REVIEW

Deep-sea discovery and detective work: towards solving the hemicalide structural enigma through computational NMR analysis and stereocontrolled synthesis

Nelson Y. S. Lam* and Ian Paterson*[a]

Dedicated to Prof. Cesare Gennari in celebration of his 70th birthday

[a] Dr. N. Y. S. Lam, Prof. Dr. I. Paterson
Yusuf Hamied Department of Chemistry,
Lensfield Road, Cambridge, CB2 1EW, United Kingdom
Homepage: http://www-paterson.ch.cam.ac.uk/

Abstract: The marine natural product hemicalide displays potent anticancer activity. While detailed NMR experiments established the planar carbon skeleton, its full configurational assignment proved enigmatic. This Review summarises ongoing synthetic efforts and NMR studies on hemicalide, which have succeeded in reducing the number of possible stereoisomers from more than one million to a more manageable eight structures, setting the stage for a focused total synthesis campaign.

1. Introduction

Marine organisms are a valuable source of novel biologically active compounds and potential drug candidates, where their diverse range of bioactivities have resulted in the clinical validation of many promising compounds.[1] In particular, cytotoxic natural products isolated from marine organisms have provided a rich source of novel chemotherapeutic agents, especially in an anticancer context.[2] Postulated to be a source of chemical defense for sessile organisms, these marine-derived secondary metabolites have evolved to be highly efficient at target binding.[3,4] Their potent biological activity typically only results in minute quantities generated by the microbial producer, which consequently can result in a vanishingly low isolation yield, insufficient for comprehensive structural or mode of action elucidation. Their scarcity, coupled with the structural complexity that often typifies these natural products, renders full or partial chemical synthesis a necessity to solve their 3D molecular structures—a prerequisite for downstream studies in the context of their potential development as novel therapeutic agents.[5]

First appearing in the patent literature in 2011, hemicalide (**1**, Figure 1) is no exception to the challenges outlined above. The complex marine polyketide was isolated from the sponge *Hemimycale* sp. collected around the Torres Islands of Vanuatu by a team supported by the CNRS-Pierre Fabre Laboratories and the Institut de Recherche pour le Développement. Bioassay-guided fractionation from 5 kg of *hemimycale* sp. afforded 0.5 mg of the active component hemicalide.[6] From this material, mass spectrometry and extensive 1D and 2D NMR spectroscopic studies enabled the derivation of a planar structure consisting of an elaborate, highly oxygenated, 46-carbon chain backbone. An account of this remarkable feat of structural assignment is provided by Marcourt and Massiot with a companion article in this issue.[7] In brief, their detailed interrogation of the NMR data revealed four major stereoclusters: a contiguous polypropionate stereohexad (C8-C13), an α,β-dihydroxylactone-containing stereopentad (C18-C24), a polyacetate-derived stereotetrad (C27-C32) and an α-hydroxylactone-containing stereopentad (C36-C42).[6] Alongside an isolated C45 hydroxyl-bearing stereocentre, this structural analysis indicated that hemicalide possesses 21 stereocentres. Unfortunately, further derivatisation or degradation studies to ascertain both the relative and absolute configuration were not conducted owing to the extremely meagre material supply, leaving all 21 stereocentres unassigned.

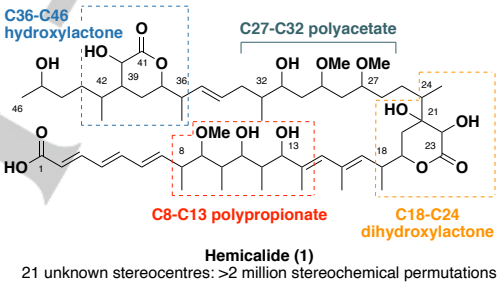
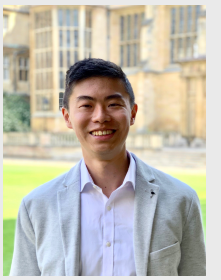
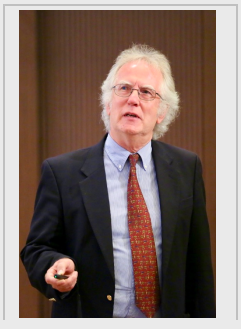


Figure 1. Planar structure of hemicalide (**1**) highlighting major stereoclusters

From the submilligram sample of isolated material, however, hemicalide was found to be highly cytotoxic, registering picomolar IC_{50} levels across a number of cancer cell lines. Preliminary mechanistic studies pointed towards a novel mode of microtubule destabilisation.[6] At a dose of 5-50 nM, immunohistochemical labeling of α-tubulin in HeLa cells resulted in no microtubular skeleton for cells in interphase. Mitotic cells additionally were found arrested in prometaphase without microtubules.[6] Other well-known antimitotic natural products, such as paclitaxel, epothilone B and discodermolide inhibit microtubule formation during the metaphase/anaphase boundary.[8] These observations suggest that hemicalide inhibits the formation of the microtubular network in a novel manner; a valuable property as a potential anticancer agent in light of rising drug resistance towards existing cancer chemotherapeutics.[9] Despite the promising biological effects exhibited by hemicalide from nascent investigations, further elucidation of its mode of action and, critically, subsequent preclinical development cannot occur if its precise 3D molecular architecture is left undetermined. In this case, the 21 unassigned stereocentres of hemicalide produce 2^{21} possible stereoisomers. Excluding enantiomers, this results in *over one million (1048576) possible diastereomeric permutations for the natural product*.



Nelson Lam is a graduate of the University of Auckland, where he completed a BMus and BSc (Hons), completing his Honours project with Prof. Christian Hartinger in 2014. He completed his PhD at the University of Cambridge (2019) under the supervision of Prof. Ian Paterson as a Woolf Fisher Scholar, working on the synthesis of stereochemically ambiguous polyketide natural products. Following postdoctoral research as a Lindemann Fellow with Prof. Jin-Quan Yu (Scripps Research), Nelson is currently a Junior Research Fellow in Chemistry at Trinity Hall, Cambridge developing novel selective transformations harnessing non-covalent interactions.

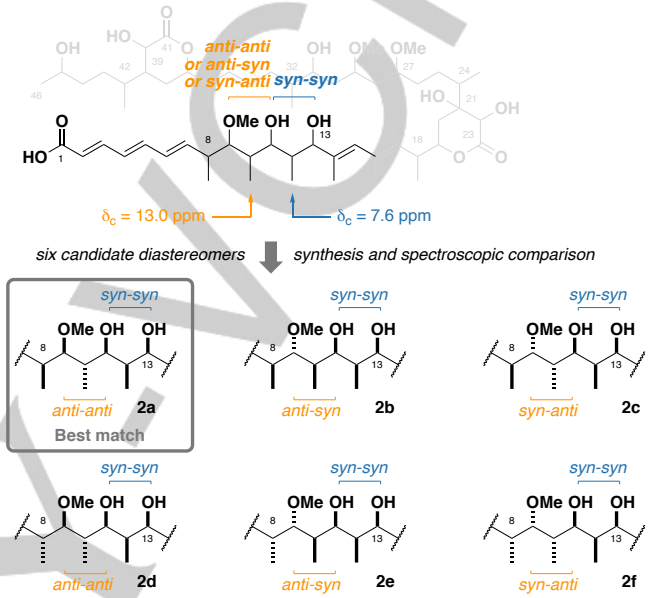


Ian Paterson received his BSc (Hons) degree from the University of St. Andrews and PhD from the University of Cambridge, working with Prof. Ian Fleming. After a postdoctoral period with Prof. Gilbert Stork at Columbia University, he joined the faculty at University College London. In 1983, he returned to Cambridge, where he is now Emeritus Professor of Organic Chemistry and a Fellow of Jesus College. His research interests are centred on novel synthetic methods for the efficient control of stereochemistry and the total synthesis of bioactive natural products, particularly anticancer agents. His research achievements have been recognised by various awards, and he is a Fellow of the Royal Society and the Royal Society of Edinburgh.

The alluring combination of hemicalide's natural scarcity and impressive biological properties, encased in a complex yet intractably enigmatic structure, led to our involvement in pursuing its complete configurational assignment,[10–13] and ultimately targeting its total synthesis. Armed with the 1D ^{1}H and ^{13}C NMR data listed in the patent application as the only leads in this "stereo-mystery" detective case, we embarked on this adventurous journey of discovery in 2013. This Review outlines how multipronged *ab initio* NMR studies, in synergy with carefully planned fragment syntheses, have enabled the confident assignment of the relative configuration of 19 of the 21 stereocentres in hemicalide. Together with all possible remaining permutations between the stereoclusters, these cumulative studies have enabled the possible set of diastereomers of hemicalide to be narrowed down to eight, setting the stage now for a focused total synthesis campaign.

2. Assignment and Synthesis of the C8-C13 Stereohexad

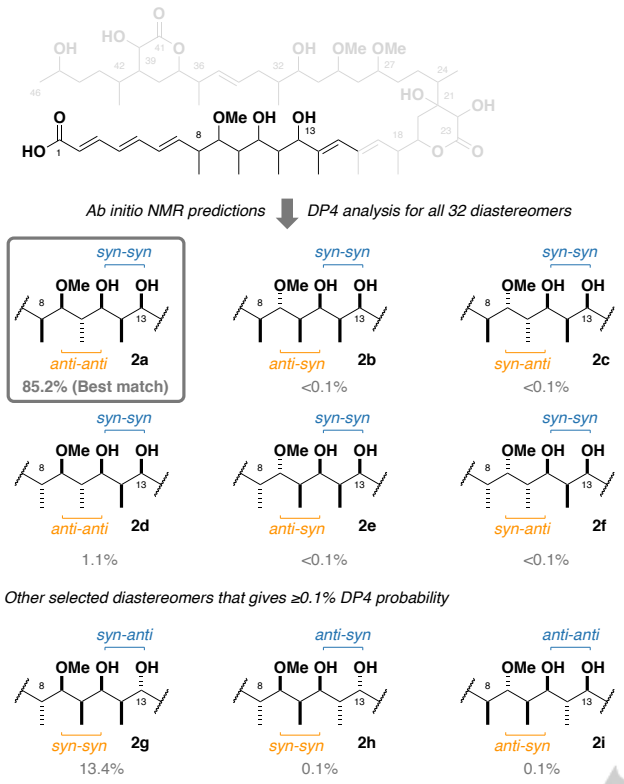
Our involvement towards the stereochemical determination of hemicalide follows on the back of insightful studies conducted by Ardisson *et al*. Operating under the initial assumption that remote stereocentres minimally affect the chemical shifts of isolated stereoclusters, the Ardisson group commenced their investigations. Within the C8-C13 stereohexad, an astute observation of an unusually upfield shifted Me12 in the ^{13}C NMR spectra alluded to a *syn-syn* relationship for the C10-C13 stereotriad.[14,15] The same phenomenon was not observed for Me10, ruling out a *syn-syn* relationship for the substituents across C9-C11 and indicating that the stereotriad likely contains a *syn-anti*, *anti-syn* or an *anti-anti* configuration (Scheme 1). These two key observations reduced the number of diastereomers to consider from 32 down to six. Following the synthesis of all six candidate diastereomers (**2a-2f**) and detailed NMR spectroscopic comparisons with the natural product data, Ardisson concluded diastereomer **2a** was the best match to hemicalide.



Scheme 1. Diagnostic methyl ^{13}C NMR resonances narrows the polypropionate region down to six diastereomers (**2a**-**2f**) ; synthesis and spectroscopic comparison revealed **2a** being the best match for the natural product

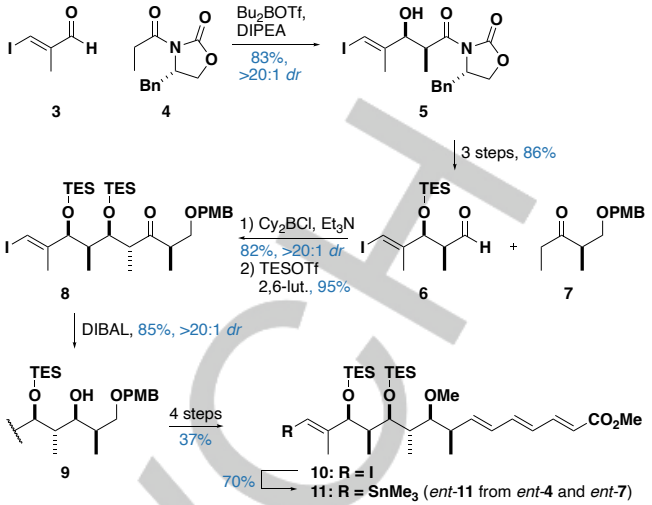
During this time, advances in quantum mechanical NMR chemical shift predictions led to the development of the DP4 algorithm, enabling the assignment of a probability score to a given candidate diastereomer being correct relative to the experimental spectra. This is achieved through *ab initio* calculations of ^{1}H and ^{13}C NMR chemical shifts for each candidate diastereomer. Comparing the differences between the theoretical chemical shifts with the experimental data *via* the DP4 algorithm gives rise to an overall probability for the candidate diastereomer fitting the experimental data. Operating under the same assumption that remote stereocentres minimally affect the chemical shifts of a region under investigation, Smith and Goodman examined this truncated fragment *via* a Gauge-including atomic orbital (GIAO)-NMR shift prediction methodology (termed DP4f for virtual fragments).[16] Through a comprehensive analysis of all 32 candidate diastereomers (a representative set shown in Scheme 2; **2a-2g**), they independently arrived at the same conclusion regarding the most likely relative configuration of the stereohexad in the C1-C15 fragment in **2a**.[16] At this point, spectroscopic validation provided in the form of stereodefined synthesis by the Ardisson group, and independently verified through *ab initio* computational means by Smith and Goodman enabled the determination of six stereocentres (C8-C13) of hemicalide. This preliminary detective work served to reduce the number of stereochemical permutations down to 2^{16} (15 unknown

stereocentres + unknown absolute configuration of C8-C13 region), or 32768 (2^{15}) excluding enantiomers.



Scheme 2. DP4 analysis of all 32 diastereomers conducted by Smith and Goodman (representative structures **2a-2i**) provided orthogonal corroboration for the assignment in **2a**

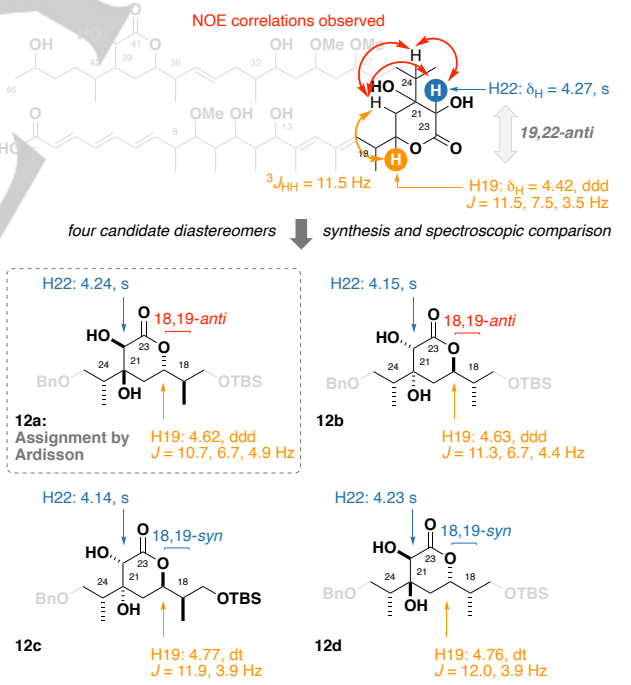
Building on these secure foundations, we then embarked on a focused synthetic campaign, paying particular attention towards developing a route that would enable the facile synthesis of both candidate enantiomers required to probe the diastereomeric relationship between this fragment with the remainder of the natural product.[17] The synthesis of the C1-C15 fragment began with an Evans *syn*-aldol reaction[56] of **3** with **4** to generate adduct **5** (Scheme 3).[57] Subsequent manipulations revealed aldehyde **6**, which engaged with ketone **7** *via* a 1,4-*syn* boron-mediated aldol reaction[58] to give **8**. Following protection, the final stereocentre was set through a diastereoselective reduction to generate fragment **9.** The full C1-C15 vinyl iodide **10** was then generated in four steps, and could be transformed to the corresponding vinyl stannane **11** to ensure flexibility in the anticipated downstream fragment union step. The uniformly high diastereoselectivity of this route permitted the synthesis of both enantiomers of the C1-C15 fragment (**10**, **11** and *ent*-**10,** *ent*-**11**).



Scheme 3. Synthesis of vinyl iodide **10** and vinyl stannane **11**

3. DP4-Enabled Reassignment of the C18-C24 Region

For the C18-C24 region, NOE and coupling constant analyses strongly suggested a 19,22-*anti* relationship, enabling Ardisson to narrow down the set of possible diastereomers to four candidates.[18]

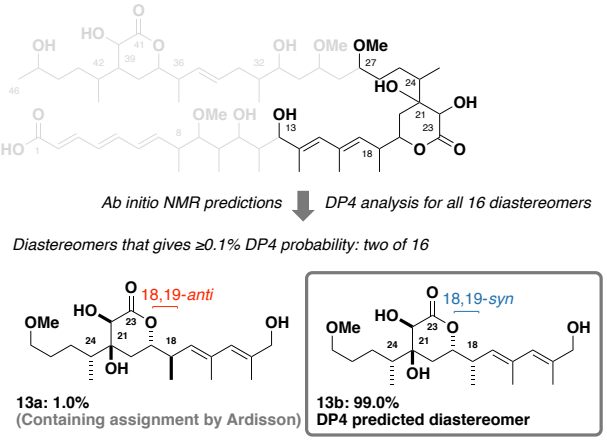


Scheme 4. Ardisson's assignment of the C18-C24 dihydroxylactone

Synthesis of all four diastereomers **12a-d** and comparing their NMR data to hemicalide led to the conclusion that the diastereomer **12a**, containing the 18,19-*anti* relationship, was the most likely candidate (Scheme 4). Ardisson's 18,19-*anti* assignment hinged upon a qualitative comparison of the H19 proton signal in their model fragment, reasoning that the 18,19-*anti* diastereomer, which exhibited a ddd multiplicity, was a closer

match for the experimental hemicalide data compared to the 18,19-*syn* relationship.

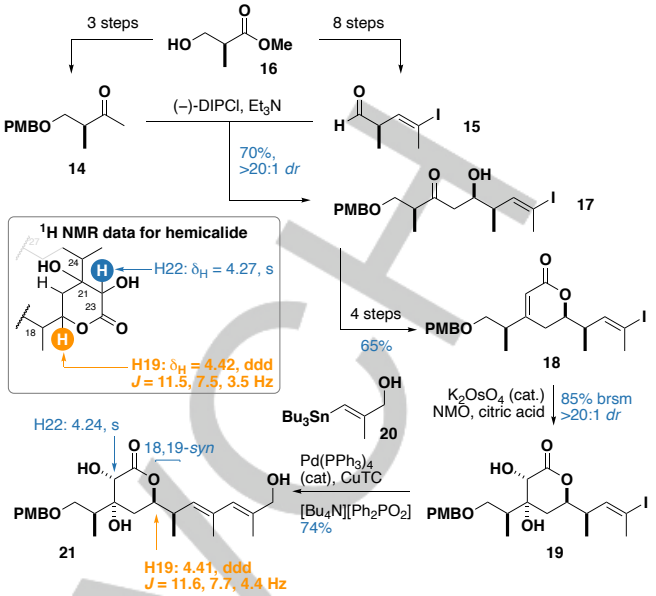
Analogously, we applied the DP4 methodology to ascertain the relative configuration within this region. Using a model diene fragment, *ab initio* NMR calculations and analysis of all 16 diastereomers revealed only two candidates **13a** and **13b** that scored meaningful probabilities for consideration. Interestingly, however, DP4 calculations identified diastereomer **13b** (Scheme 5) as the most likely candidate with very high confidence (99%),[19] which contradicted Ardisson's assignment with its key 18,19-*syn* relationship (**13a**). Notably, the latter assignment was considered highy improbable through DP4 analysis (1%).



Scheme 5. DP4 predicted diastereomer for the C18-C24 dihydroxylactone

Seeking to resolve these contradictory observations, we sought to corroborate the DP4 prediction through a targeted synthesis of a model dihydroxylactone truncate bearing an 18,19-*syn* configuration (Scheme 6). This commenced with a boron-mediated aldol reaction between ketone **14** and aldehyde **15**—derived from the same Roche ester **16**—to form adduct **17**. A four-step sequence (protection, aldol addition, lactonisation and dehydration) gave **18**. This was then dihydroxylated to install the remaining two stereocentres in **19**. From here, a Stille cross-coupling with stannane **20** afforded model fragment **21** containing the 18,19-*syn* configuration, which was a much closer spectroscopic match to hemicalide than Ardisson's original proposed assignment in **12a**. In particular, the diagnostic H19 proton signal possessed near identical chemical shift, multiplicity and coupling constant measurements. Independently, Ardisson and Cossy *et al.* arrived at the same conclusion as indicated in the reassignment to a 18,19-syn configuration in their later publication.[20]

Assisted by DP4 calculations, this synthetic effort provided a more confident elucidation of this region, securing the identity of five more stereocentres leading to 11 of the 21 stereocentres assigned. This then reduces the number of possible candidate diastereomers down to 2^{12} (10 unknown stereocentres + unknown absolute configurations for the C8-C13 and C18-C24 regions), or 2048 (2^{11}) excluding enantiomers. Notably, the route to assemble intermediate **19** from the chiral pool precursor **16** enabled the ready generation of the enantiomer, which was executed *en route* towards solving the configurational relationship between the C1-C13 and C18-C24 regions.



Scheme 6. Synthesis of model truncate **21** for NMR comparison revealed that hemicalide possesses the 18,19-*syn* configuration

4. Synthesis-Enabled Assignment Between the C1-C13 and C18-C24 Regions

With the confident assignment of the C1-C13 and the C18-C24 regions, the next unknown was the diastereomeric relationship between the two stereoclusters. Initial forays into this were undertaken by the Ardisson and Cossy groups, who first reported a synthesis of the C1-C25 truncate **22** (Figure 2) featuring the initially-assigned 18,19-*anti* configuration in the dihydroxylactone fragment.[21] Following our reassignment, the Ardisson and Cossy groups disclosed a revised synthesis of their C1-C27 fragment to contain the updated 18,19-*syn* configuration in **23**.[20] Analysis of the NMR deviations between truncates **22** and **23** (containing the revised 18,19-*syn* configuration) provided further affirmation for our dihydroxylactone reassignment (Table 1), with **23** bearing the revised stereochemistry resulting in a far closer match to the spectroscopic data for hemicalide (**23**: $\Sigma|\Delta\delta_H|$ = 0.26 ppm; $\Sigma|\Delta\delta_C|$ = 3.8 ppm) compared with truncate **22** ($\Sigma|\Delta\delta_H|$ = 0.76 ppm; $\Sigma|\Delta\delta_C|$ = 10.9 ppm). In both **22** and **23**, the Ardisson and Cossy groups targeted a 13,18-*anti* relationship between the two fragments; one of the two possible permutations present in hemicalide. Considering that the revised truncate represents over half of the molecular structure, the comparatively large total ^{1}H and ^{13}C deviations obtained in **23** ($\Sigma|\Delta\delta_H|$ = 0.26 ppm; $\Sigma|\Delta\delta_C|$ = 3.8 ppm) and the absence of the alternative 13,18-*syn* diastereomer for comparison led us to suspect that the reported 13,18-*anti* configuration may be a red herring.

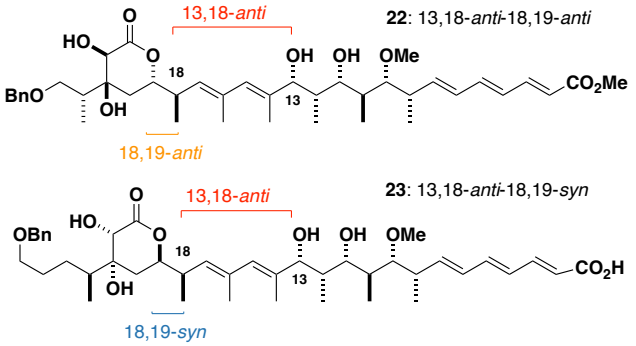


Figure 2. Ardisson and Cossy's 13,18-*anti* configured truncates.

Table 1. Sum of absolute errors |Δ| (ppm)[a][b] for **22** and **23** synthesised by Ardisson/Cossy compared to hemicalide (**1**).[6]

Entry	Sum |Δ| ^{1}H	Max |Δ| ^{1}H	Sum |Δ| ^{13}C	Max |Δ| ^{13}C
1. (13,18-*anti*-) **18,19-*syn* acid 23**[21]	**0.26**	**0.04**	**3.8**	**0.5**
2. (13,18-*anti*-) 18,19-*anti* acid **22**[20]	0.76	0.12	10.9	1.5

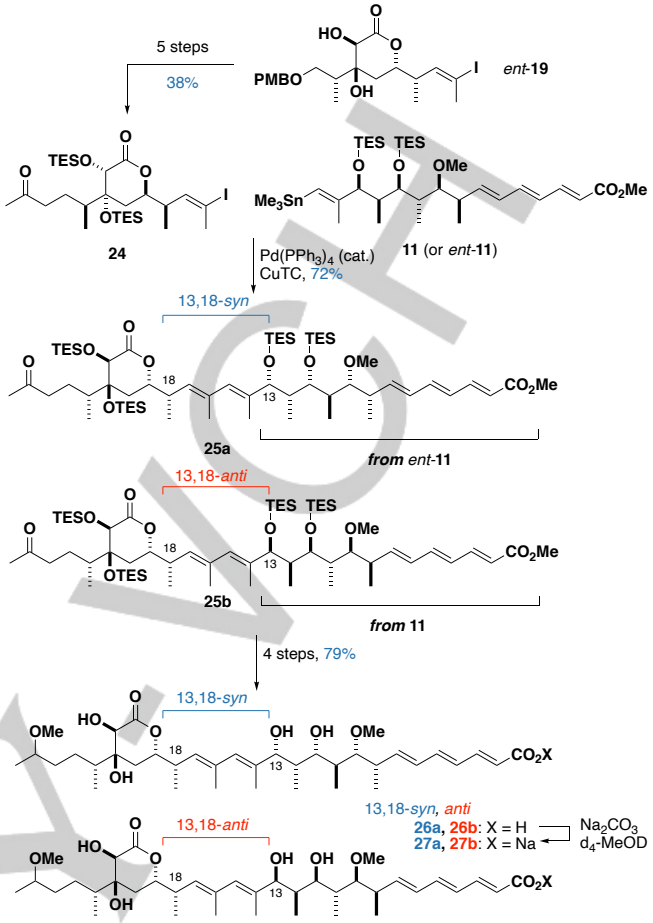
[a] Absolute errors taken for NMR shifts between C8-Me24

[b] |Δ| = δ(Experimental shift) – δ(Reported shift), errors in ppm

We initially sought to probe this problem computationally, but this was found to be impractical given the size and, in particular, the flexibility of the truncates in question. As such, we turned towards stereodefined synthesis to assist the relative configurational assignment between the C1-C13 and the C18-C24 fragment.[17] Specifically, we aimed to generate two candidate diastereomers (13,18-*syn* and 13,18-*anti*) for the C1-C28 fragment containing the C1-C15 region and the revised C16-C24 region to undertake high-resolution NMR spectroscopic comparisons

From intermediate *ent*-**19**, completion of the C16-C28 fragment **24** required five steps (Scheme 7).[55] Exploratory studies indicated that the planned Stille cross coupling was best executed using a C1-C15 stannane **11**. By coupling **24** with either **11** or *ent*-**11**, the 13,18-*syn* and 13,18-*anti* diastereomers **25a** and **25b** of the C1-C28 region were obtained. These were then transformed into model acid truncates **26a** and **26b** in four steps. Detailed ^{1}H and ^{13}C NMR comparisons between 13,18-*syn* **26a**, 13,18-*anti* **26b** with hemicalide revealed that the 13,18-*syn* configuration in **26a** ($\Sigma|\Delta\delta_H|$ = 0.05 ppm; $\Sigma|\Delta\delta_C|$ = 0.8 ppm) was a much closer match than the alternative 13,18-*anti* configuration in **26b** ($\Sigma|\Delta\delta_H|$ = 0.08 ppm; $\Sigma|\Delta\delta_C|$ = 1.7 ppm) and in acid **23** containing the 13,18-*anti* configuration[21] (Table 2, Figure 3).

During the initial investigation into the C1-C13 region of hemicalide, Ardisson *et al.* noticed significant deviations in chemical shifts in the C1-C7 trienoic acid region of their fragments from hemicalide, though notably the perturbations did not extend into the aliphatic region (C8 onwards).[14] They proposed that these deviations were attributed to hemicalide likely being isolated as the carboxylate salt rather than the free acid. To test this hypothesis, we fully deprotonated our acid fragments **26a** and **26b** to obtain their corresponding sodium salts (**27a** and **27b**) to assess if these samples would then provide closer matches to the natural product, to which we observed poorer spectroscopic



Scheme 7. Synthesis of model truncate **26a** and **26b** for NMR comparison

Table 2. Sum of absolute errors |Δ| (ppm)[a][b] for each diastereomer compared to the reported spectra of hemicalide (**1**).[6]

Entry		Sum |Δ| ^{1}H	Max |Δ| ^{1}H	Sum |Δ| ^{13}C	Max |Δ| ^{13}C
1.	**13,18-*syn* acid 26a**	**0.05**	**0.01**	**0.8**	**0.1**
2.	13,18-*syn* salt **27a**	0.08	0.02	1.7	0.2
3.	Acid **23**[21]	0.26	0.04	3.8	0.5
4.	13,18-*anti* acid **26b**	0.22	0.04	4.1	0.6
5.	13,18-*anti* salt **27b**	0.22	0.04	4.3	0.7

[a] Absolute errors taken for NMR shifts between C8-Me24

[b] |Δ| = δ(Experimental shift) – δ(Reported shift), errors in ppm

correlations. In conjunction with the observation that the chemical shifts for the C1-C7 trienoic acid region differed after HCl work up, after chromatography and after complete deprotonation, this led us to believe that hemicalide was analysed in its partially deprotonated form likely to arise from multiple rounds of chromatography during its purification. At this point, detailed spectroscopic comparison through a synthesis-enabled venture enabled the confident determination of the C1-C24 region of hemicalide down to a single diastereomer, reducing the number of possible candidate diastereomers for hemicalide down to 2^{11} (10 unknown stereocentres + unknown absolute configuration for C1-C24 region), or 1024 (2^{10}) excluding enantiomers.

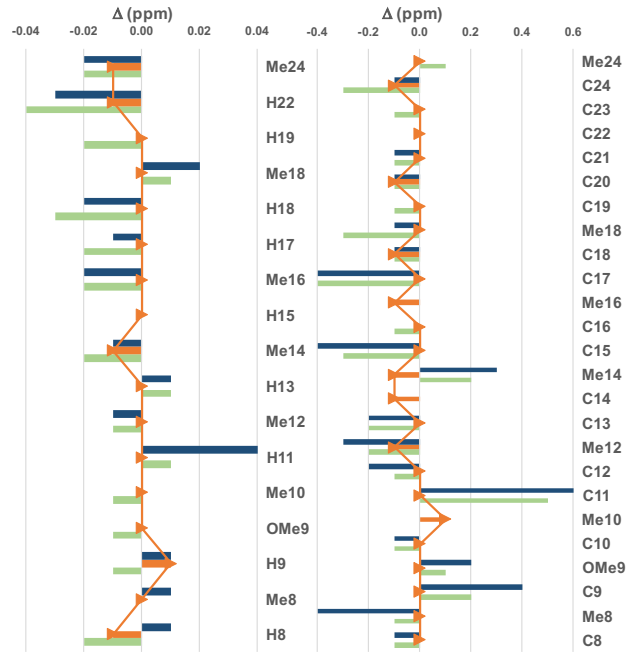
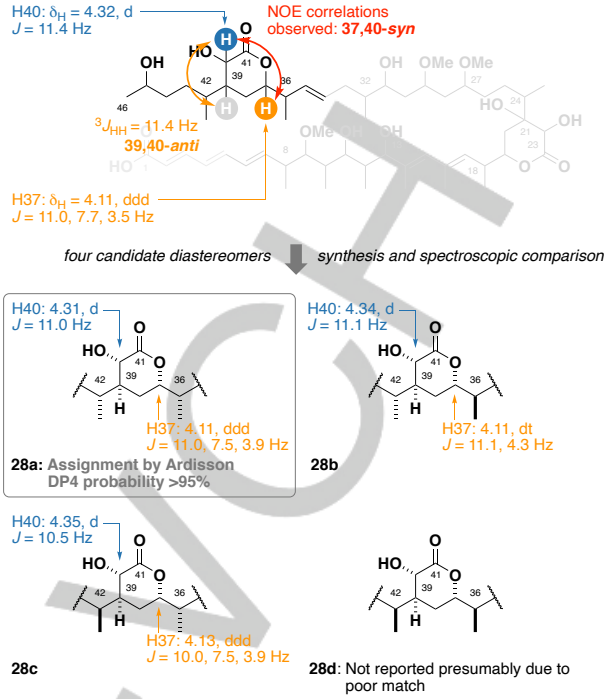


Figure 3. Bar graph highlighting the ^{1}H (left) and ^{13}C (right) NMR chemical shift differences between the reported 13,18-*anti* acid **23**[21] (**green**), model 13,18-*anti* acid **26b** (**blue**), model 13,18-*syn* acid **26a** (**orange**) and hemicalide (**1**) between H8/C8 and Me24, overlaid with a line graph for acid **26a**.[17]

5. Assignment and Synthesis of the C36-C46 Region

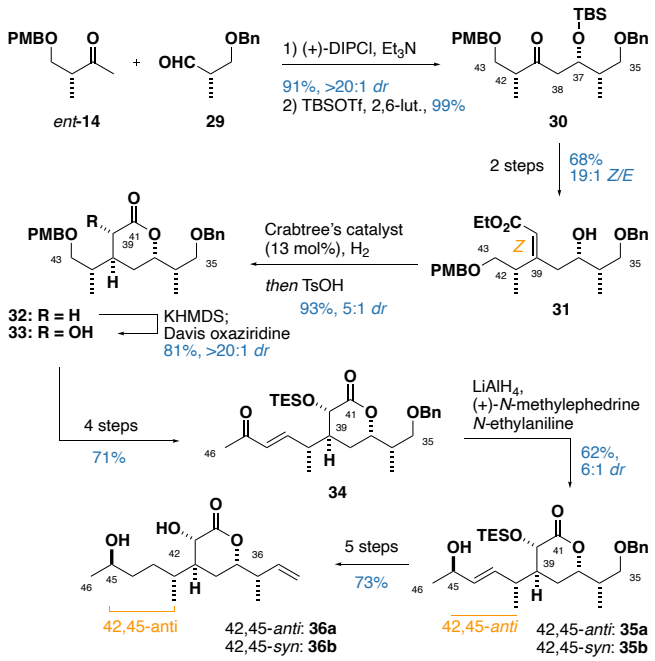
The C36-C46 hydroxylactone contains the final stereochemically rich fragment in hemicalide. Analogous to the assignment of the C18-C25 dihydroxylactone, observation of key NOE correlations and coupling constants allowed Ardisson/Cossy *et al.* to conclusively secure the configuration of the δ-lactone and reduce the number of diastereomers from 32 to eight (Scheme 8).[22] During the course of their study, they found that the distal nature of the C45 carbinol stereocentre relative to the C36-C42 hydroxylactone meant that the two candidate C45 epimers were challenging to resolve through high-field NMR spectroscopy. By arbitrarily defining the C45 stereocentre, they further reduced the number of candidate diastereomers to four, whereby the synthesis and in depth characterisation[23] of candidate diastereomers **28a-28d** and their spectroscopic comparisons with hemicalide revealed **28a** to be the best match. In parallel, DP4 calculations performed on a model C36-C46 fragment corroborated the relative configuration in **28a** proposed by the Ardisson/Cossy groups.[19,24] This provided further orthogonal evidence for the C36-C46 region, enabling the assignment of 16 of the 21 stereocentres in hemicalide, leaving 2^7 (5 unknown stereocentres + unknown absolute configurations for C1-C24 and C36-C42 regions) possible stereoisomers, or 64 (2^6) candidates excluding enantiomers.



Scheme 8. Assignment of the C36-C42 hydroxylactone region

With a confident assignment, we next embarked on a focused synthetic campaign targeting the configuration present in diastereomer **28a**.[25] Synthesis commenced with a boron-mediated aldol reaction between Roche ester-derived ketone *ent*-**14**[19] and aldehyde **29**,[26] followed by silyl protection, to give **30** with high diastereoselectivity (Scheme 9). Next, a two-step manipulation involving a Peterson olefination afforded the acyclic *Z*-enoate **31**. At this point, a hydroxyl-directed hydrogenation of the enoate olefin catalysed by Crabtree's catalyst configured the C39 stereocentre. The intermediate saturated ester was then cyclised *in situ* under acidic conditions to give the requisite δ-lactone **32**. Diastereoselective α-oxidation mediated by Davis oxaziridine installed the final lactone stereocentre in **33**.

At this point, we recalled an earlier observation regarding the C27 ketone reduction towards generating model C1-C28 fragments. Echoing the observations reported by Ardisson/Cossy *et al.*, the distal epimeric C27 mixtures obtained minimally impacted the chemical shifts for the remainder of the molecule.[22] However, subtle and minute chemical shift differences were seen, alluding to the possibility of a synthesis-enabled elucidation of the remote C45 carbinol stereocentre if an in-house processed FID file for hemicalide could be employed for direct NMR spectroscopic comparisons. With this goal in mind, a four-step sequence gave enone **34**, where stereoselective reduction under Terashima conditions[27] was found to best install the C45 stereocentre, flexibly delivering the 42,45-*anti* (**35a**) and 42,45-*syn* diastereomer (**35b**) through the choice of either antipode for the *N*-methylephedrine ligand. A five-step sequence then afforded both diastereomers **36a** and **36b** of the model diol for spectroscopic comparisons with the natural product, enabled by the timely provision of the original FID files from Dr Georges Massiot.



Scheme 9. Synthesis of model C36-C46 fragments for NMR comparison

While both epimers were indeed spectroscopically similar, they were distinguishable particularly with respect to in house processed NMR spectra for hemicalide (Table 3). Extracting both ^{1}H and ^{13}C NMR chemical shift data to two decimal places for both C45-epimeric C35-C46 fragments revealed that a 42,45-*anti* relationship (**36a**; entry 1: $\Sigma|\Delta_H|$ = 0.05 ppm; $\Sigma|\Delta_C|$ = 0.38 ppm) was a closer fit than the alternative 42,45-*syn* configuration (**36b**; entry 2: $\Sigma|\Delta_H|$ = 0.13 ppm; $\Sigma|\Delta_C|$ = 0.67 ppm).[25] With four unknown stereocentres remaining, this served to reduce the number of possible diastereomers for hemicalide down to 32.

Table 3. Determination of the 42,45-*anti* configuration in hemicalide. Sum of absolute errors |Δ| (ppm)[a][b] for each diastereomer compared to the reported spectra of hemicalide (**1**)

Entry	Sum |Δ| ^{1}H	Max |Δ| ^{1}H	Sum |Δ| ^{13}C	Max |Δ| ^{13}C
1. 42,45-*anti* 36a	**0.05**	**0.02**	**0.38**	**0.10**
2. 42,45-*syn* **36b**	0.13	0.05	0.67	0.20

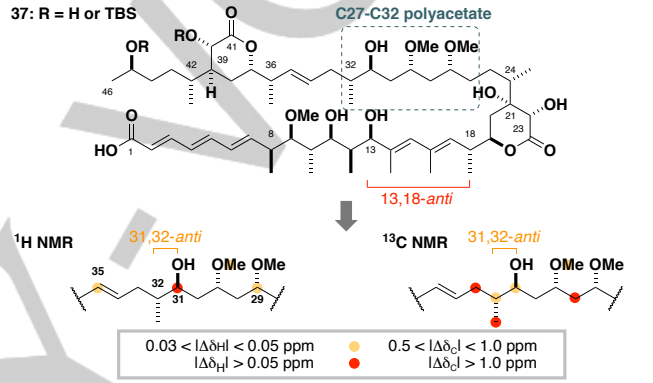
[a] Absolute errors taken for NMR shifts between H/C39-H/C46

[b] |Δ| = δ(experimental shift) – δ(reported shift), errors in ppm

6. Synthesis-Enabled Assignment of the C31-C46 Region

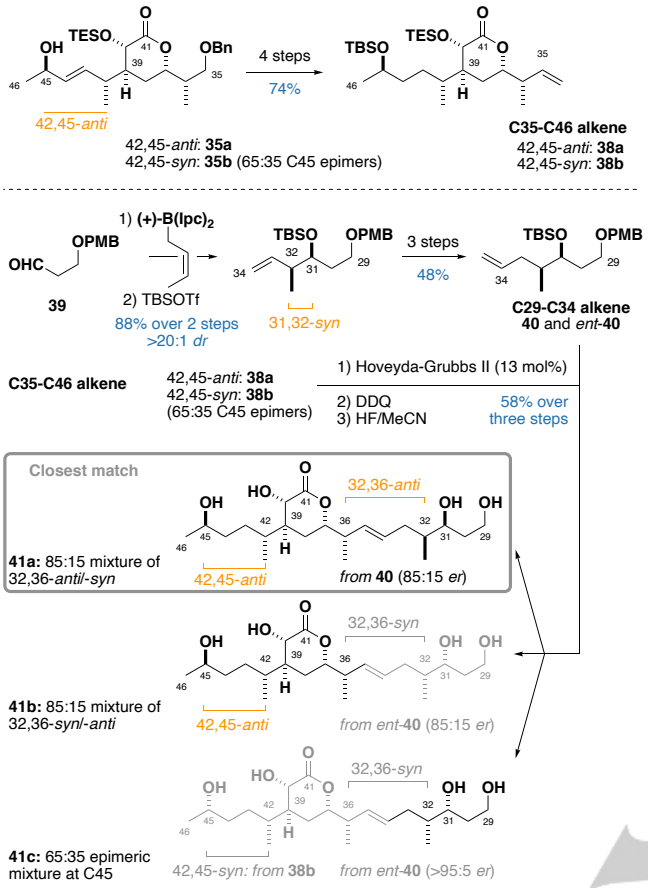
At this stage, 17 of the 21 stereocentres in hemicalide had been assigned, corroborated both by synthesis and computational studies. The four remaining unassigned stereocentres are contained within the C27-C34 polyacetate region. We attempted DP4 calculations to determine the most likely diastereomer for the C27-C34 region. However, the region's inherent flexibility led to multiple low energy conformations, which could not be further refined as the chemical shifts for the methylene protons in this region were obscured in a 22H multiplet in the ^{1}H NMR spectrum for hemicalide. These limitations meant that no candidate C25-C34 diastereomers could be assigned with high confidence.[24]

In 2019, the Ardisson and Cossy teams reported the synthesis of **37**, corresponding to the full carbon skeleton of hemicalide (Scheme 10) but with one residual silyl protecting group still in place. This synthetic venture represented the first time the polyacetate region was targeted, even though the four stereocentres appeared to be randomly assigned.[28] While large global chemical shift deviations between **37** and hemicalide provided compelling evidence that the natural product is likely to not contain this configuration, this timely report gave valuable clues towards the likely configuration within the C27-C34 region. Notably, the large ^{13}C chemical shift deviations, particularly clustering around the 31,32-*anti* related stereocentres, was considered to be a diagnostic indication that hemicalide likely possesses the opposite 31,32-*syn* configuration relative to the candidate structure pursued by the Ardisson and Cossy teams.



Scheme 10. Full carbon skeleton **37** of hemicalide as synthesised by Cossy, Meyer *et al.* Comparison of ^{1}H and ^{13}C NMR chemical shifts for C29-C35 region with the corresponding data for hemicalide does not support the 31,32-*anti* configuration.

With this hypothesis in hand, we focused our campaign to prepare both enantiomers of a suitable 31,32-*syn* fragment, with the view to validate and establish both the configuration at C31 and C32, as well as establishing its relationship to the distal 1,5-related C36-C46 region previously established. We envisaged fragment union with either antipode of the C29-C35 region could be readily achieved through olefin cross metathesis. To this end, suitable C35-C46 alkene coupling partners **38a** and **38b** were generated from intermediate **35a** and **35b**. The synthesis of the corresponding C29-C35 alkene coupling partner commenced with a Brown *syn*-crotylation of aldehyde **39** to set the 31,32-*syn* configuration (Scheme 11). From here, four steps then afforded the required fragment **40**, with the opposite enantiomer (*ent*-**40**) obtained using the enantiomeric crotylation agent.[25]



Scheme 11. Synthesis of model C31-C46 triols for NMR comparison, each containing encoded diastereomeric mixtures to aid spectroscopic differentiation

At the outset, we were aware that such distally located 1,4- and 1,5-related stereoclusters would afford only minute chemical shift differences. To secure a further form of spectroscopic validation, we planned to generate three "encoded" fragments with defined diastereomeric ratios such that each resonance could be calibrated to the alternative diastereomer.[29] In addition to this form of internal verification, this critically enabled us to resolve each resonance to two decimal points. In conjunction with an in house processed FID file for hemicalide, this strategy was viewed to enable a more confident assertion for any stereochemical conclusions to be made. To this end, we performed parallel cross-metathesis operations with a configurationally pure 42,45-*anti* C35-C46 alkene **38a** with 31,32-*syn* C29-C35 alkene (85:15 *er*), which generated the C29-C46 fragment **41a** after deprotection as an 85:15 mixture of 32,36-*anti* and -*syn* diastereomer (with 32,36-*syn* being the minor component). An analogous cross metathesis-deprotection sequence involving 31,32-*syn* *ent*-**40** (85:15 *er*) generated the alternative C29-C46 fragment **41b** as an 85:15 mixture of 32,36-*syn* and -*anti* diastereomer, with the minor component matching **41a**. Finally, to further confirm the veracity of C45, 42,45-*syn* C35-C46 alkene **38b** containing an encoded 65:35 mixture of epimers at C45 was coupled with 31,32-*syn* alkene *ent*-**40** to generate C29-C46 fragment **41c** bearing 31,32-*syn*, 32,36-*anti*, 42,45-*syn* configuration as a 65:35 mixture of epimers at C45. All three candidate diastereomers were subjected to high-resolution NMR comparison with hemicalide (**1**; Table 4, Figure 4).[25]

Detailed comparisons with hemicalide first revealed that all 31,32-*syn* diastereomers (entries 1 to 3, $\Sigma|\Delta_H| < 0.11$ ppm; $\Sigma|\Delta_C| < 2.86$ ppm) contained much smaller chemical shift deviations relative to the previously reported 31,32-*anti* configuration in **37** (entry 4: $\Sigma|\Delta_H| = 0.33$ ppm; $\Sigma|\Delta_C| = 8.48$ ppm).[9] These observations strongly supported our hypothesis in which the natural product contains a 31,32-*syn* configuration. The ^{13}C NMR data proved especially diagnostic for the remaining stereocentres; a comparison of fragments **41c** (42,45-*syn*) and **41b** (42,45-*anti*) with hemicalide further corroborated the 42,45-*anti* configuration as determined above (entry 2: $\Sigma|\Delta_C| = 2.72$ ppm) over the alternative 42,45-*syn* isomer (entry 3: $\Sigma|\Delta_C| = 2.86$ ppm). Further comparison between 32,36-*syn* **41b** (entry 2), 32,36-*anti* **41a** (entry 1) with the parent natural product conclusively gave smaller chemical shift deviations for the 32,36-*anti* diastereomer (entry 1: $\Sigma|\Delta_H| = 0.06$ ppm; $\Sigma|\Delta_C| = 1.57$ ppm) over the alternative 32,36-*syn* diastereomer (entry 2: $\Sigma|\Delta_H| = 0.11$ ppm; $\Sigma|\Delta_C| = 2.72$ ppm).[25] This synthesis-enabled venture, crucially facilitated through judicious incorporation of diastereomeric mixtures, further illuminates two of the four remaining stereocentres. In combination with the C1-C24 region and now a single diastereomer of the C31-C46 region, this leaves 2^4 (16) remaining stereoisomers, or eight possible diastereomers.

Table 4. Statistical summary of absolute errors |Δ| (ppm)[a][b] for each diastereomer compared to the reported NMR spectra of hemicalide (**1**)

Entry	Sum |Δ| ^1H	Max |Δ| ^1H	Sum |Δ| ^{13}C	Max |Δ| ^{13}C
1. 31,32-*syn*, 32,36-*anti*, 42,45-*anti* (41a)	**0.06**	**0.04**	**1.57**	**0.81**
2. 31,32-*syn*, 32,36-*syn*, 42,45-*anti* (**41b**)	0.11	0.04	2.72	0.72
3. 31,32-*syn*, 32,36-*syn*, 42,45-*syn* (**41c**)	0.10	0.03	2.86	0.73
4. Full skeleton diastereomer **37**[28] 31,32-*anti*, 32,36-*syn*, 42,45-*anti*	0.33	0.14	8.48	1.32

[a] Absolute errors taken for $^1H/^{13}C$ NMR chemical shifts between H/C31-H/C46

[b] |Δ| = δ(experimental shift) – δ(reported shift), errors in ppm

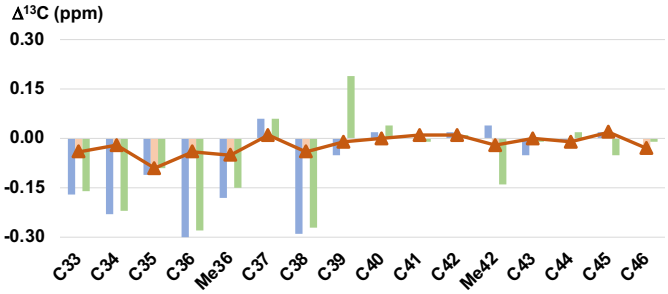


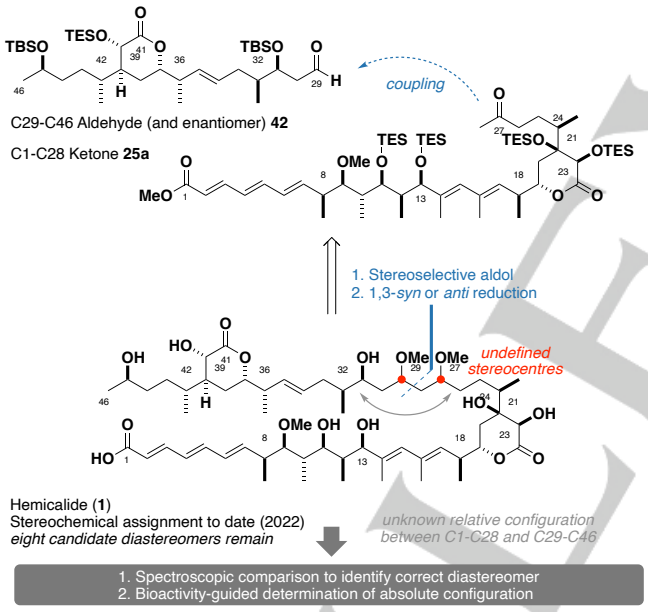
Figure 4. Bar graph highlighting the ^{13}C NMR chemical shift differences between the C33-C46 diastereomers **41a** (orange), **41b** (blue) and **41c** (green) relative to hemicalide (**1**).[25]

7. Conclusions and Outlook

With 21 unknown stereocentres and over a million permutations, the incremental yet definitive assignment of 19 out of 21 stereocentres, corresponding to eight remaining diastereomers, helps clear the impenetratable fog that has enveloped this enigmatic natural product. Crucially, the path paved towards

visualising hemicalide's tantalising molecular architecture was achieved through the synergism of contemporary forensic tools in the form of *ab initio* DP4 calculations, as well as the judicious execution of stereocontrolled synthesis.[30,31] Notably, the omission of either technique would have significantly confounded the progress towards solving this seemingly intractable problem.

With only eight diastereomers remaining, a focused total synthesis campaign can now commence. Noting that the remaining stereocentres (C27 and C29) are 1,3-related, a proposed endgame strategy entails a diastereoselective aldol coupling between a C1-C28 ketone **25a** and a C29-C46 aldehyde **42** to first configure C29 (Scheme 12). A subsequent diastereoselective reduction would set the final C27 stereocentre. From here, *O*-methylation, global desilylation followed by ester hydrolysis should enable controlled access to all possible diastereomeric permutations for high-resolution NMR spectroscopic comparisons. Importantly, these proposed advanced intermediates are readily attainable from our current synthetic routes. With the goal towards potential preclinical development of this intriguing natural product, any particularly close NMR hits with the natural product data can also be tested in cancer cell lines. This "activity-guided" approach is crucial, particularly in the absence of chiroptical data, where the absolute configuration will need to be inferred from biological data.



Scheme 12. Proposed endgame for hemicalide to access all eight remaining diastereomers, and proposed workflow towards securing the absolute configuration of hemicalide

In conjunction with Marcourt and Massiot's companion article[7] detailing its isolation and spectroscopic analysis, this Review provides a case report on the detective narrative of hemicalide. While its structural enigma still requires a final chapter to be fully solved, these collective discoveries exemplify the power of stereocontrolled synthesis[5,32] in combination with computational and experimental NMR analysis in assigning the 3D molecular architecture to such a hot target molecule.[33] Of note, judiciously executed fragment synthesis remains especially pertinent and exceedingly powerful in this context, where the absence of genetic data of the producing organism,[34] or structurally analogous natural products fully preclude the application of contemporary analytical techniques (*e.g.* gene cluster analysis[34] and structure homology-assisted assignments[35,36]) to aid hemicalide's configurational elucidation.[5, 33] These preliminary results give confidence that a full structural assignment of a compound as complex as hemicalide, from exceedingly sparse initial data, is indeed possible by adopting this modus operandi. As demonstrated by the ever improving NMR correlations presented herein, we are optimistic that a fruitful pursuit towards unravelling the final unknown issues in the hemicalide enigma will avoid its relegation to a list of "molecules that were never there".[37]

Acknowledgements

The work at Cambridge summarised in this Review was conducted by a group of talented PhD students: Steven Smith, Callum MacGregor, Bing Yuan Han, Nelson Lam and Tegan Stockdale, as well as MSci students, James Thompson and Ciaran Lunt. Their dedication, creativity and hard work are gratefully acknowledged. We thank the Woolf Fisher Trust and Trinity Hall for funding (fellowship to NYSL), Prof. Jonathan Goodman for valuable support throughout, Dr Georges Massiot for the provision of raw NMR spectroscopic files for hemicalide and, together with Prof. Janine Cossy, for helpful discussions.

Keywords: Configuration determination • Natural products • NMR spectroscopy • Structure elucidation • Total Synthesis

References

[1] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2020**, *83*, 770–803.

[2] A. R. Carroll, B. R. Copp, R. A. Davis, R. A. Keyzers, M. R. Prinsep, *Nat. Prod. Rep.* **2022**, DOI: 10.1039/D1NP00076D.

[3] M. P. Puglisi, J. M. Sneed, R. Ritson-Williams, R. Young, *Nat. Prod. Rep.* **2019**, *36*, 410–429.

[4] D. H. Williams, M. J. Stone, P. R. Hauck, S. K. Rahman, *J. Nat. Prod.* **1989**, *52*, 1189–1208.

[5] For a comprehensive discussion on the role of synthesis in the stereochemical assignment of complex natural products, see: (a) N. Y. S. Lam, I. Paterson, *Eur. J. Org. Chem.* **2020**, *2020*, 2310–2320. (b) N. Y. S. Lam, PhD Thesis, *The Synthesis-Enabled Stereochemical Elucidation of the Marine Natural Products Hemicalide and Phormidolide A,* University of Cambridge, **2019**, DOI: 10.17863/CAM.45911.

[6] I. Carletti, C. Debitus, G. Massiot, *Molécules Polykétides Comme Agents Anticancéreux.* WO2011051380 (A1), **2011**.

[7] L. Marcourt, G. Massiot, *Eur. J. Org. Chem.* **2022**, *DOI: 10.1002/ejoc.202200208.*

[8] K. N. Bhalla, *Oncogene* **2003**, *22*, 9075–9086.

[9] C. Holohan, S. Van Schaeybroeck, D. B. Longley, P. G. Johnston, *Nat. Rev. Cancer* **2013**, *13*, 714–726.

[10] I. Paterson, S. Williams, *Isr. J. Chem.* **2017**, *57*, 192–201.

[11] I. Paterson, S. M. Dalby, J. C. Roberts, G. J. Naylor, E. A.

Guzmán, R. Isbrucker, T. P. Pitts, P. Linley, D. Divlianska, J. K. Reed, A. E. Wright, *Angew. Chem. Int. Ed.* **2011**, *50*, 3219–3223.

[12] I. Paterson, K. K. H. Ng, S. Williams, D. C. Millican, S. M. Dalby, *Angew. Chem. Int. Ed.* **2014**, *53*, 2692–2695.

[13] I. Paterson, R. Britton, O. Delgado, N. M. Gardner, A. Meyer, G. J. Naylor, K. G. Poullennec, *Tetrahedron* **2010**, *66*, 6534–6545.

[14] E. Fleury, M. I. Lannou, O. Bistri, F. Sautel, G. Massiot, A. Pancrazi, J. Ardisson, *J. Org. Chem.* **2009**, *74*, 7034–7045.

[15] R. W. Hoffmann, U. Weidmann, *Chem. Ber.* **1985**, *118*, 3980–3992.

[16] S. G. Smith, J. M. Goodman, *J. Am. Chem. Soc.* **2010**, *132*, 12946–12959.

[17] B. Y. Han, N. Y. S. Lam, C. I. MacGregor, J. M. Goodman, I. Paterson, *Chem. Commun.* **2018**, *54*, 3247–3250.

[18] E. Fleury, G. Sorin, E. Prost, A. Pancrazi, F. Sautel, G. Massiot, M. I. Lannou, J. Ardisson, *J. Org. Chem.* **2013**, *78*, 855–864.

[19] C. I. MacGregor, B. Y. Han, J. M. Goodman, I. Paterson, *Chem. Commun.* **2016**, *52*, 4632–4635.

[20] C. Lecourt, S. Boinapally, S. Dhambri, G. Boissonnat, C. Meyer, J. Cossy, F. Sautel, G. Massiot, J. Ardisson, G. Sorin, M. I. Lannou, *J. Org. Chem.* **2016**, *81*, 12275–12290.

[21] G. Sorin, E. Fleury, C. Tran, E. Prost, N. Molinier, F. Sautel, G. Massiot, S. Specklin, C. Meyer, J. Cossy, M. I. Lannou, J. Ardisson, *Org. Lett.* **2013**, *15*, 4734–4737.

[22] S. Specklin, G. Boissonnat, C. Lecourt, G. Sorin, M. I. Lannou, J. Ardisson, F. Sautel, G. Massiot, C. Meyer, J. Cossy, *Org. Lett.* **2015**, *17*, 2446–2449.

[23] E. De Gussem, W. Herrebout, S. Specklin, C. Meyer, J. Cossy, P. Bultinck, *Chem. Eur. J.* **2014**, *20*, 17385–17394.

[24] C. I. MacGregor, PhD Thesis, *Studies Towards the Structural Elucidation and Total Synthesis of Hemicalide*, University of Cambridge, **2015**.

[25] T. P. Stockdale, N. Y. S. Lam, I. Paterson, *Chem. Commun.* **2022**, *58,* 5729–5732.

[26] I. Paterson, K. S. Yeung, *Tetrahedron Lett.* **1993**, *34*, 5347–5350.

[27] S. Terashima, N. Tanno, K. Koga, *J. Chem. Soc. Chem. Commun.* **1980**, 1026–1027.

[28] C. Lecourt, S. Dhambri, K. Yamani, G. Boissonnat, S. Specklin, E. Fleury, K. Hammad, E. Auclair, S. Sablé, A. Grondin, P. B. Arimondo, F. Sautel, G. Massiot, C. Meyer, J. Cossy,G. Sorin, M. I. Lannou, J. Ardisson, *Chem. Eur. J.* **2019**, *25*, 2745–2749.

[29] J. Wu, P. Lorenzo, S. Zhong, M. Ali, C. P. Butts, E. L. Myers, V. K. Aggarwal, *Nature* **2017**, *547*, 436–440.

[30] N. Y. S. Lam, T. P. Stockdale, M. J. Anketell, I. Paterson, *Chem. Commun.* **2021**, *57*, 3171–3189.

[31] T. P. Stockdale, N. Y. S. Lam, M. J. Anketell, I. Paterson, *Bull. Chem. Soc. Jpn.* **2021**, *94*, 713–731.

[32] N. Y. S. Lam, G. Muir, V. R. Challa, R. Britton, I. Paterson, *Chem. Commun.* **2019**, *55*, 9717–9720.

[33] I. E. Ndukwe, X. Wang, N. Y. S. Lam, K. Ermanis, K. L. Alexander, M. J. Bertin, G. E. Martin, G. Muir, I. Paterson, R. Britton, J. M. Goodman, E. J. N. Helfrich, J. Piel, W. H. Gerwick, R. T. Williamson, *Chem. Commun.* **2020**, *56*, 7565–7568. This study is notable given that, even with independent comprehensive genetic analyses undertaken, a stereochemical reassignment was rendered necessary through the combination of synthesis and computational methods.

[34] Personal communication with Dr. Georges Massiot indicated that initial attempts at recultivating the parent organism were unsuccessful, and therefore data on hemicalide's biosynthesis remain as yet elusive.

[35] E. J. N. Helfrich, R. Ueoka, A. Dolev, M. Rust, R. A. Meoded, A. Bhushan, G. Califano, R. Costa, M. Gugger, C. Steinbeck, P. Moreno and J. Piel, *Nat. Chem. Biol*., **2019**, *15*, 813–821.

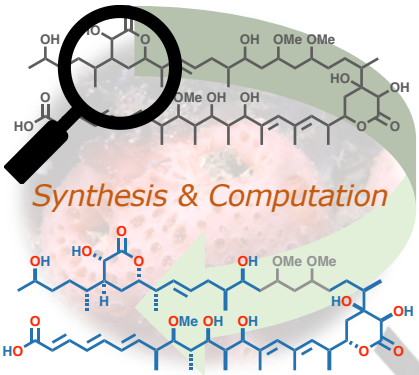
[36] For an example of a proposed biogenetic hypothesis enabled by structural homology analysis, see: I. Paterson, R. Britton, O. Delgado, A. E. Wright, Chem. Commun. **2004**, 632–633.

[37] K. C. Nicolaou, S. A. Snyder, *Angew. Chem. Int. Ed.* **2005**, *44*, 1012–1044.

WILEY-VCH

Entry for the Table of Contents

This Review highlights how the judicious and synergistic application of computational NMR analysis and stereocontrolled synthesis has enabled the incremental assignment of hemicalide—a complex and extraordinarily bioactive marine polyketide containing 21 unknown stereocentres.



Total synthesis, Polyketides*

Nelson Y. S. Lam, Ian Paterson**

Page No. – Page No.

Deep-sea discovery and detective work: towards solving the hemicalide structural enigma through computational NMR analysis and stereocontrolled synthesis.

Institute and/or researcher Twitter usernames: @nelsonyslam @chemcambridge

WILEY-VCH