
This version is available at [https://strathprints.strath.ac.uk/63300/](https://strathprints.strath.ac.uk/63300/)

**Strathprints** is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url ([https://strathprints.strath.ac.uk/](https://strathprints.strath.ac.uk/)) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

---

The Strathprints institutional repository ([https://strathprints.strath.ac.uk/](https://strathprints.strath.ac.uk/)) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.
Scalable total synthesis and comprehensive structure–activity relationship studies of the phytotoxin coronatine

Mairi M. Littleson,1 Christopher M. Baker,2 Anne J. Dalencon,2 Elizabeth C. Frye,2 Craig Jamieson,1 Alan R. Kennedy,1 Kenneth B. Ling,2 Matthew M. McLachlan,2 Mark G. Montgomery,2 Claire J. Russell,2 and Allan J. B. Watson1*

Natural phytotoxins are valuable starting points for agrochemical design. Acting as a jasmonate agonist, coronatine represents an attractive herbicidal lead with novel mode of action and has been an important synthetic target for agrochemical development. However, both restricted access to quantities of coronatine and a lack of a suitably scalable and flexible synthetic approach to its constituent natural product components, coronafacic and coronamic acids, has frustrated development of this target. Here, we report gram-scale production of coronafacic acid that allows a comprehensive structure–activity relationship study of this target. Biological assessment of a >120 member library combined with computational studies have revealed the key determinants of potency, rationalising hypotheses held for decades, and allowing future rational design of new herbicidal leads based on this template.

Food security is recognised as a global concern due to a growing population increasing food consumption, and various factors that diminish production, such as arable land desertification and infestation by pests.1 The requirement for effective herbicides for improved weed control and crop yield is essential to meet global food demand.2 Resistance to traditionally used herbicides is an increasing problem and there is increasing regulatory pressure on the current crop protection products available to the farmer,3 which has resulted in a pressing need for the development of novel and safe agrochemicals with new modes of action (MOA).4 In this regard, natural products are valuable starting points for agrochemical design as they often allow the targeting of distinct biological space;5,6 mesotrione is a pertinent example of how natural assets can be leveraged to new herbicidal agents.7

Coronatine (COR, 1; Figure 1a) is produced by several strains of Pseudomonas syringae and has attracted attention both synthetically and biologically due to its chemical structure8 and promising phytotoxic properties.9 Known to be a non-host specific agonist of the active plant hormone (+)-7-iso-jasmonoyl-L-isoleucine (JA-Ile; Figure 1a),10 1 has been found to induce a range of stress-response and defence-related activity in plants by interaction with the jasmonate receptor COR-insensitive 1 (COI1), and inducing phytotoxic effects through activation of the JA-signalling pathway.11 Through this biological pathway, COR has been reported to exhibit a range of phytotoxic activity across several plant species, including leaf tissue chlorosis12 and senescence,13 root stunting,14,15 increased ethylene production,16 production of defence-related secondary metabolites,17 induction of hypertrophy,18 and stomatal opening.19 The jasmonate receptor represents a novel MOA not currently exploited by commercial phytotoxins, and as such the development of a COR-based herbicide is highly desirable.4

COR (1) is composed of two constituent natural products: the cis-fused 5,6-bicyclic polyketide core unit, coronafacic acid (CFA, 2), coupled to an isoleucine (Ile)-derived amino acid, coronamic acid (CMA, 3), through an amide linkage (Figure 1a).8 Since the discovery of COR 40 years ago,8 considerable synthetic efforts have been directed towards the synthesis of both 2 and 3.20-57 However, access to useful quantities of 1 either by bacterial fermentation or synthetically, has been challenging and has afforded only relatively limited structure-activity relationship (SAR) studies (Figure 1b, 1c; vide infra).41,58-69 In addition, 2 has long been viewed as a principal component from which the bioactivity of 1 is derived; however, to date, the reported cumulative production of 2 by chemical synthesis is less than 1 g over nine separate synthesis campaigns. Moreover, hydrolysis of natural 1 is both atom inefficient and prohibitively costly.64,65 Lastly, there is no substantive quantitative biological data across the intended targets (weed species).

1Department of Pure and Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow, G1 1XL, UK.
2Syngenta, Jealott’s Hill International Research Centre, Bracknell, Berkshire, RG42 6EY, UK. *e-mail: allan.watson.100@strath.ac.uk
To summarise reported SAR data (Figure 1b, 1c), both CFA and CMA moieties confer phytotoxic activity separately, however, this is greatly enhanced when the components are coupled to give the parent structure.60 With regard to the core moiety, it is known that the cis-stereochemistry of the ring junction is important for biological activity, mimicking the side chain configuration of JA-Ile.10,65,70 Substitution at the C6 position has also been shown to be required for activity in potato tuber inducing assays.61 Reduction of the carbonyl moiety has been reported to lead to reduced volatile inducing activity in rice leaves with respect to COR,63 however; there have been reports of retained activity of this compound. The analogue where the ſ,ȫ-unsaturated amide has been reduced to afford the fully saturated 6,5-bicycle has been reported and found to be highly active in volatile emission assays.63 With regard to the amino acid portion, it has been widely reported that the free carboxyl terminus of the amino acid is required for maximal activity,60 and substitution which retains the S-stereochemistry of CMA at the ſ-carbon is important for activity, as has been demonstrated through the synthesis of other COR stereoisomers.60 Tolerance for alternative amino acid substitution with both natural and non-natural amino acids has been demonstrated, however, at the outset of this study, a complete SAR for this portion of the molecule was unclear.65,68

Figure 1 | (+)-7-iso-JA-Ile (JA-Ile) and coronatine (1): structure and route design plan. Structural similarities between the natural bioactive ligand JA-Ile and 1 are highlighted. Coronatine can be considered as comprising of two component parts; the bicyclic core, coronafacic acid 2, and amino acid moiety, coronamic acid 3.

To enable a comprehensive SAR exploration, a scalable, tractable, and flexible synthesis of 2 is required. Herein, we report a collaborative industry-academia approach71 that has provided a practical, gram-scale synthesis of (+)-2, enabling the subsequent preparation of a >120 member library of analogues of 1. Access to grams of (+)-2 has allowed array synthesis of
amide analogues of 1 to explore the binding region around the CMA motif and the flexibility of the synthetic route has allowed SAR charting around the CFA unit, both through single point changes to the bicyclic structure and more significant structural modifications of the core scaffold. This library has been assessed for herbicidal activity against several weed species and, using computational modelling of the active site, has allowed the principal drivers of potency to be revealed, allowing a more rational approach to herbicide discovery using this template.

**Results and Discussion**

From the outset, our synthetic strategy was focussed on scalability, to enable preparation of a library of amide analogues of 1 (i.e., variation of the CMA region), and flexibility, to allow SAR interrogation of 1 (i.e., the CFA region). It was our intention that the synthetic campaign and subsequent biological evaluation of COR analogues would inform the design and synthesis of structurally less complex COR derivatives, ideally with the retention or enhancement of phytotoxic potency. Based on the lack of robust SAR data, as the largest fragment, 2 has been assumed to be the key driver of the potency of 1. With the total production of synthetic 2 less than one gram over decades of investigation, access to quantities of this fragment suitable for SAR interrogation has been the most significant challenge in developing COR as an agrochemical lead. As such, our approach had synthetic expediency and flexibility embedded from the outset.

The cyclohexene scaffold of 2 clearly codifies for an intramolecular Diels-Alder (IMDA) disconnection (Figure 1a) and, indeed, this approach has been used in previous syntheses of (±)-2. The requisite diene would be accessed by the aldol disconnection employed by Charette. This ultimately provides a convergent synthesis using two fragments that are readily modifiable and therefore impart the flexibility required of the SAR objectives. The flexibility and mapping strategy of the fragment approach is shown in Figure 1d. Access to 3 was not problematic and was generated via a modified variant of an established dialkylation process (Figure 1a).

It has been reported that (+)-1 is significantly more potent than (−)-1 with respect to stomatal opening activity. Accordingly, the SAR associated with each of the stereoisomers was an important aspect of the investigation; however, investment into asymmetric routes at this stage was an inefficient use of resource where the SAR was largely unknown. Based on this, we elected to prepare all compounds as racemates to enable expedient analogue synthesis and, after initial triage in the biological assays, assess single enantiomers by separation of the racemic material.

Our optimised scalable route to (±)-2 is shown in Figure 2. The aldehyde fragment required for the aldol addition (7) was obtained in five steps and 37% overall yield from the readily available 1,4-butane diol 4. Mono-protection of 4 with dihydropyran proceeded in high yield, allowing isolation of alcohol 5. Swern oxidation afforded the corresponding aldehyde, which was immediately reacted with vinylmagnesium bromide and quenched with acetic anhydride to give 6 in 63% over 2 steps. THP deprotection followed by further Swern oxidation gave access to aldehyde 7 on multigram scale. The route to this fragment generated in excess of 44 g of 7 for this campaign.
HODHO

10. Aldol addition product 9 previously reported requirement of highly elevated temperatures and a pressure-sealed vessel to allow the cyclization of this desired class of triene (We had initially viewed the IMDA reaction as being particularly challenging with respect to the scalability of the route. The we found that stereospecific dehydration of the subsequent dehydration and IMDA reaction. ultimately allowing access to over 50 g of aldol adduct reversed selectivity, in favour of the been reported, junction.

analogues (Figure 1d). Variation of the ester used in the aldol addition (The flexibility offered by this synthetic sequence allowed single point changes to allow the synthesis of a series of CFA structural modification at C7, leading to regioisomeric and ring expanded (decalin) cores.

With access to sufficient quantities of (±)-2 as well as derivatives with variation of the core CFA template, we prepared a library of analogues to prosecute the SAR objectives (Figure 3). It has been reported that the enzyme responsible for the linkage of 2 and 3, coronafacate ligase, has a degree of tolerance around the amino acid structure, as evidenced through.
the isolation of several \(N\)-coronafacoyl compounds alongside COR.\(^{77-80}\) Accordingly, we determined it appropriate to prepare a range of coronafacoyl amide analogues, maintaining \((\pm)\)-2 as the common core unit (Figure 3a). To ensure breadth in our SAR study, a variety of natural and non-natural amino acids were incorporated using straightforward HATU-mediated coupling on the amino acid methyl esters, followed by hydrolysis under basic conditions to afford the desired acidergic compounds (16-37).\(^{68}\) For the CFA analogues with single point changes and variation of the carbonyl unit, we prepared both the CMA- and \(L\)-Lle-derived \(N\)-coronafacoyl amides using the same amidation procedure (38-45), including the decalin and aromatised analogues (43-45). Stereoisomers of interest following initial triage (\textit{vide infra}) were separated by chiral preparative HPLC and evaluated (several examples shown: 46-51). Lastly, two arrays were generated using automated synthesis with (i) variation of CMA to a range of nonnatural amino acids on the aromatised CFA core and (ii) variation of CFA to non-CFA acids on the CMA residue (not shown, see ESI).

Overall, 127 analogues of COR were successfully prepared and assessed using a raft of phenotypic assessments against several weed species. A selection of this library is presented in Table 1.
Figure 3 | Representative examples of COR analogue synthesis. a, HATU coupling of (±)-2 and amino acid methyl esters, which were then hydrolysed to afford the free-acids. HATU, 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate. b, Alternative core coupling to (±)-CMA and L-Ile. c, CFA core oxime and core analogues.
Table 1 | Biological data. Scoring of active compounds from SAR screening. In initial greenhouse screening (GH1), compounds are assessed for pre- and post-emergence activity against four weed species, and scored visually for % phytotoxicity (0-100, where 100 is complete control of the target and 0 is no control). % phytotoxicity is colour coded: no colour: inactive compound, yellow: 40-50% phytotoxicity, pale green: 60-70% phytotoxicity, dark green: 80-100% phytotoxicity. Key compounds are outlined. Test species: *Amaranthus retroflexus* (AMARE), *Lolium perenne* (LOLPE), *Stellaria media* (STEME), *Digitaria sanguinalis* (DIGSA). DS = desiccation, GI = germination inhibition, NC = necrosis, NT = not tested, ST = stunting.

<table>
<thead>
<tr>
<th>Compound</th>
<th>AMARE</th>
<th>LOLPE</th>
<th>STEME</th>
<th>DIGSA</th>
<th>Symptom</th>
<th>AMARE</th>
<th>LOLPE</th>
<th>STEME</th>
<th>DIGSA</th>
<th>Symptom</th>
</tr>
</thead>
<tbody>
<tr>
<td>(+)-1</td>
<td>90</td>
<td>60</td>
<td>NT</td>
<td>90</td>
<td>ST/DS</td>
<td>80</td>
<td>90</td>
<td>NT</td>
<td>100</td>
<td>ST/DS</td>
</tr>
<tr>
<td>(-)-1</td>
<td>40</td>
<td>0</td>
<td>50</td>
<td>60</td>
<td>NC/ST</td>
<td>70</td>
<td>40</td>
<td>70</td>
<td>80</td>
<td>NC/ST</td>
</tr>
<tr>
<td>(-)-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>(-)-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ST</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>ST</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>20</td>
<td>NC/ST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NC/ST</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ST</td>
<td>50</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>ST</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>20</td>
<td>NC/ST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NC/ST</td>
</tr>
<tr>
<td>36</td>
<td>50</td>
<td>40</td>
<td>0</td>
<td>60</td>
<td>ST</td>
<td>40</td>
<td>30</td>
<td>50</td>
<td>50</td>
<td>ST</td>
</tr>
<tr>
<td>38a</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>80</td>
<td>NC/ST</td>
<td>80</td>
<td>60</td>
<td>80</td>
<td>80</td>
<td>NC/ST</td>
</tr>
<tr>
<td>41a</td>
<td>30</td>
<td>20</td>
<td>30</td>
<td>60</td>
<td>GI/ST</td>
<td>0</td>
<td>20</td>
<td>80</td>
<td>0</td>
<td>GI/ST</td>
</tr>
<tr>
<td>43a</td>
<td>30</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>GI/ST</td>
<td>30</td>
<td>60</td>
<td>40</td>
<td>80</td>
<td>GI/ST</td>
</tr>
<tr>
<td>45a</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>100</td>
<td>NC/ST</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>40</td>
<td>NC/ST</td>
</tr>
<tr>
<td>39b</td>
<td>20</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>ST</td>
<td>30</td>
<td>40</td>
<td>70</td>
<td>0</td>
<td>ST</td>
</tr>
</tbody>
</table>

The obtained biological data has allowed mapping of SAR around the natural product scaffold (see ESI for the full data set):

1) With regard to the amino acid substituent, there appears to be little tolerance for structural modification away from the CMA motif; variation of this region produced inactive compounds, or analogues of significantly reduced potency (e.g., compare 1 vs. 16, 21, 32, 35, and 36). This is likely due to the increased bulk in the amino acid region preventing binding to the COI1-JAZ co-receptor. Typically, moderate levels of phytotoxicity were observed with quaternary substituted amino acids, e.g., 32, aligning this portion of the molecule more closely to the structural features of CMA. In agreement with previous reports, we observed that S-stereochemistry at the α-carbon is important for activity, as demonstrated through comparison of 32 and the respective R-configured analogue (31) which is inactive (See ESI). These results point towards the importance of the amino acid residue to achieve significant levels of potency, and more specifically the importance of a structurally intact CMA unit.

2) Several of our core-modified CMA-conjugates showed significant phytotoxic activity. As previously mentioned, it is known that the cis-stereochemistry of the fused ring-junction is important for biological activity. Compound 43a, featuring the cis-decalin core, showed good levels of activity, however the trans-decalin structure 44a was inactive, implying that the cyclopentenone ring is tolerant of variation but that the cis-ring junction is required; however, we cannot rule out that the lack of activity is due to the alternative carbonyl placement (*vide infra*). Substitution at the C6 position is required for activity (41a vs. 42a). The alcohol derivative, 38a, showed good levels of activity, while oxime compound 39b was weakly active, suggesting that variation to the ketone moiety is tolerated; however, this could be a function of metabolism to
the ketone in planta. The CMA analogue with aromatic core, 45a, showed significant levels of activity (Table 1). The activity of this compound also demonstrates the importance of the CMA moiety, as our array analogue synthesis maintaining the aromatic core moiety failed to deliver analogues of significant activity (See ESI). This result demonstrates the potential for CFA simplification with the retention of phytotoxic activity if the CMA moiety is maintained, a highly desirable outcome of our SAR study as it renders the preparation of 2 unnecessary and replaceable with simpler analogues. The general inactivity of our L-Ile conjugates (38-45b; see ESI for the complete data set) with CFA analogues in comparison with their CMA substituted counterparts again bolsters importance of the CMA residue. Further attempts to simplify the CFA scaffold with significant structural modifications and with retention of the CMA amino acid were largely unsuccessful (see ESI).

In light of these results, compounds 43a and 45a were separated into their component enantiomers by chiral HPLC, and the single enantiomers assessed for phytotoxicity. Moderate activity levels were observed; however, in both cases activity levels obtained were weaker than (+)-1 (see ESI). The activity observed from (+)-45a in comparison with the complete inactivity of (−)-45a demonstrates that the potency of 45a is derived from only one enantiomer. This is further demonstrated in the variation of activity levels from the enantiomers of 43a. Of the stereoisomers with the homologated core (48-51), none were more effective than (+)-1, with trends as expected (see ESI).

To rationalise these observations, we undertook molecular modelling of COR in the active site of COI1 (Figure 4). Both the crystal structure and the computational calculations indicate that the binding of COR is chiefly driven by the formation of strong H-bonding interactions with three arginine residues in the active site (ARG85, ARG348, and ARG409) from the amide carbonyl and CMA-based carboxylate group. Favourable interactions also come from the formation of a H-bond with the CFA cyclopentanone carbonyl group and TYR444 as well as a number of hydrophobic interactions, including the ethyl unit of CFA with the lipophilic region consisting of LEU91, PHE89, and ALA86, and insertion of the cyclopropyl-ethane tail into a hydrophobic pocket consisting of ALA384, VAL385, and TYR386.

To rationalise the observed generally detrimental impact of variation of CMA to L-Ile (and the majority of other amino acids), we compared the docking of 1 and 16 (Figure 4, structures (a) and (b)). Docking calculations were performed against the binding site of subunit B from PDB structure 3OGK, using the program Glide in “standard precision” mode. In this approach, the receptor was treated as rigid, while ligands were docked with full conformational flexibility, including sampling of ring conformations, and an additional energy penalty was included to discourage formation of non-planar amide bonds. Following the initial docking, the five highest-scoring poses for each ligand were identified using the Glide “docking score”, a metric that accounts for both the ligand-receptor molecular mechanics interaction energy and the ligand strain energy: these five poses were then subjected to a full force field minimization, with the resulting, minimized, poses re-scored and only the single highest-scoring docked pose for each ligand retained for analysis. This model indicates that inserting the branched alkyl chain of the Ile unit of 16 into the hydrophobic pocket formed by ALA384, VAL385, and TYR386 is much less favourable than for the CMA unit of 1. The unfavourable binding observed results from multiple steric clashes between the alkyl chain of the Ile unit and the protein. In addition, we expect that the increased flexibility found in Ile vs. CMA would result in a higher entropic cost of binding, negatively impacting on affinity.

With regards to the more tolerant CFA region, the principal interactions are the H-bonding from the CFA cyclopentanone carbonyl group and TYR444, and the hydrophobic interactions from the ethyl unit of CFA with the lipophilic region consisting of LEU91, PHE89, and ALA86. In addition, ARG496 is proximal to the carbonyl of the CFA cyclopentanone; examination of the crystal structure suggests that this is not at a range to form a H-bond; however, we believe the proximity allows this interaction in the dynamic setting. Aromatised compounds (45a) are effective since this placement of functional groups and key interactions are conserved. Removal of either/both of these functional groups induces penalties but so long as they are maintained, there is considerable flexibility in this region, explaining the activity of the decalin and stereoisomeric compounds (38a, 41a, 45a).
Considering all of these data, key potential lead structures that could form the basis of an optimization program are 32, 36, 38a, 43a, and 39b (Table 1). The SAR information codified by structures is aligned with the ligand-based toxicophore model described in Figure (c), below based on consideration of the biological data obtained.
Conclusion
In conclusion, an extensive SAR investigation around the phytotoxic natural product COR has been carried out. The multigram-scale synthesis of (±)-CFA enabled array synthesis of a broad scoped amino acid screen, using CFA as the common core unit. Investigation of CFA modifications through alterations to the bicyclic structure has allowed for a mapping of SAR around the core-motif. Typically, we have observed that although incorporation of alternative amino acids onto the CFA core can result in low levels of phytotoxic activity, the key determinant of activity appears to be the CMA moiety, and a greater tolerance for modification has been observed around the CFA core. These results demonstrate that further studies featuring the CMA unit and CFA replacements may be beneficial, and we would suggest that future research efforts in the area focus on these derivatives.

References

Acknowledgements

We thank the EPSRC UK National Mass Spectrometry Facility at Swansea University for analyses, the University of Strathclyde for PhD studentship (M. M. L), and Syngenta for financial and chemical support. We thank Dr M. E. Watson (University of Strathclyde) for supplying amino acids used to prepare 31 and 32. We thank Thorsten Platz and Russell Ellis (Syngenta) for assistance with compound purification and Matthew Plane and David Pearce (Syngenta) for assistance with robot-enabled parallel synthesis.