.
- [17] Brown S, Cruickshank E, Storey JMD, et al. Multiple polar and non-polar nematic phases. *Chem Phys Chem*. 2021;22(24):2506–2510.
- [18] Li J, Wang Z, Deng M, et al. General phase-structure relationship in polar rod-shaped liquid crystals: importance of shape anisotropy and dipolar strength. *Giant*. 2022;11:100109.
- [19] Song Y, Li J, Xia R, et al. Development of emergent ferroelectric nematic liquid crystals with highly fluorinated and rigid mesogens. *Phys Chem Chem Phys*. 2022;24(19):11536–11543.
- [20] Dai S, Li J, Kougo J, et al. Polar liquid crystalline polymers bearing mesogenic side chains with large dipole moment. *Macromolecules*. 2021;54(13):6045–6051.
- [21] Cruickshank E, Walker R, Storey JMD, et al. The effect of a lateral alkyloxy chain on the ferroelectric nematic phase. *RSC Adv*. 2022;12(45):29482–29490.
- [22] Cigl M, Podoliak N, Landovský T, et al. Dimethylamino terminated ferroelectric nematogens revealing high permittivity. *arXiv e-prints*. 2023:arXiv:2301.04865.
- [23] Nishikawa H, Shiroshita K, Higuchi H, et al. A fluid liquid-crystal material with highly polar order. *Adv Mater*. 2017;29(43):1702354.
- [24] Mandle RJ, Goodby JW. Inexpensive, operationally simple flow reactor for rapid, chromatography free synthesis of siloxane based liquid crystals. *React Chem Eng*. 2018;3(4):515–519.
- [25] Xia S, Ding X, Wang Y, et al. Continuous-flow synthesis of an important liquid-crystal intermediate using a microreaction system. *Ind Eng Chem Res*. 2018;57(36):12113–12121.
- [26] Chen Q, Xia S, Luo G, et al. Continuous-flow synthesis of the liquid crystal intermediate 1-(4-ethoxy-2,3-difluorobenzyl)-4-propylcyclohexan-1-ol in a microfluidic system: experimental and numerical studies. *Chem Eng Sci*. 2022;254:117645.
- [27] Zak M, Herman J, Czerwiński M, et al. Deuterated liquid crystals - design and synthesis of deuterium labelled 4,4''-dialkyl-2',3'-difluoro-[1,1': 4',1'']terphenyls using batch and continuous flow systems. *Liq Cryst*. 2022;1–14. DOI:10.1080/02678292.2022.2145379
- [28] Gómez-Bombarelli R, Wei JN, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Sci*. 2018;4(2):268–276.
- [29] Zhou Z, Kearnes S, Li L, et al. Optimization of molecules via deep reinforcement learning. *Sci Rep-UK*. 2019;9(1):10752.
- [30] Zhavoronkov A, Ivanenkov YA, Aliper A, et al. Deep learning enables rapid identification of potent DDR1

- kinase inhibitors. *Nat Biotechnol.* **2019**;37(9):1038–1040.
- [31] Brown N, Fiscato M, Segler MHS, et al. GuacaMol: benchmarking models for de novo molecular design. *J Chem Inf Model.* **2019**;59(3):1096–1108.
- [32] Gao W, Coley CW. The synthesizability of molecules proposed by generative models. *J Chem Inf Model.* **2020**;60(12):5714–5723.
- [33] Cheng AH, Cai A, Miret S, et al. Group SELFIES: a robust fragment-based molecular string representation. *Digital Discovery.* **2023**. DOI:10.1039/D3DD00012E
- [34] Nigam A, Pollice R, Krenn M, et al. Beyond generative models: superfast traversal, optimization, novelty, exploration and discovery (STONED) algorithm for molecules using SELFIES. *Chem Sci.* **2021**;12(20):7079–7090.
- [35] Krenn M, Häse F, Nigam A, et al. Self-referencing embedded strings (SELFIES): a 100% robust molecular string representation. *Mach Learn.* **2020**;1(4):045024.
- [36] Frisch MJ, Trucks GW, Schlegel HB, et al. Gaussian 09. Gaussian 09. **2009**.
- [37] Zhou S, Xin Y, Wang J, et al. Application of continuous flow in tazobactam synthesis. *Org Process Res Dev.* **2021**;25(7):1648–1657.
- [38] Shirono K, Morimatsu T, Takemura F. Gas solubilities (CO₂, O₂, Ar, N₂, H₂, and He) in liquid chlorinated methanes. *J Chem Eng Data.* **2008**;53(8):1867–1871.