DoE-Driven Development of an Organocatalytic Enantioselective Addition of Acetaldehyde to Nitrostyrenes in Water**

Valeria Nori,[a] Arianna Sinibaldi,[a] Giuliana Giorgianni,[a] Fabio Pesciaioli,[a] Francesca Di Donato,[a] Emanuele Cocco,[a] Alessandra Biancolillo,[a] Aitor Landa,[b] and Armando Carlone*[a]

Abstract: The development of an enantioselective enamincatalysed addition of masked acetaldehyde to nitroalkenes via a rational approach helped to move away from the use of chloroform. The presented research allows the use of water as a reaction medium, therefore improving the industrial relevance of a protocol to access very important pharmaceutical intermediates. Critical to the success is the use of chemometrics-assisted ‘Design of Experiments’ (DoE) optimisation during the development of the presented new synthetic approach, which allows to investigate the chemical space in a rational way.

In the last twenty years an impressive number of organic transformations exploiting organocatalysis has been reported.[1] This technology platform has gained more and more recognition in virtue of both its greenness[2] and its potential to mimic enzymes, which makes it ideal for industrial applications, where it is now gaining increasing traction.[3] A cost-efficient strategy to access valuable γ-aminoacids, such as baclofen and pregabalin, is the enamine-catalysed addition of acetaldehyde to nitroalkenes; however, this approach is not without its challenges. In fact, acetaldehyde tends to form oligomers, is highly reactive, toxic, and flammable. It is not surprising that it took several years to develop protocols for this reaction; seminal works by Hayashi[4] (Scheme 1a) and List[5] (Scheme 1b) showed that an aminocatalytic enantioselective addition of acetaldehyde to nitroalkenes can be performed efficiently. However, a high catalyst loading and a large excess of acetaldehyde were needed to counterbalance the challenges posed by the reagents. Hayashi subsequently reported a protocol where they were able to use 2 equiv. of acetaldehyde for the addition on a nitrostyrene derivative.[6] Few years later, Pericàs ingeniously employed paraldehyde which can release free acetaldehyde in situ (Scheme 1c).[7] Unfortunately, 10 equivalents of acetaldehyde (3.3 equiv. of paraldehyde) and a relatively high catalyst loading of two supported organocatalysts were still needed. Recently, our group succeeded to replace free acetaldehyde with acetaldehyde dimethyl acetal;
this way, it was possible to lower both the equivalents of acetaldehyde and the organocatalyst loading, obtaining the desired products in high yields and ee (Scheme 1d). However, a limitation of the report is the use of chloroform, a class 2 solvent. To overcome this hurdle, we decided to explore the possibility to carry out the reaction in water. The use of water is becoming increasingly common in organocatalysis,[10] since Barbas reported for the first time that water was a good media for organocatalytic reactions.[10] Water has unique characteristics as a solvent. It is cheap, non-toxic, available in bulk, eco-friendly, hazardless and non-flammable.

These properties make it an ideal candidate as a solvent or cosolvent from an industrial standpoint.[11] In our instance, it became obvious that many parameters needed to be assessed to optimise the reaction in water (e.g.: catalyst type and loading, additives, concentration, cosolvent). Therefore, instead of proceeding with a trial-and-error approach following the one-variable-at-a-time (OVAT) method, it was chosen to explore the chemical space in a rational way using Design of Experiments (DoE).[12,13] In fact, most of the times, the variables influencing the reaction interact with each other, so they need to be changed simultaneously to find the actual optimum conditions.

Herein we report the rational development of the organocatalytic asymmetric enamine reaction of acetaldehyde dimethyl acetal 5 in water. The desired γ-nitroaldehydes derivatives were obtained in good yield and enantioselectivities using an extremely simple, safe, and green protocol (Scheme 1).

A preliminary catalyst screening was performed to test proline-based organocatalyst, previously shown to perform well in water,[14] bearing an array of substituents I–IX. 4-Chloro-β-nitrostyrene 1a was chosen as reaction partner since it would lead to the baclofen precursor 3a. The screening was performed employing acetaldehyde dimethyl acetal 5, capable to release in situ free acetaldehyde under the effect of Amberlyst-15.[15] The results obtained are summarised in terms of conversion and enantioimetric excess of the desired product 3a (Table 1).

The organocatalysed Michael addition of 5 (2 equiv.) to 1a (1 equiv.) in water was sluggish, albeit with moderately good ee, when using catalyst 1 (entry 1). A preliminary catalyst screening was performed to test whether longer aliphatic substituents could improve the desired reactivity in water (entries 2–9). It was indeed shown that VII, bearing the longest aliphatic chain, could promote the reaction at a faster rate and with higher ee (entry 7), while others did not show any appreciable conversion (entries 2–6), or produced a lower conversion (entry 8–9).

Decreasing the concentration of the reaction did not show any effect on the conversion and ee (entries 7, 10, and 11). Challenged by these results, it was decided to pursue the optimisation of the reaction looking at diverse variables; however, a rather high number of parameters could have been screened and an OVAT approach would not have proven rational. For the first optimisation, a full factorial design would have required 43200 experiments, whereas a D-optimal design drastically reduced the number of experiments to 34.

### Table 1. Preliminary data for the Michael addition of masked acetaldehyde to nitroalkenes in water.[28]

| Entry | Catalyst | M [mol/L] | Conv. [%] | ee [%] |
|-------|----------|-----------|-----------|--------|
| 1     | I        | 1.6       | 27        | 62     |
| 2     | II       | 1.6       | <5        | n.d.   |
| 3     | III      | 1.6       | 6         | 79     |
| 4     | IV       | 1.6       | <5        | n.d.   |
| 5     | V        | 1.6       | <5        | n.d.   |
| 6     | VI       | 1.6       | <5        | n.d.   |
| 7     | VII      | 1.6       | 20        | 79     |
| 8     | VIII     | 1.6       | 12        | 72     |
| 9     | IX       | 1.6       | 17        | 61     |
| 10    | X        | 0.8       | 19        | 80     |
| 11    | XI       | 0.4       | 18        | 79     |

[a] Reactions performed on 0.4 mmol scale; catalyst (0.02 mmol, 5 mol%), 4-chloro-β-nitrostyrene 1a (60 mg, 0.4 mmol, 1 equiv.), acetaldehyde dimethyl acetal 5 (85 μL, 0.8 mmol, 2 equiv.), Amberlyst-15 (14 mg, 10 mol%) and water as solvent at room temperature. [b] Measured by 1H NMR spectroscopy. [c] Determined by chiral HPLC analysis after conversion of the aldehyde into the corresponding alcohol by reduction with NaBH₄.

The first DoE was applied to evaluate how different variables affect the desired reactivity and if their effects could be linked. The selected parameters were: 1) type of organocatalysts; 2) catalyst loading; 3) equivalents of S; 4) reaction time; 5) reaction concentration; 6) cosolvents; 7) acidic catalysts to deprotect 5 (organic, inorganic, and immobilised acids were tested); 8) acidic catalyst loading; 9) ionic strength of water (for further details, see Supporting Information). The quantitative variables were investigated using the usual coding for the levels (−1 for low and +1 for high level) whereas the implicit level was used to evaluate the qualitative ones (I for (1), absence of co-solvent for (6), Amberlite 2900 for (7), absence of salt for (9)); in other words, each coefficient is to be considered with respect to the implicit level performances.

The responses were analysed for 3a in terms of conversion and enantioimetric excess (Figure 1).

The model generated showed that the best results were obtained with organocatalysts I and VII: furthermore, it was decided to include II as well in the next optimisation step, given its structural similarity with VII. Both the catalyst loading, and the concentration were significant parameters. On the other hand, cosolvents and salts showed no effect. Other parameters, such as equivalents of 5, time, and amount of acidic catalyst, were not significant; however, they were included in the following screening to explore any arising contribution. In terms
of the acidic catalyst, the resin Amberlite 2900 gave the best results; Amberlyst-15, Amberlite 1200 and HCl, cost effective catalysts, were, however, also included in the next screening. A second design of experiments (D-Optimal) was performed investigating: 1) organocatalysts (I, II and VII); 2) catalyst loading; 3) equivalents of 5; 4) reaction time; 5) reaction concentration; 6) acidic catalysts; 7) acidic catalyst loading (for further details, see Supporting Information). This second DoE drastically reduced the number of experiments from 288 for a full factorial design to 42 experiments, including 6 replicates.

The second model built on PC2-scores as response showed that the best results were obtained with organocatalysts I and VII (Figure 2). Choosing either catalyst is arbitrary, following the responses of the model. Moreover, the best acidic catalyst in terms of conversion and enantiomeric excess was Amberlite120. Concerning the organocatalyst loading, its coefficient is significant and negative; it means that it contributes positively to the increase of the final response (10 mol% seemed to be the best compromise). On the contrary, the coefficient related to time of reaction is positive but not significant; as consequence, its variations do not affect the final responses to be optimised. This variable does not correlate significantly with the other ones, so its level can be decided independently of the conditions used for the other variables. 24 h were chosen as optimum because of practical and economical points of view. The loading of the acidic catalyst, the concentration of the reaction mixture and the equivalents of 5 correlate with each other: the response is maximised when at the same time the equivalents of 5 is at its highest level (2.5 equiv.), the concentration of the limiting reactant in the aqueous medium is at its highest level (1.2 M) and loading of the acidic resin is at its lowest level (5 mol%). Both the catalyst loading, and the concentration of the reaction mixture were significant parameters and contributed positively to the increase in response.

Similarly, equivalents of 5, time, and amount of acidic catalyst were not significant, so their variations did not affect the final responses, but it was decided to continue exploring their contribution. The replications (inclusion of repeated experiments) have been used to evaluate experimental variation and the reproducibility of the reaction.

Analysis of the results showed agreement between the replicates which are close in the space of the main components PC1 and PC2; the trend showed by the samples justifies the choice of the scores on PC2 as response (to be minimised in order to maximise Y_{NMR} and ee) (Figure 3). To further optimise the reaction conditions for all the responses under study, a third DoE (full factorial design with two variables and two levels) was performed investigating: 1) organocatalysts (I and VII); 2) catalyst loading; 3) equivalents of 5; 4) reaction time; 5) reaction concentration; 6) acidic catalysts; 7) acidic catalyst loading (for further details, see Supporting Information). This third DoE drastically reduced the number of experiments from 288 for a full factorial design to 36 experiments, including 6 replicates.
performed, one for each remaining organocatalysts (I and VII). The variable investigated were: 1) reaction time and 2) loading of the acidic catalyst.

Organocatalyst I was rejected because the resulting chemometric model was not significant given that none of the coefficients associated to the investigated variables is significant. Conversely, the model generated to optimise yield and ee employing organocatalyst VII resulted significative only for the former. This means that no significative improvements for the ee occur in the investigated domain. The yield, on the contrary, can be optimised looking at the value of the coefficients. As shown in Figure 4, the coefficient related to the reaction time is positive; the response is maximised when time is at its highest level (36 h). On the other hand, the coefficient related to the percentage of the acidic catalyst is not significative; therefore, the best choice is the lowest loading (2.5 mol%). These two variables do not correlate with each other, in fact the resulting chemometric model is linear. The mathematical model (Eq. (1)) was validated to demonstrate its capability in response prediction and optimisation.

\[
Y_{\text{NMR}} \% = 43 + 15.5 \times \text{time} + 0.5 \times \%\text{mol acid catalyst} - 5 \times \text{time} \times \%\text{mol acid catalyst}
\] (1)

Based on the chemometric results, the conditions optimised via the last Design of Experiments were chosen to evaluate the generality of the reaction (Scheme 2).

As expected, and in agreement with previous reports, nitrostyrene 1b was found to be the most reactive among all other aromatic nitroalkenes. Nevertheless, both electron-withdrawing and electron-donating groups are well tolerated and several nitrostyrene derivatives successfully afforded the desired Michael adducts in good yields and enantioselectivity. Very interestingly, our method development proved to be robust; in fact, the results are consistent going from 0.2 mmol to 1 mmol to 1 g scale. This is even more striking given the biphase nature of the reaction mixture, along with the stirring and mixing issues that may arise on such different scales, without further development.

Unfortunately, aliphatic nitroalkenes (Scheme 3) proved to be a limitation of the developed protocol; in fact, in contrast with the reaction run in CHCl₃, the formation of several by-products, when running the reaction in water, resulted in very low yields.

Therefore, it was decided to further optimise the reaction conditions for these substrates, focusing on 1f. Two factors (at two different levels), acidic catalyst loading (10 mol% and 20 mol%) and acetaldehyde dimethyl acetal 5 (5 equiv. and 10 equiv.), were taken into account into a further full-factorial design. Consequently, 7 additional experiments (4 + 3 replicates of the central point) were carried out. Nevertheless, the coefficients resulted to be non-significant (Figure 5), meaning that it is not possible to significantly improve the yield by modifying the selected variables in the local investigated domain and the general outcome of the analysis indicated the
The Michael addition of acetaldehyde to nitroalkenes in water was developed. The investigation was performed with the aim of moving from the use of chloroform to water, and via a rational exploration of the chemical space by using DoE. While a current limitation remains the application to aliphatic nitroalkenes, the corresponding aromatic products were obtained in good yields and high ee.

Acknowledgements

G.G. is grateful to PON-DOT13OV2OC for an industrial Ph.D. fellowship. F.P. and A.B. thank PON-AIM grant number 1842894 for funding this research. A.L. thanks the University of the Basque Country UPV/EHU (UFIQO5YC11/22), Basque Government (GVgrant IT1236-19), and Ministerio de Ciencia e Innovación (grant PID2019-109633GB-C21) for financial support. Open Access Funding provided by Università degli Studi dell'Aquila within the CRUI-CARE Agreement. Open Access Funding provided by Università degli Studi dell'Aquila within the CRUI-CARE Agreement.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: acetaldehyde · asymmetric organocatalysis · design of experiments · optimisation · water

[1] S. H. Xiang, B. Tan, Nat. Commun. 2020, 11, 3786–3790.
[2] a) Green Chemistry: Theory and Practice (Eds.: P. T. Anastas, J. C. Warner), Oxford University Press, New York, 1998; b) P. Anastas, N. Eghbali, Chem. Soc. Rev. 2010, 39, 301–312; c) A. Antenucci, S. Dughera, P. Renzi, ChemSusChem 2021, 14, 2785–2853.
[3] a) A. Carlone, L. Bernardi, Phys. Sci. Rev. 2019, 4, 20180097–20180117; b) D. Wallace, S. Challenger, Z. D. Ding, W. E. G. Osminski, H. Ren, W. D. Wulff, A. A. Desai, M. Munmun, J. P. Scott, M. Alam et al., Org. Process Res. Dev. 2011, 15, 1088–1211.
[4] a) Y. Hayashi, T. Itoh, S. Aratake, H. Ishikawa, Angew. Chem. Int. Ed. 2008, 47, 2082–2084; Angew. Chem. 2008, 120, 2112–2114; b) Y. Hayashi, T. Itoh, M. Ohkubo, H. Ishikawa, Angew. Chem. Int. Ed. 2008, 47, 4722–4724; Angew. Chem. 2008, 120, 4800–4802.
[5] P. Garcia-Garcia, A. Ladépêche, R. Balder, B. List, Angew. Chem. Int. Ed. 2008, 47, 4719–4721; Angew. Chem. 2008, 120, 4797–4799.
[6] H. Ishikawa, M. Homma, Y. Hayashi, Angew. Chem. Int. Ed. 2011, 50, 2824–2827; Angew. Chem. 2011, 123, 2876–2879.
[7] X. Fan, C. Rodríguez-Escrich, S. Sayalero, M. A. Pericás, Chem. Eur. J. 2013, 19, 10814–10817.
[8] G. Giorgianni, V. Nori, A. Baschieri, L. Palombi, A. Carlone, Catalysts 2020, 10, 1296–1300.
[9] General reviews: a) Organic Synthesis in Water (Eds.: P. A. Grieco), Blackie Academic & Professional, London, 1998; b) U. M. Lindström, Chem. Rev. 2002, 102, 2751–2772; c) N. Pinault, D. W. Bruce, Coord. Chem. Rev. 2003, 241, 1–25; d) C.-J. Li, Chem. Rev. 2005, 105, 3095–3165; e) C.-J. Li, L. Cheng, Chem. Soc. Rev. 2006, 35, 68–82; f) C. Pana, Z. Wang, Coord. Chem. Rev. 2008, 252, 736–750; g) M. Gruttadauria, F. Giacalone, R. Noto, Adv. Synth. Catal. 2009, 351, 33–57; h) M. Raj, V. K. Singh, Chem. Commun. 2009, 6687–6703; i) R. N. Butler, A. G. Coyne, Chem. Rev. 2010, 110, 6302–6337; j) N. Mase, C. F. Barbas III, Org. Biomol. Chem. 2010, 8,
[10] N. Mase, K. Watanabe, H. Yoda, K. Takabe, F. Tanaka, C. F. Barbas, J. Am. Chem. Soc. 2006, 128, 4966–4967.

[11] D. G. Blackmond, A. Armstrong, V. Coombe, A. Wells, Angew. Chem. Int. Ed. 2007, 46, 3798–3800; Angew. Chem. 2007, 119, 3872–3874.

[12] a) T. Lundstedt, E. Seifert, L. Abramo, B. Thelin, Å. Nyström, J. Pettersen, R. Bergman, Chemometr. Intell. Lab. 1998, 42, 3–40; b) R. Carlson, Chemometr. Intell. Lab. 2004, 73, 151–166; c) R. Leardi, Anal. Chem. Acta 2009, 652, 161–172; d) S. A. Weissman, N. G. Anderson, Org. Process Res. Dev. 2015, 19, 1605–1633. For reviews on the application of DoE in academia, please see: e) P. M. Murray, F. Bellany, L. Benhamou, D. K. Bučar, A. B. Taborb, T. D. Sheppard, Org. Biomol. Chem. 2016, 14, 2373–2384; f) P. Renzi, M. Bella, Synlett 2017, 28, 306–315; g) B. Benedetti, V. Caponigro, F. Ardini, Crit. Rev. Anal. Chem. 2020, 1–14; h) V. Nori, A. Sinibaldi, F. Pesciaioli, A. Car lone, Synthesis 2022, https://doi.org/10.1055/a-1736–6703.

[13] For selected examples on the application of DoE in academia, see: a) C. Jamieson, M. S. Congreve, D. F. Emi bata-Smith, S. V. Ley, Synlett 2000, 1603–1607; b) C. Jamieson, M. S. Congreve, D. F. Emi bata-Smith, S. V. Ley, J. J. Scicinski, Org. Process Res. Dev. 2002, 6, 823–825; c) M. D. Evans, J. Ring, A. Schoen, A. Bell, P. Edwards, D. Berthelot, R. Nicewonger, C. M. Baldino, Tetrahedron Lett. 2003, 44, 9337–9341; d) L. Veum, S. R. M. Pereira, J. C. Van der Waal, U. Hanefeld, Eur. J. Org. Chem. 2006, 1664–1671; e) T. N. Glasnow, H. Tye, C. O. Kappe, Tetrahedron 2008, 64, 2035–2041; f) M. Silvi, P. Renzi, D. Rosato, C. Margarita, A. Vecchioni, I. Bordacchini, D. Morra, A. Nicolosi, R. Cari, F. Scibba, D. M. Scarpino Schi etroma, M. Bella, Chem. Eur. J. 2013, 19, 9973–9978; g) P. Renzi, C. Kronig, A. Car lone, S. Eröksüz, A. Berkes sel, M. Bella, Chem. Eur. J. 2014, 20, 11768–11775; h) V. Hajzer, P. Alexy, A. Latika, J. Durmis, R. Sebesta, Monatsh. Chem. 2015, 146, 1541–1545; i) A. Ekebergh, C. Lingblom, P. Sandin, C. Wenneräs, J. Mårtensson, Org. Biomol. Chem. 2015, 13, 3382–3392; j) R. Salvo, S. Placidi, A. Carlone, A. Vecchioni, A. Di Sabato, D. C. Buscem i, A. Rossi, A. Antenucci, A. Malkov, M. Bella, J. Org. Chem. 2019, 84, 7395–7404.

[14] C. Palomo, A. Landa, A. Mielgo, M. Diarbide, Á. Puente, S. Vera, Angew. Chem. Int. Ed. 2007, 46, 8431–8435.

Manuscript received: December 21, 2021
Accepted manuscript online: March 1, 2022
Version of record online: March 22, 2022