Accepted Manuscript

Assessment of the regioselectivity in the condensation reaction of unsymmetrical *o*-phthaldialdehydes with alanine

Agathe C.A. D'Hollander, Nicholas J. Westwood

PII: S0040-4020(17)31183-3

DOI: 10.1016/j.tet.2017.11.035

Reference: TET 29115

To appear in: *Tetrahedron*

Received Date: 27 July 2017

Revised Date: 5 November 2017

Accepted Date: 13 November 2017

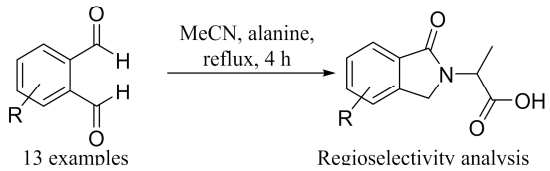
Please cite this article as: D'Hollander ACA, Westwood NJ, Assessment of the regioselectivity in the condensation reaction of unsymmetrical *o*-phthaldialdehydes with alanine, *Tetrahedron* (2017), doi: 10.1016/j.tet.2017.11.035.

Graphical Abstract

Assessment of Regioselectivity in the Condensation Reaction of Unsymmetrical *o*-Phthaldialdehydes with Alanine

Leave this area blank for abstract info.

Agathe C.A. D'Hollander and Nicholas J. Westwood*
School of Chemistry, University of St Andrews and EaStCHEM, North Haugh, St Andrews, Fife, KY16 9ST, UK.



Design of a highly regioselective substrate

Increased mechanistic understanding

Assessment of the Regioselectivity in the Condensation Reaction of Unsymmetrical *o*-Phthaldialdehydes with Alanine

Agathe C.A. D'Hollander[a] and Nicholas J. Westwood[a,*]

[a]*School of Chemistry and Biomedical Sciences Research Complex, University of St Andrews and EaStCHEM, North Haugh, St Andrews, Fife, KY16 9ST, UK.*

Keywords:
o-phthaldialdehyde
Condensation reaction
Regioselectivity
Mechanistic understanding

Abstract:

One approach for the synthesis of isoindolinones, a privileged bioactive heterocyclic core structure, involves a condensation reaction of *o*-phthaldialdehydes with a suitable nitrogen-containing nucleophile. This fascinating reaction is revisited here in the context of the use of *o*-phthaldialdehydes that contain additional substituents in the aromatic ring leading to a detailed analysis of the regioselectivity of the reaction. Eleven monosubstituted *o*-phthaldialdehydes were synthesised and reacted with alanine. The regioselectivity observed across the eleven substrates led to the design of a disubstituted substrate that reacted with very high control. A gram-scale reaction followed by esterification gave one major regioisomer in high yield. In addition, the regioselectivity observed on reaction of two novel monodeuterated substrates led to an increased mechanistic understanding.

∗ Corresponding author. Tel.: +44-1334-46-3816; fax: +44-1334-46-2595; e-mail: njw3@st-andrews.ac.uk

1. Introduction

The isoindolinones make up an important class of bioactive molecules that includes the known drugs Pazinaclone (**1**),[1a] Indoprofen (**2**)[1b] and Chlorthalidone (**3**)[1c] (Figure 1a).

<<single column figure>>

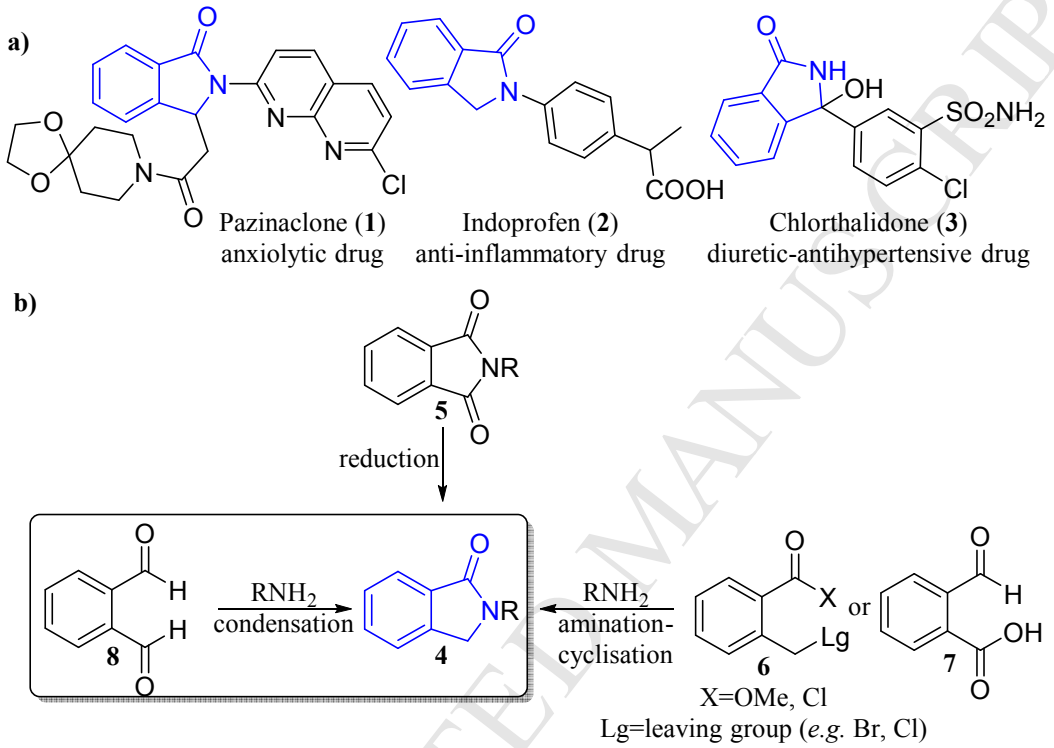


Figure 1: a) Structure of known bioactive isoindolinone containing drugs;[1] **b)** Overview of some of the different routes used to prepare the isoindolinone core **4.**[2-6a]

Common methods of obtaining isoindolinones that are unsubstituted in the aromatic ring, for example compound **4** (Figure 1b), include selective reduction of **5**,[2] reductive amination-cyclisation of **6**[3] or **7**[4] with a primary amine (RNH_2) and, of interest here, the condensation reaction of *o*-phthaldialdehyde (**8**) with a primary amine (RNH_2).[5,6a]

To date, the majority of studies performed on this condensation reaction have focused on evaluating the scope of the amine nucleophile that can be tolerated in the reaction[5a,5b,5d,6a,6b] and/or proposing potential reaction mechanisms.[5a,5b,6] In contrast, examples of the use of this

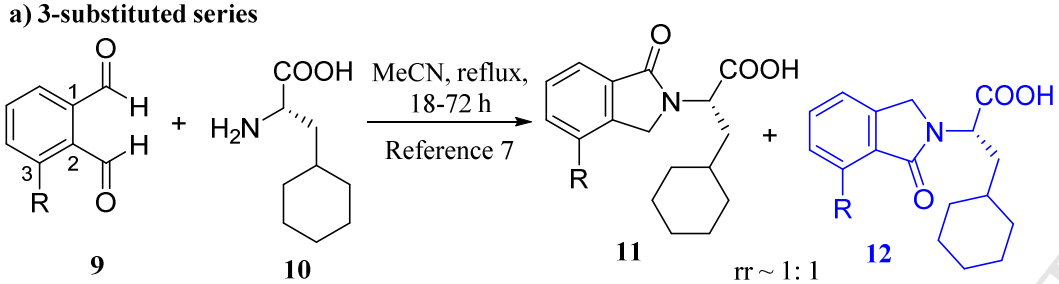
condensation reaction with monosubstituted *o*-phthaldialdehydes are rare (SI1 part I). One report describes a regioselectivity of 1:1 for the products **11**:**12** resulting from the condensation of **9** with **10** (Scheme 1a).[7] However, the observed regioselectivity was measured only after filtration or purification by column chromatography. Isolated yields for the formation of a single isomer, **14** in most cases, resulting from the condensation of **13** with various amines have also been reported (Scheme 1b).[8] Other studies have provided only isolated yield(s) after purification (for one or for each isomer), incomplete regioisomeric ratio (rr) data within a series or have claimed to form a single regioisomer (no yield provided) without discussing the other possible isomer (SI1 part I).[7-9]

The work reported here revisits this issue by presenting a detailed study of the regiochemical outcome of the condensation of alanine (**16**) with 3-monosubstituted *o*-phthaldialdehydes **17** (to give **18** and **19**, Scheme 1c) and with 4-monosubstituted *o*-phthaldialdehydes **20** (to give **21** and **22**, Scheme 1d). Based on the initial results, the design of a highly regioselective substrate was achieved consistent with an improved understanding of the reaction. Further mechanistic insights were provided by the use of novel mono-deuterated substrates.

<<Single column scheme>>

Previous Work

a) 3-substituted series



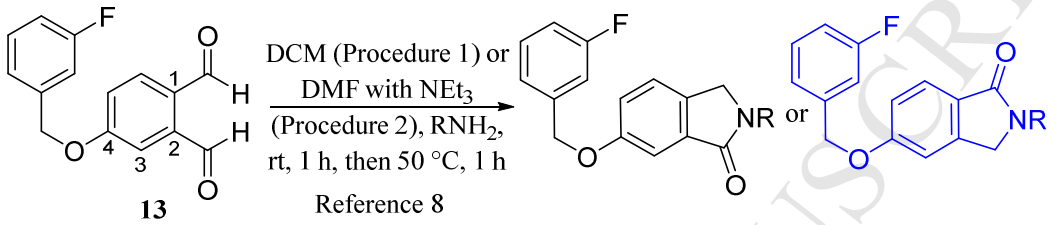
9 R=F, Cl, SMe

10

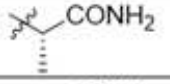
11 + **12**

rr ~ 1: 1
R=F, 77% yield
R=SMe, quant. yield
R=Cl, no yield reported

b) 4-substituted series

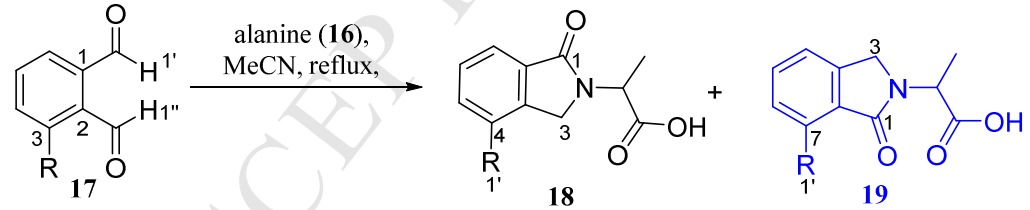


13

R	Procedure	**14**	**15**
$(CH_2)_2OMe$	(1)	19%	-
	(1)	25%	-
	(2)	24%	-
	(1)	20%	-
	(2)	20%	-
CH_2COOMe	(1)	-	22%
	(2)	17%	-
$CH_2C{\equiv}N$	(1)	13%	-

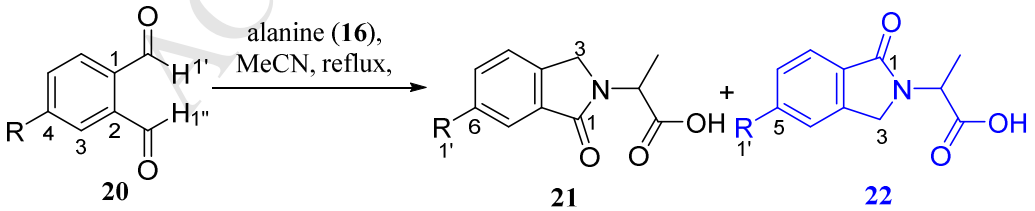
This Work

c) 3-substituted series



17 **18** **19**

d) 4-substituted series



20 **21** **22**

Scheme 1: a) and **b)** eported in the literature.[7-8] Apparently no regioselectivity was observed for the condensation of **9** with **10** (ratio of **11**:**12** of 1:1 obtained after filtration (for R=F, Cl)

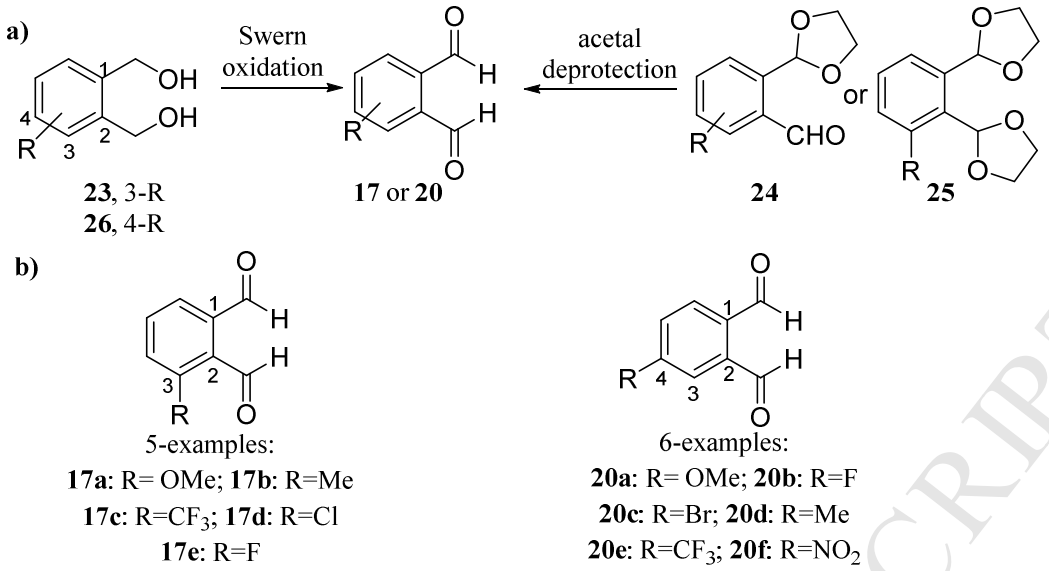
or after purification of the crude reaction mixture by column chromatography (for R=SMe)).[7] Condensation of **13** with various amines (RNH_2) followed by purification by column chromatography mainly led to isomer **14** (no comment was made regarding the apparent change of regioselectivity to give isomer **15** as the major isomer in one case).[8] **c)** and **d)** A summary of the study reported in this work to explore in more detail the influence of a substituent at the 3- or 4-position of the starting *o*-phthaldialdehyde on the observed regioselectivity.

2. Results and Discussions

2.1. Synthesis of Monosubstituted o-Phthaldialdehydes

Five 3-substituted *o*-phthaldialdehydes **17a-e** were synthesised using 2-5 step routes involving either a Swern oxidation of the corresponding diol **23** or an acetal deprotection of the corresponding monoacetal **24** or diacetal **25** (Scheme 2 and SI1 part II.1 for more details). It should be noted that the synthesis of pure samples of **17a-e** was particularly challenging (in line with literature reports[10]) with significant decomposition occurring during purification attempts and on storage. In several cases freshly prepared crude samples of the dialdehydes were used (Table 1, footnote c). Additionally, six 4-monosubstituted *o*-phthaldialdehydes **20a-f** were synthesised using 2-3 step routes either involving a Swern oxidation of the corresponding diol **26** or an acetal deprotection of the corresponding monoacetal **24** (Scheme 2 and SI1 part II.2 for more details).

<<single column scheme>>



Scheme 2: a) General synthesis scheme describing the approaches used to prepare mono-substituted *o*-phthaldialdehydes **17** or **20**. **17** and **20** were synthesised by Swern oxidation of **23** or **26**, or by acetal deprotection of **24** or **25**; b) Structures of the monosubstituted *o*-phthaldialdehyde substrates **17** and **20** synthesised in this study.

2.2 Regioselectivity of the Condensation Reaction of Mono-substituted o-Phthaldialdehydes

The mono-substituted *o*-phthaldialdehydes **17a-e** and **20a-f** were refluxed for 4 h with alanine (**16**, 1.2 equivalents) in anhydrous acetonitrile before the reaction was concentrated *in vacuo*. The crude reaction mixtures (except when specified, Table 1) were then analysed using a quantitative ^{1}H NMR experiment. A baseline correction was applied using MestReNova-9 software and integrations were calculated relative to one proton on deconvoluted peaks (see Figure 2 for an example of the analysis applied to the formation of **18a**/**19a** and SI1 part III.1 for the rest of the NMR analysis; also see the experimental section below for a detailed explanation of the analytical protocol used).

In two of the condensation reactions the structure of the major regioisomer was identified by comparison with the ^{1}H NMR spectrum of a pure sample of one of the regioisomers (for **18a**/**19a**, **21a**/**22a**, for the synthesis of authentic isomers see SI1 part III.2). In the rest of the

cases, advanced NMR techniques (HSQC, HMBC, COSY) applied to the crude reaction mixture were used to assign the structure of the major regioisomer. Considering the analysis of the regioisomeric mixture of **18b/19b** as an example (Figure 3), the proximity of a carbonyl was observed to shift the signal corresponding to the aromatic H7 proton in **18b** and the methyl H1' protons in **19b** downfield (Figures 3a and 3b). Identification of H7 in **18b** was further validated by its correlation with C1 in the HMBC analysis of the regioisomeric mixture (Figure 3a). In contrast, H4 in **19b** showed a correlation with C3 in this HMBC analysis (Figure 3c). Using the correlations observed in the COSY spectrum (Figure 3d), the signals corresponding to H5 and H6 for **18b** and **19b** were finally assigned. The value of the integrals in the 1D quantitative ^{1}H NMR spectra enabled the identification of **18b** as major isomer.

Having assigned the signals corresponding to the two regioisomeric products in each case, and using as many peaks as possible, an average percentage of the major isomer with its 95% confidence interval was then calculated for each reaction (SI1 part III.3). The reaction and its analysis were also carried out in duplicate for each substrate demonstrating high reproducibility (SI1 part III.3).

<< double column width >>

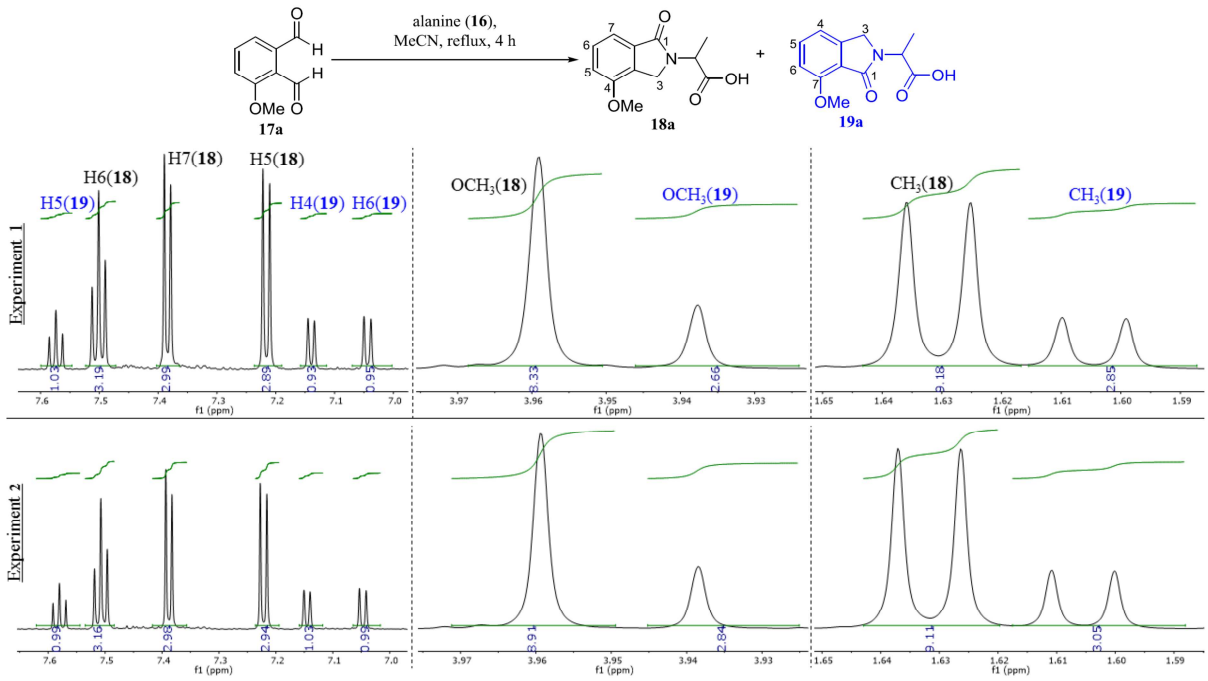


Figure 2: Quantitative ^{1}H NMR spectra of the product mixture containing **18a** and **19a** formed on reaction of **17a** with alanine (**16**) (Table 1 entry 1). From left to right the spectra represent the aromatic, the methoxy and the alanine methyl protons. Overlap between the other signals corresponding to **18a** and **19a** prevented the use of the other proton signals in the calculations. The integrals and assignments considered for the calculation of the average percentage of the major regioisomer with its 95% confidence interval (and therefore the regioisomeric ratio, see Table 1) are shown here. The top spectra represent the first experiment and the spectra at the bottom represent the repeat reaction.

<< double column width >>

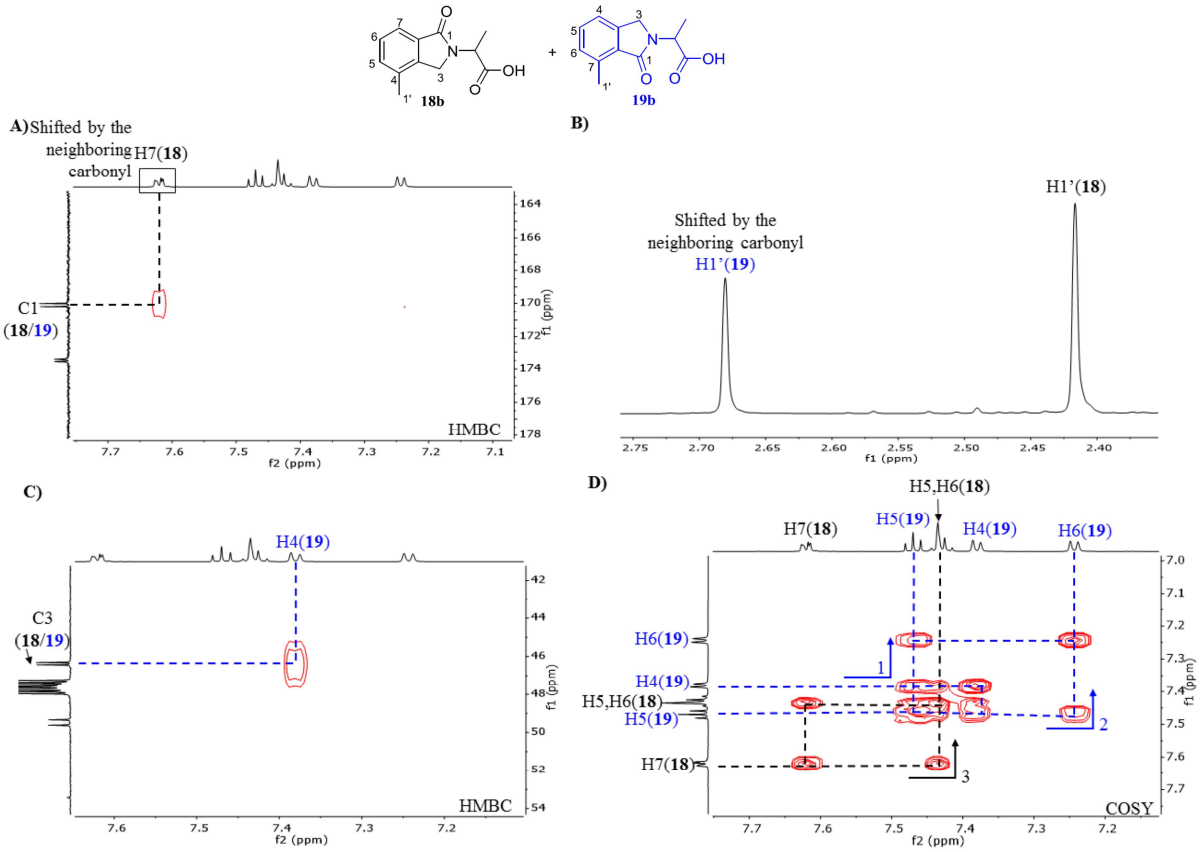
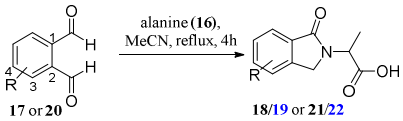
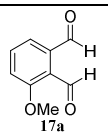
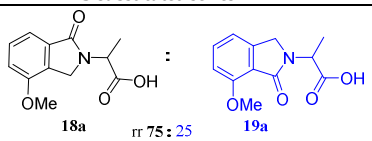
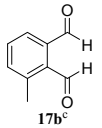
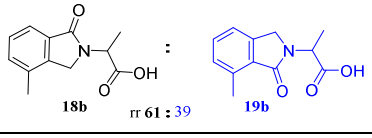
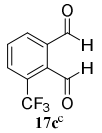
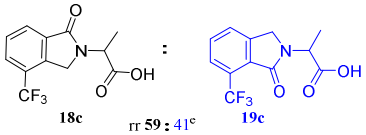
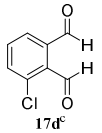
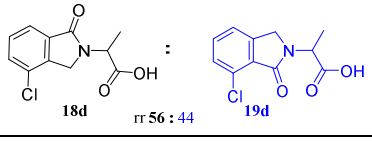
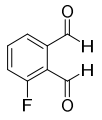
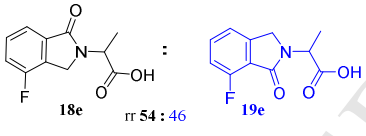
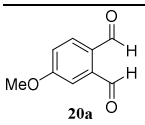
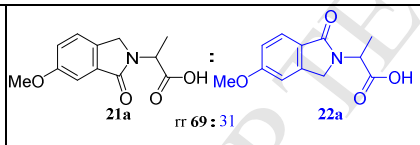
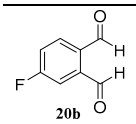
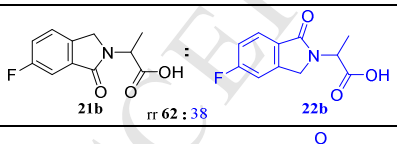
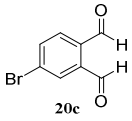
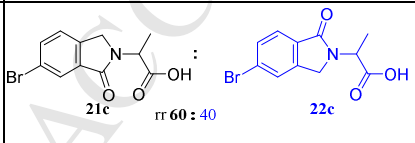
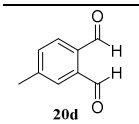
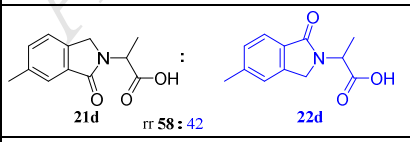
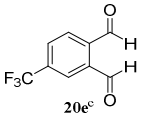
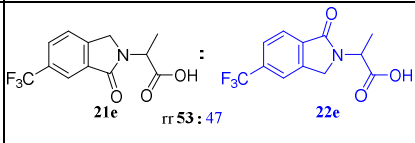


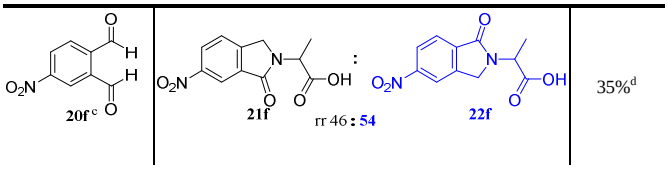
Figure 3: NMR data used to identify the structure of the major isomer in the mixture of **18b** and **19b** regioisomers. H7 in **18b** was identified by its downfield shift, due to the neighbouring carbonyl group, and its correlation to C1 in the HMBC spectrum mixture (spectrum A). H1' from **19b** was identified by its downfield shift in the ^{1}H NMR spectrum due to the neighbouring carbonyl group (spectrum B). H4 from **19b** was identified by its correlation to C3 in the HMBC spectrum of the regioisomeric mixture (spectrum C). Starting with H4 from **19b**, the correlations observed in the COSY spectrum (spectrum D) identified H5 (arrow 1) then H6 (arrow 2). Similarly, the correlations observed in the COSY spectrum for H7 from **18b** led to the identification H5 and H6 (arrow 3).

<<single column>>

Table 1: Condensation reaction of monosubstituted *o*-phthaldialdehydes **17** or **20**.



Substrate	Products ; regioisomeric ratio (rr)[a]	Yield[b] (%)
3-substituted series		
 **17a**	 **18a** : **19a** rr **75** : 25	78%
 **17b**[c]	 **18b** : **19b** rr **61** : 39	49%[d]
 **17c**[c]	 **18c** : **19c** rr **59** : 41[e]	11%[f]
 **17d**[c]	 **18d** : **19d** rr **56** : 44	33%[d]
 **17e**	 **18e** : **19e** rr **54** : 46	60%
4-substituted series		
 **20a**	 **21a** : **22a** rr **69** : 31	90%
 **20b**	 **21b** : **22b** rr **62** : 38	93%
 **20c**	 **21c** : **22c** rr **60** : 40	90%
 **20d**	 **21d** : **22d** rr **58** : 42	73%
 **20e**[c]	 **21e** : **22e** rr **53** : 47	55%[d]



[a] based on analysis of crude reaction before purification (SI1 part III.3). [b] isolated yield of the mixture of the two inseparable regioisomers. [c] crude dialdehyde was used. [d] extrapolated yield over two steps. [e] a significant amount of impurities in the crude reaction mixture led to an inaccurate determination of rr (57:43 – 60:40, SI1 part III.1). The rr is reported after purification by chromatography with collection of as much product as possible. [f] the batch of dialdehyde **17c** was divided into two portions to carry out the duplicate reactions. The resulting products were combined to give a yield over two steps.

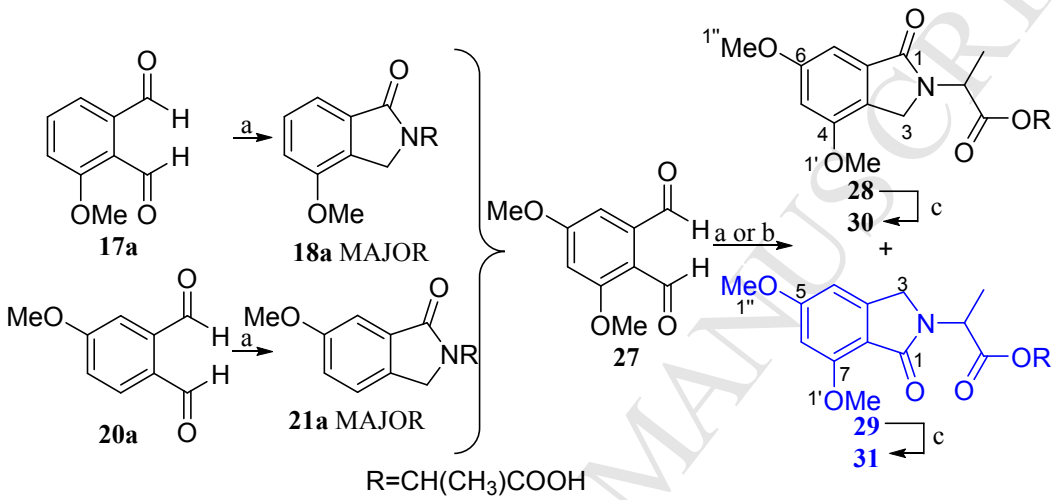
In both the 3- and 4-substituted series, the regioselectivity of the condensation reaction was dependent on the substituent employed, as expected. In the 3-substituted series, the presence of either an electron-donating or an electron-withdrawing substituent favoured the formation of regioisomer **18** although a decrease in the electron-donating strength led to a clear decrease in the regioselectivity. In the 4-substituted series it was observed that as the electron-donating properties of the substituent decreased (OMe > F/Br/Me > CF_3 > NO_2), the regioselectivity decreased and then switched from **21** being the major isomer to its regioisomer **22** being the dominant product. In both series, the methoxy substituent gave the highest regioselectivity leading to the proposal that a dimethoxysubstituted *o*-phthaldialdehyde **27** (Scheme 3) should react with very high regioselectivity.

2.3 Dimethoxysubstituted o-Phthaldialdehyde

As the monomethoxysubstituted substrates **17a** and **20a** led to the major regioisomers **18a** and **21a** respectively, it was proposed that a 3,5-dimethoxysubstituted dialdehyde **27** would

react to give **28** with an increased regioselectivity compared to **17a** and **20a** (Scheme 3). A 6-step route was developed to obtain **27** (SI1 part IV.1) and pleasingly, its subsequent reaction with alanine (**16**) gave **28** as the major regioisomer in a **88**:**12** ratio of **28**:**29** (Scheme 3 and SI1 part IV.2).

<<single column>>



Scheme 3: Rationale behind the design of substrate **27**, a 3,5-dimethoxysubstituted o-phthaldialdehyde expected to favour the regioisomer **28** with a high regioselectivity. Reagents and conditions: a) alanine (**16**), anhyd. MeCN, reflux, 4 h, 93% yield of **28**:**29** with rr 88:12 (from 0.03 g of **27**), 90% yield of **28** and **29** (1.2 g, from 0.97 g of **27**); (b) alanine (**16**), anhyd. MeCN, rt, 29 h, quant. yield of **28**:**29** with rr 97:3 (from 0.03 g of **27**); (c) $SOCl_2$, anhyd. MeOH, rt, 12 h, 84% (1.06 g) for **30**, 16% (0.20 g) for **31**.

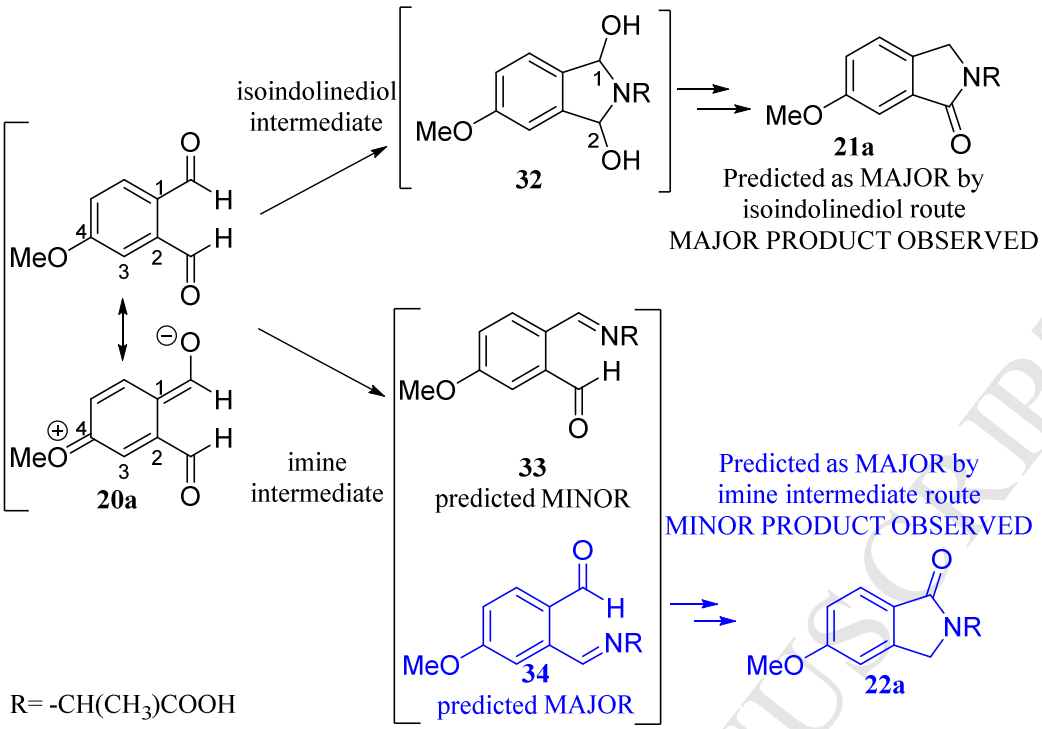
On scaling up the condensation reaction of **27** (0.97 g instead of 30 mg initially used), the quantitative ^{1}H NMR analysis of the crude reaction mixture was not carried out in this case due to the low solubility of the isoindolinone products requiring a large amount of deuterated solvent for their complete dissolution (Scheme 3). Purification by column chromatography

led to the pure and inseparable regioisomers **28** and **29** in good yield (90%, 1.2 g). Derivatisation using thionyl chloride and methanol gave the now separable regioisomeric esters isolated in 1.06 g for ester **30** (84% yield) and 0.2 gram for ester **31** (16% yield) respectively. The overall esterification yield was quantitative and thus the initial ratio of **28** and **29** for the larger scale was calculated as 84:16 which was comparable to the small scale result (88:12 of **28**:**29**). Inspired by this result, the condensation reaction was attempted on a small scale (30 mg of **27**) at room temperature (as opposed to reflux) and led to an excellent regioselectivity of **97**:**3** of **28**:**29** (Scheme 3 and SI1 part IV.3).

2.4 Mechanism Discussion

A continuing discussion on the likely mechanism of this condensation reaction has occurred in the literature over the last decade.[5-6] The early stages of the proposed mechanisms can be divided into two categories involving either an isoindolinediol intermediate[5] (such as **32** from **20a**) or imine intermediates[6] (such as **33** and **34** from **20a**) (Scheme 4).

<<single column>>



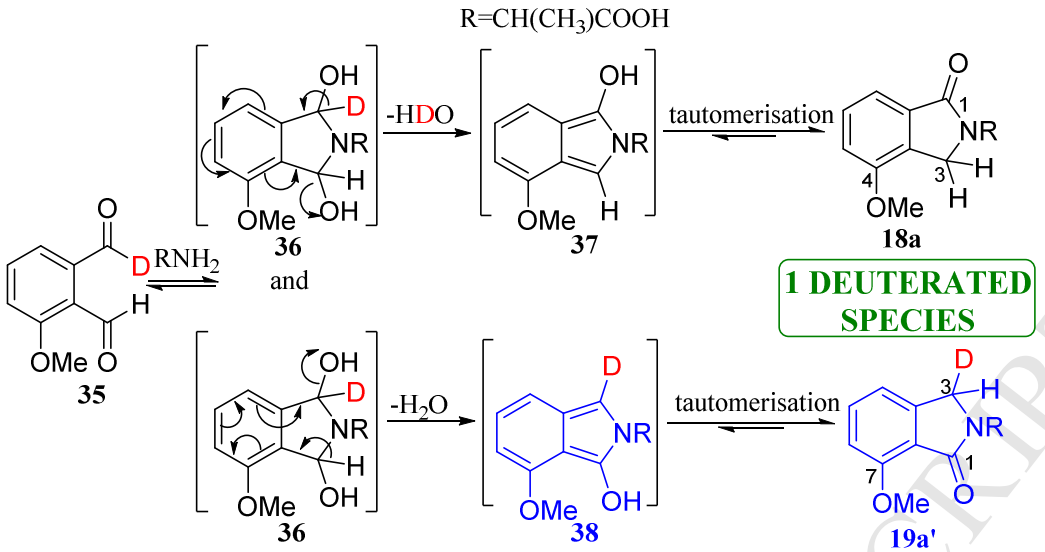
Scheme 4: Considering the two most frequently proposed mechanisms in the literature[5-6]: (i) if the reaction proceeds via **32**, the electron-donating properties of the methoxy-substituent should favour the displacement of the C1-hydroxyl in **32**, leading to the major isomer **21a** (as observed) or (ii) if the reaction proceeds via **33** and **34**, the electron-donating properties of the methoxy-substituent should decrease the reactivity of the C1-formyl in **20a**, favouring **34** and leading to **22a** as the major product (not observed).

Whilst drawing clear mechanistic conclusion from the 3-monosubstituted *o*-phthaldialdehyde series proved challenging (see SI1 part V.1. for a more detailed discussion), the results obtained with the 4-substituted series proved insightful (Table 1). Considering **20a** as an example (Scheme 4), the electron-donating properties of the methoxy-substituent should favour the displacement of the C1-hydroxyl group in the proposed isoindolinediol intermediate **32** leading to the predicted formation of regioisomer **21a** as the major product (as observed, see Table 1). In contrast, the electron-donating properties of the methoxy-substituent would be expected to decrease the reactivity of the C1-formyl group[11] in **20a**

favouring the formation of the proposed imine intermediate **34** over imine **33** and therefore predicting that **22a** should be the major product (not observed). In this study, **21a** was obtained as the major regioisomer providing evidence to support the initial conversion of **20a** to the isoindolinediol intermediate **32**. The same reasoning could explain the switch for the formation of the regioisomer **22e** as the major product when the strongly electron-withdrawing nitro-substituent was used (substrate **20e**, Table 1 and SI1 part V.2 for a more detailed discussion). This conclusion is also supported by Pan *et al.* and Jones *et al.*'s studies reporting the ESI-MS^2 and the ^{1}H NMR spectroscopy detection of an isoindolinediol intermediate in a related transformation.[5a, 5b]

Assuming an isoindolinediol intermediate is initially formed in this reaction, two routes have been proposed in the literature to form the final isoindolinones.[5a,5c,5e,5f] One of these involving a [1,3]-hydride shift was ruled out based on stereoelectronic arguments (see reference 5(f) for additional comment). Detailed consideration of the other possible mechanism and inspired by the reaction of the mono-substituted *o*-phthaldialdehyde **17a**, it was proposed that the use of a mono-deuterated mono-methoxysubstituted substrate such as **35** (Scheme 5) could potentially provide additional mechanistic information. If the proposed reaction pathway[5a, 5d] was followed, the isoindolinediol intermediate **36** formed from **35** could undergo either a loss of HDO leading to **37**, and subsequently to **18a**, or a loss of water leading to **38**, and subsequently to the deuterated **19a'** (Scheme 5). In this case, only one of the two predicted products is deuterated. Analogous reasoning can be applied to the reaction of mono-deuterated **39** (Scheme 6) and would predict the formation of one mono-deuterated isomer **18a'** along with the non-deuterated **19a**.

<<Single column>>



Scheme 5: The predicted products formed from mono-deuterated **35** considering the likely isoindolinediol intermediate **36**. Loss of HDO or water from **36** would lead to **37** or **38** and then to **18a** and **19a'** respectively via tautomerisation. An alternative mechanism proposed in the literature[5a,5c,5e,5f] involving a a [1,3]-deuteride or a [1,3]-hydride shift was ruled out based on stereoelectronic considerations in response to a comment during the review of this paper (see reference 5(f)).

Mono-deuterated **35** and **39** were successfully synthesised with the key deuterium incorporation step being achieved using sodium borodeuteride reduction of lactones **40** and **41** (Figure 4a, SI1 part II.1.1. and part VI.2.)[12] followed by Swern oxidation[13] of **42** and **43** respectively. ^{1}H NMR spectra of **35** and **39** confirmed the presence of a monodeuterated aldehyde (Figure 4b). However, a signal consistent with potential traces of non-deuterated **17a** was observed in the ^{13}C NMR spectrum of **39** (see SI2 for spectrum). Analysis of the condensation of **35** and **39** with alanine (**16**) at reflux was carried out by ^{13}C NMR spectroscopy (Figure 4c). The spectra associated with both purified reaction mixtures differed from that of pure **18a** and **19a** in the signals that corresponded to the $\underline{C}H_2$ carbon only (red circle, Figure 4c). The ^{13}C NMR analysis of the reaction of **35** (spectrum ii) showed a major

isomer with a singlet corresponding to the $\underline{C}H_2$ signal of **18a** (*c.f.* spectrum i and ii). The observed triplet (blue bars, spectrum ii) suggested that the minor isomer was monodeuterated and thus corresponded to **19a'**. Similarly, ^{13}C NMR analysis of the reaction of **39** (spectrum iii) showed the presence of the undeuterated minor isomer **19a** (singlet, *c.f.* spectrum iv with iii) and the monodeuterated major isomer **18a'** (triplet, blue bars, spectrum iii) with a small additional signal (labelled * in spectrum iii) assigned to the formation of trace amounts of undeuterated **18a** (*c.f.* spectrum iii with i). Therefore, in both reactions, only one monodeuterated product was formed consistent with the reaction occurring as shown in Scheme 5. Further evidence to support this comes from work previously published by Pan *et al.*[5a] The presence of trace amounts of undeuterated **18a** when using substrate **39** could potentially be explained by the initial contamination of substrate **39** with traces of undeuterated substrate **17a** or by the reversibility of the final tautomerisation step. Interestingly, the regioisomeric ratios obtained with substrates **35** and **39** were slightly affected by the presence of a deuterium atom suggesting a kinetic isotope effect (SI1 part VI.3.).

3. Conclusions

Although unstable and challenging to obtain, eleven monosubstituted *o*-phthaldialdehydes were successfully synthesised. The result of their condensation reaction with alanine (**16**) led to the design of a disubstituted analogue **27** that reacted with **16** in a highly regioselective manner (rr of 88:12 on 30 mg scale). A gram-scale reaction using this substrate **27** followed by esterification led to the isolation of pure isoindolinone regioisomers in high yield and selectivity (rr of 84:16). Performing this condensation at room temperature provided an excellent regioisomeric ratio of 97:3.

<< single column>>

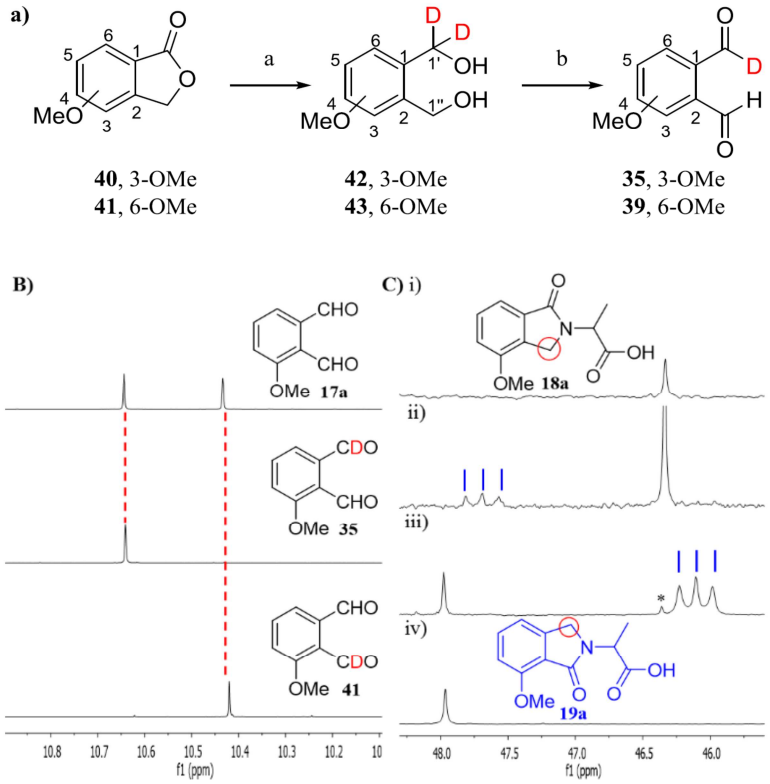
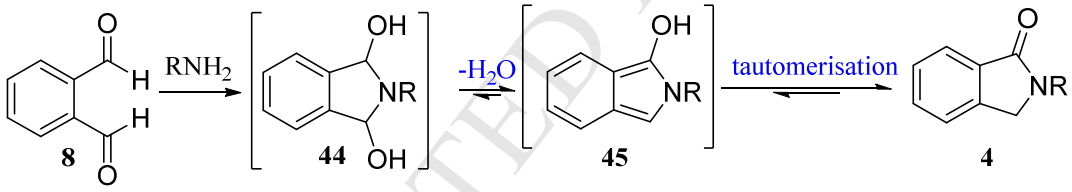


Figure 4: a) Synthesis of monodeuterated substrates **35** and **39**. Reagents and conditions: (a) $NaBD_4$, $ZnCl_2$, *N*,*N*-dimethylaniline, anhyd. THF, reflux, 20 h, 62% for **42**, 49% for **43**; (b) $(COCl)_2$, anhyd. DCM, anhyd. DMSO, Et_3N, rt, 15–17 h, 67% for **35**, 86% for **39**; **b)** Aldehyde region of the ^1H NMR spectra of **17a**, **35** and **39**. Only one aldehyde signal is present in the ^1H NMR spectra of **35** and **39** (red lines) compared to the ^1H NMR spectrum of **17a** supporting the formation of monodeuterated substrates; **C)** CH_2 region of the ^{13}C NMR spectrum of: i) **18a**; ii) the purified mixture of isoindolinones obtained from the condensation of **35** with alanine (**16**); iii) the purified mixture of isoindolinones obtained from the condensation of **39** with alanine (**16**); iv) **19a**. The condensation reaction of **35** led to a major isomer with a singlet corresponding to **18a** and a minor isomer with a triplet (blue bars, spectrum ii) corresponding to monodeuterated **19a'**. Similarly, the condensation reaction of **39** led to a minor isomer with a singlet corresponding to **19a** and a major isomer with a triplet

(blue bars, spectrum iii) corresponding to monodeuterated **18a'**. *^{13}C NMR signal consistent with traces of undeuterated **18a**.

In addition, the structure of the major regioisomers of the condensation reactions of 4-monosubstituted substrates supported the view that this reaction most likely occurs through an isoindolinediol intermediate (**44** from **8**,Scheme 6). Two routes have been proposed in the literature to form the final isoindolinone product from such an intermediate.[5a,5c,5d] The synthesis and condensation reaction of two additional monodeuterated substrates supported a mechanism that most likely occurs *via* loss of water from **44** and a subsequent tautomerisation of **45** to give the more thermodynamically stable **4** (Scheme 6).

<< single column>>



Scheme 6: Proposed condensation reaction mechanism based on this study. Addition of amine RNH_2 to o-phthaldialdehyde (**8**) leads to isoindolinediol intermediate **44.** A water loss followed by tautomerisation of **45** would then provide the desired product **4**.

4. Experimental section

General methods: All solvents and chemicals were purchased from Sigma Aldrich, Alfa Aeser, Acros Organics, Fluorochem, Apollo Scientific or Fisher Scientific and used without further purification. All air or moisture sensitive reactions were carried out in flame or oven-

dried glassware under a positive pressure of nitrogen. Thin layer chromatography (TLC) analysis was performed using glass plates coated with silica gel (with fluorescent indicator UV_{254}). Developed plates were air dried and analysed under a UV lamp (254/365 nm). Flash chromatography was performed using silica gel (40-62 μm, Fluorochem). Melting points were measured using Electrothermal 9100 capillary melting point apparatus. Values are quoted to the nearest 1 °C and are uncorrected. Fourier Transform infra-red spectra (FT IR) were acquired on a Shimadzu IR Affinity-1 Fourier transform IR spectrophotometer using thin films, a Pike MIRacle ATR accessory and Shimadzu IR solution v1.50 for analysis. Absorption maxima (ν_{max}) are reported in wavenumbers (cm^{-1}). Broad signals were assigned with as br. Nuclear magnetic resonance (NMR) spectra were acquired at room temperature on a Brüker Advance 500 (^{1}H, 500 MHz; ^{13}C, 125 MHz), a Brüker Advance 400 (^{1}H, 400 MHz; ^{13}C, 100 MHz) or a Brüker Advance 300 (^{1}H, 300 MHz; ^{13}C, 75 MHz) instruments. Deuterium solvents were used as lock for all NMR spectra and ^{1}H NMR shifts were internally referenced to the solvent. Chemical shifts are expressed as δ in units of ppm. ^{13}C NMR spectra were recorded using the PENDANT sequence mode. Data processing was carried out using MestReNova 9.0 NMR program (Brüker UK Ltd). Signals for protons and carbons were assigned, as far as possible, by using the following two-dimensional NMR spectroscopy techniques: [^{1}H, ^{1}H] COSY (COrrelation SpectroscopY), [^{1}H, ^{13}C] HSQC (Heteronuclear Single Quantum Coherence) and [^{1}H, ^{13}C] HMBC (Heteronuclear Multiple Bond Connectivity). For ^{1}H NMR, the multiplicity is indicated by the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br. = broad. The CH_2 proton NMR signals corresponding to the isoindolinone ring appeared as a singlet, a multiplet or as two apparent doublets with a roofing effect which are referred to as doublets hereafter. All ^{13}C NMR signals were singlets (except for compounds containing F) and correspond to one carbon unless otherwise stated. Mass spectrometric analysis (electrospray

mode, ES, or atmospheric solids analysis probe mode, ASAP) was carried out on a high performance orthogonal acceleration reflecting TOF (Time Of Flight) mass spectrometer operating in positive and negative mode, coupled to a Waters 2975 HPLC.

General procedures

General procedure A: Swern oxidation

To a solution of $(COCl)_2$ (2.6 equiv.) in anhydrous DCM (1.4 mL per mmol of diol) was added dropwise a solution of anhydrous DMSO (5.2 equiv.) in anhydrous DCM (2.2 mL per mmol of diol) at –78 °C and under a nitrogen atmosphere. After 10 mins. stirring at –78 °C, a solution of diol (1.0 equiv.) in anhydrous DCM (2.2 mL per mmol of diol) was added at this temperature. The reaction mixture was stirred at –78 °C for 1 h under a nitrogen atmosphere before Et_3N (1.4 mL per mmol of diol) was added dropwise. The reaction mixture was warmed to rt and stirred for 2–18 h under a nitrogen atmosphere. The reaction was then quenched by addition of H_2O. The organics were extracted with DCM, combined, washed with water, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo* to afford the desired *o*-phthaldialdehyde.

General procedure B: Condensation reaction

Alanine (**16**, 1.2 equiv.) was added to a solution of unsymmetrical *o*-phthaldialdehyde (1.0 equiv.) in anhydrous MeCN (3.8 mL per mmol of *o*-phthaldialdehyde). The reaction mixture was heated at reflux for 4 h under a nitrogen atmosphere. The solution was then cooled to rt before being concentrated *in vacuo* to afford the crude mixture of regioisomers.

Quantitative ^{1}H NMR analysis (SI1 part III.1) was carried out on this crude reaction mixture to obtain the average percentage of the major regioisomer with its 95% confidence interval (SI1 part III.3) and subsequently the regioisomeric ratio (rr, Table 1). The NMR spectrum was processed using MestReNova 9.0 as follows: (i) the baseline was corrected using the Whittaker smoother; (ii) the peaks were selected by manual threshold following a Global

Spectral Deconvolution (GSD) with a refinement level of 4 (20 fitting cycles). When the peaks were too overlapped to allow an accurate deconvolution, they were not used in the subsequent steps; (iii) the integral for each peak was calculated using the peaks method and applying an automatic integration; (iv) when the integrals of two distinguishable signals were overlapping, manual removal of one signal allowed the integral of the remaining signal to be obtained. For each signal, the contribution was calculated relative to one proton (CH_2 integrals were divided by 2, CH_3 integrals were divided by 3 and the integration of an aromatic signal corresponding to two protons belonging to the same regioisomer was divided by 2). Using as many signals as possible, an average regioisomeric ratio with its standard deviation was then calculated using Excel Microsoft Office software. The desired 95% confidence interval was then calculated using the following formula: =CONFIDENCE.NORM(0,05;standard deviation; number of ratio calculated). The crude product was then purified by column chromatography (0–10% MeOH in DCM then 0–10% MeOH in DCM with 1% CH_3COOH) to afford the pure regioisomeric mixture and to enable calculation of the yield.

General procedure C: Reduction of lactone using $NaBD_4$

To a solution of lactone (1 equiv.) in anhydrous THF (1 mL per mmol of lactone) under a nitrogen atmosphere were added cautiously $NaBD_4$ (2 equiv.), $ZnCl_2$ (1 equiv.) and *N*,*N*-dimethylaniline (1 equiv.). The reaction was refluxed for 20 h under a nitrogen atmosphere before being cooled to 0 °C and quenched cautiously with an aqueous solution of NH_4Cl (10%). The grey solid formed was filtered off and washed with DCM. The filtrate was extracted with DCM. The organics were combined, washed with water, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo* to afford the crude product. Purification by

column chromatography (0–100% EtOAc in petroleum ether) afforded the desired deuterated diol.

Compounds synthesised

3-Methoxyphthalaldehyde (17a)[13]

17a was synthesised according to general procedure **A** using (3-methoxy-1,2-phenylene)dimethanol (**23a**, 1.0 equiv., 500 mg, 2.97 mmol), $(COCl)_2$ (2.6 equiv., 0.66 mL, 7.73 mmol), DMSO (5.2 equiv., 1.10 mL, 15.40 mmol). Additional DMSO (1 mL) was required to help **23a** solubilisation. The reaction was stirred at rt for 17 h. Pure **17a** was obtained after recrystallisation from EtOAc as a yellow gum (249 mg, 1.52 mmol, 51%).

^{1}H NMR (400 MHz, $CDCl_3$) δ_H 10.62 (s, 1H, H1''), 10.41 (s, 1H, H1'), 7.63 (t, *J* = 8.0 Hz, 1H, H5), 7.42 (d, *J* = 8.0 Hz, 1H, H6), 7.23 (d, *J* = 8.0 Hz, 1H, H4), 3.97 (s, 3H, OCH_3). Spectral data in accordance with those reported in the literature.[13]

2-(4-Methoxy-1-oxoisoindolin-2-yl)propanoic acid (18a) with 2-(7-methoxy-1-oxoisoindolin-2-yl)propanoic acid (19a)

A mixture of **18a** and **19a** was synthesised according to general procedure **B** using 4-methoxyphthalaldehyde (**17a**, 1.0 equiv., 30 mg, 0.18 mmol) and alanine (**16**, 1.2 equiv., 19 mg, 0.22 mmol). A pure regioisomeric mixture of **18a** and **19a** was obtained as a light yellow solid (32 mg, 0.14 mmol, 78%). Each regioisomer was also synthesised in pure form *via* a different method (SI1 part III.2) and their characterisation is reported here.

18a: **Mp:** 205–208 °C; **IR** ν_{max} cm^{-1} (thin film) 2943 (O-H stretch), 1721 (C=O stretch), 1628 (C=O stretch), 1605 (C=C stretch), 1497 (C=C stretch), 1275 (C-O stretch), 1061, 806; **^{1}H NMR** (400 MHz, CD_3OD) δ_H 7.48 (t, *J* = 8.0 Hz, 1H, H6), 7.36 (d, *J* = 7.8 Hz, 1H, H7), 7.19 (d, *J* = 8.0 Hz, 1H, H5), 4.99 (q, *J* = 7.5 Hz, 1H, CH), 4.53 (d, *J* = 17.5 Hz, 1H, H3), 4.47 (d, *J* = 17.5 Hz, 1H, H3), 3.93 (s, 3H, OCH_3), 1.62 (d, *J* = 7.5 Hz, 3H, CH_3); **^{13}C NMR** (100 MHz, CD_3OD) δ_c 174.7 (COOH), 171.0 (C1), 156.2 (C4), 134.5 (C7a), 131.4 (C3a), 131.1

(C6), 116.2 (C7), 114.6 (C5), 56.1 (OCH_3), 51.0 (CH), 46.3 (C3), 15.9 (CH_3); **HRMS** (ES^+) *m/z* calculated for $C_{12}H_{14}NO_4$ $[M+H]^+$: 236.0917; found: 236.0918. See SI1 part IX for experimental procedure and SI2 for ^1H and ^{13}C NMR spectra.

19a: **Mp**: 218–220 °C; **IR** ν_{max} cm^{-1} (thin film) 2934 (O-H stretch), 1728 (C=O stretch), 1636 (C=O stretch), 1612 (C=C stretch), 1489 (C-H bend), 1456 (C-H bend), 1279 (C-N stretch), 1204 (C-O stretch), 1180 (C-O stretch), 1084, 773; **^{1}H NMR** (500 MHz, CD_3OD) δ_H 7.55 (t, *J* = 8.0 Hz, 1H, H5), 7.11 (d, *J* = 8.0 Hz, 1H, H4), 7.02 (d, *J* = 8.0 Hz, 1H, H6), 4.96 (q, *J* = 7.5 Hz, 1H, CH), 4.54 (d, *J* = 17.5 Hz, 1H, H3), 4.47 (d, *J* = 17.5 Hz, 1H, H3), 3.91 (s, 3H, OCH_3), 1.59 (d, *J* = 7.5 Hz, 3H, CH_3); **^{13}C NMR** (125 MHz, CD_3OD) δ_c 174.9 (COOH), 170.0 (C1), 158.8 (C7), 146.3 (C3a), 135.0 (C5), 120.1 (C7a), 116.2 (C4), 111.2 (C6), 56.0 (OCH_3), 50.7 (CH), 48.0 (C3), 15.8 (CH_3); **HRMS** (ES^+) *m/z* calculated for $C_{12}H_{14}NO_4$ $[M+H]^+$: 236.0917; found: 236.0917. See SI1 part IX for experimental procedure and SI2 for ^1H and ^{13}C NMR spectra.

2-(4-Methyl-1-oxoisoindolin-2-yl)propanoic acid (18b) with 2-(7-methyl-1-oxoisoindolin-2-yl)propanoic acid (19b) *via* 3-methylphtalaldehyde (17b)

17b was synthesised according to general procedure **A** using (3-methyl-1,2-phenylene)dimethanol (**23b**, 1.0 equiv., 100 mg, 0.66 mmol), $(COCl)_2$ (2.6 equiv., 0.15 mL, 1.77 mmol) and DMSO (5.2 equiv., 0.24 mL, 3.43 mmol). The reaction was stirred at rt for 2 h. Crude **17b** was obtained contaminated with a small amount of impurities (73 mg, 0.49 mmol, 75% assuming pure **17b**) Characterisation was performed on this slightly impure mixture with the ^1H and ^{13}C NMR signals reported for **17b** only.

IR ν_{Max} cm^{-1} (Thin Film) 2922 (C-H stretch), 1690 (C=O stretch), 1481 (C-H bend), 1358, 1260, 1086, 1067, 1038; 907, 866, 775; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 10.59 (s, 1H, H1’’), 10.21 (s, 1H, H1’), 7.72 (d, *J* = 7.5 Hz, 1H, H6), 7.57 (t, *J* = 7.5 Hz, 1H, H5), 7.49 (d, *J* = 7.5 Hz, 1H, H4), 2.59 (s, 3H, CH_3); **^{13}C NMR** (125 MHz, $CDCl_3$) δ_C 193.7 (C1’’), 192.8 (C1’),

140.0 (C3), 137.3 (C1), 136.7 (C4), 135.9 (C2), 132.4 (C5), 129.3 (C6), 19.5 (CH_3); **HRMS** (ES^+) *m/z* calculated for $C_9H_9O_2$ $[M+H]^+$: 149.0597; found: 149.0595. See SI2 for ^1H and ^{13}C NMR spectra.

A portion of this impure mixture (1.0 equiv., 30 mg, 0.20 mmol assuming pure **17b**) was treated according to general procedure **B** using alanine (**16**, 1.2 equiv., 22 mg, 0.24 mmol). A pure regioisomeric mixture of **18b** and **19b** was obtained as a light yellow solid (29 mg, 0.13 mmol, 49% extrapolated yield over 2 steps). The ^1H and ^{13}C NMR signals integration are given assuming a mixture of **18b**:**19b** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2916 (O-H stretch), 2361 (C-H stretch), 1734 (C=O stretch), 1636 (C=O stretch), 1609 (C=C stretch), 1456 (O-H bend), 1211 (C-O stretch), 1175 (C-O stretch); **^{1}H NMR** (400 MHz, CD_3OD) δ_H 7.60 (dd, *J* = 6.0, 2.9 Hz, 1H, H7 in **18b**), 7.45 (t, *J* = 7.5 Hz, 1H, H5 in **19b**), 7.42-7.38 (m, 2H, H5 and H6 in **18b**), 7.36 (d, *J* = 7.5 Hz, 1H, H4 in **19b**), 7.22 (d, *J* = 7.5 Hz, 1H, H6 in **19b**), 5.02 (q, *J* = 7.5 Hz, 1H, CH in **18b**), 4.98 (q, *J* = 7.5 Hz, 1H, CH in **19b**), 4.58-4.45 (m, 4H, H3 in **18b**, H3 in **19b**), 2.65 (s, 3H, H1' in **19b**), 2.39 (s, 3H, H1' in **18b**), 1.64 (d, *J* = 7.5 Hz, 3H, CH_3 in **18b**), 1.60 (d, *J* = 7.5 Hz, 3H, CH_3 in **19b**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 175.0 (COOH in **19b**), 174.8 (COOH in **18b**), 171.6 (C1 in **19b**), 171.4 (C1 in **18b**), 144.2 (C3a in **19b**), 142.5 (C4 in **18b**), 138.4 (C7 in **19b**), 134.4 (C3a or C7a in **18b**), 133.9 (C5 in **18b**), 132.7 (2C, C5 in **19b**, C3a or C7a in **18b**), 130.9 (C6 in **19b**), 130.1 (C7a in **19b**), 129.4 (C6 in **18b**), 121.73 (C4 in **19b** or C7 in **18b**), 121.67 (C4 in **19b** or C7 in **18b**), 51.0 (CH in **18b**), 50.7 (CH in **19b**), 47.8 (C3 in **19b**), 47.7 (C3 in **18b**), 17.52 (C1' in **18b** or **19b**), 17.45 (C1' in **18b** or **19b**), 16.0 (CH_3 in **18b**), 15.9 (CH_3 in **19b**); **HRMS** (ES^-) *m/z* calculated for $C_{12}H_{12}NO_3$ $[M-H]^-$: 218.0823; found: 218.0823. See SI2 for ^1H and ^{13}C NMR spectra.

2-(1-Oxo-4-(trifluoromethyl)isoindolin-2-yl)propanoic acid (18c) with 2-(1-oxo-7-(trifluoromethyl)isoindolin-2-yl)propanoic acid (19c) *via* 3-(trifluoromethyl)phthalaldehyde (17c)

To a solution of 2-(1,3-dioxolan-2-yl)-6-(trifluoromethyl)benzaldehyde (**24c**, 1 equiv., 100 mg, 0.41 mmol) in DCM (5 mL) was added an aqueous solution of HCl (2 M, 5 mL). The reaction was stirred at rt for 15 h. Water was added and the organics were extracted with DCM, combined, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. **17c** was obtained impure (24 mg). The considerable amount of impurities prevented any characterisation. Instead, this impure mixture (1.0 equiv., 24 mg, 0.12 mmol assuming pure **17c**) was directly treated according to general procedure **B** using alanine (**16**, 1.2 equiv., 13 mg, 0.14 mmol). A pure regioisomeric mixture of **18c** and **19c** was obtained as a light brown oil (12 mg, 0.04 mmol, 11% yield over 2 steps). The small amount of product available did not allow the identification of the ^{13}C NMR signals corresponding to C4, CF_3 in **18c** and C7, CF_3 in **19c** and allowed the visibility of grease impurities in the ^{1}H and ^{13}C NMR spectra. Additionally, the ^{13}C NMR signals corresponding to COOH in **18c** and in **19c** were extrapolated from the HMBC spectrum (SI1 part VII). The ^{1}H and ^{13}C NMR signal integration are given assuming a mixture of **18c**:**19c** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2922 (O-H stretch), 2359 (C-H stretch), 1684 (br. C=O stretch), 1558 (C=C stretch), 1456 (O-H bend), 1325 (C-F stretch), 1117 (C-F stretch); **^{1}H NMR** (400 MHz, CD_3OD) δ_H 8.04 (d, *J* = 7.5 Hz, 1H, H7 in **18c**), 7.93 (d, *J* = 7.5 Hz, 1H, H5 in **18c**), 7.86 (d, *J* = 7.5 Hz, 1H, H4 in **19c**), 7.82 (d, *J* = 7.5 Hz, 1H, H6 in **19c**), 7.78-7.72 (m, 2H, H5 in **19c**, H6 in **18c**), 5.03-4.97 (m, 2H, CH in **18c**, CH in **19c**), 4.83 (d, *J* = 18.0 Hz, 1H, H3 in **18c**), 4.74 (d, *J* = 18.0 Hz, 1H, H3 in **18c**), 4.70 (d, *J* = 17.5 Hz, 1H, H3 in **19c**), 4.62 (d, *J* = 17.5 Hz, 1H, H3 in **19c**), 1.64 (d, *J* = 7.5 Hz, 3H, CH_3 in **18c**), 1.63 (d, *J* = 8.0 Hz, 3H, CH_3 in **19c**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 174.0 (2C, COOH in **18c**, COOH in **19c**,

extrapolated from the HMBC spectrum, SI1 part VII), 169.2 (C1 in **18c**), 167.4 (C1 in **19c**), 146.4 (C3a in **19c**), 140.7 (C3a in **18c**), 135.1 (C7a in **18c**), 132.9 (C5 in **19c**), 130.3 (C6 in **18c**), 129.9-129.8 (m, C5 in **18c**), 128.5 (C4 in **19c**), 128.3 (C7 in **18c**), 126.8 (C7a in **19c**), 126.6-126.5 (m, C6 in **19c**), 51.7 (CH in **18c** or **19c**), 51.5 (CH in **18c** or **19c**), 48.2 (C3 in **19c**), 47.9 (C3 in **18c**), 16.03 (CH_3 in **18c** or **19c**), 16.00 (CH_3 in **18c** or **19c**); **HRMS** (ES^-) *m/z* calculated for $C_{12}H_9NO_3F_3$ $[M-H]^-$: 272.0540; found:272.0540. See SI2 for ^1H and ^{13}C NMR spectra.

2-(4-Chloro-1-oxoisoindolin-2-yl)propanoic acid (18d) with 2-(7-chloro-1-oxoisoindolin-2-yl)propanoic acid (19d) *via* 3-chlorophtalaldehyde (17d)

17d was synthesised according to general procedure **A** using (3-chloro-1,2-phenylene)dimethanol (**23d**, 1.0 equiv., 100 mg, 0.58 mmol), $(COCl)_2$ (2.6 equiv., 0.13 mL, 1.51 mmol) and DMSO (5.2 equiv., 0.21 mL, 3.02 mmol). The reaction was stirred at rt for 2 h. Crude **17d** was obtained with small impurities (51 mg, 0.30 mmol, 52% assuming pure **17d**) Characterisation was performed on this slightly impure mixture with the ^1H and ^{13}C NMR signals reported for **17d** only.

IR ν_{Max} cm^{-1} (Thin Film) 2924 (C-H stretch), 1695 (C=O stretch), 1591 (C=C stretch), 1460 (C-H bend), 1354, 1292, 1175, 1144, 1078, 910, 862, 773; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 10.64 (s, 1H, H1''), 10.32 (s, 1H, H1'), 7.81 (d, *J* = 8.0 Hz, 1H, H6), 7.69 (d, *J* = 8.0 Hz, 1H, H4), 7.64 (t, *J* = 8.0 Hz, 1H, H5); **^{13}C NMR** (125 MHz, $CDCl_3$) δ_C 191.3 (C1' or C1''), 191.2 (C1' or C1''), 139.0 (C1), 137.5 (C3), 135.1 (C4), 134.3 (C5), 133.8 (C2), 127.7 (C6); **HRMS** (ES^+) *m/z* calculated for $C_8H_6O_2{}^{35}Cl$ $[M+H]^+$: 169.0051; found: 169.0046. See SI2 for ^1H and ^{13}C NMR spectra.

A portion of this impure mixture (1.0 equiv., 30 mg, 0.18 mmol assuming pure **17d**) was treated according to general procedure **B** using alanine (**16**, 1.2 equiv., 19 mg, 0.21 mmol). A pure regioisomeric mixture of **18d** and **19d** was obtained as a yellow gum (27 mg, 0.11

mmol, 33% extrapolated yield over 2 steps). The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture of **18d**:**19d** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2922 (O-H stretch), 2359 (C-H stretch), 1653 (br. C=O stretch), 1558 (C=C stretch), 1456 (O-H bend), 1206 (C-O stretch), 1173 (C-O stretch); **^{1}H NMR** (400 MHz, CD_3OD) δ_H 7.74 (d, *J* = 7.5 Hz, 1H, H7 in **18d**), 7.64 (d, *J* = 8.0 Hz, 1H, H5 in **18d**), 7.59-7.51 (m, 3H, H4 in **19d**, H5 in **19d**, H6 in **18d**), 7.45 (d, *J* = 7.5 Hz, 1H, H6 in **19d**), 5.03-4.94 (m, 2H, CH in **18d**, CH in **19d**), 4.66 (d, *J* = 17.5 Hz, 1H, H3 in **18d**), 4.63 (d, *J* = 17.5 Hz, 1H, H3 in **19d**), 4.57 (d, *J* = 17.5 Hz, 1H, H3 in **18d**), 4.54 (d, *J* = 17.5 Hz, 1H, H3 in **19d**), 1.65 (d, *J* = 7.5 Hz, 3H, CH_3 in **18d**), 1.62 (d, *J* = 7.5 Hz, 3H, CH_3 in **19d**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 175.4 (2C, COOH in **18d**, COOH in **19d**), 169.8 (C1 in **18d**), 168.4 (C1 in **19d**), 146.5 (C3a in **19d**), 141.6 (C3a in **18d**), 135.4 (C7a in **18d**), 134.0 (C5 in **19d**), 133.0 (C5 in **18d**), 132.0 (C7 in **19d**), 131.2 (C6 in **18d**), 130.6 (C6 in **19d**), 130.4 (C4 in **18d**), 129.1 (C7a in **19d**), 123.1 (C4 in **19d**), 123.0 (C7 in **18d**), 51.6 (CH in **18d**), 51.4 (CH in **19d**), 47.71 (C3 in **18d**), 47.68 (C3 in **19d**), 16.1 (CH_3 in **18d**), 15.9 (CH_3 in **19d**); **HRMS** (ES^-) *m/z* calculated for $C_{11}H_9NO_3{}^{35}Cl$ $[M-H]^-$: 238.0276; found: 238.0277. See SI2 for ^{1}H and ^{13}C NMR spectra.

3-Fluorophthalaldehyde (17e)

To a solution of 2,2'-(3-fluoro-1,2-phenylene)bis(1,3-dioxolane) (**25e**, 1 equiv., 100 mg, 0.42 mmol) in DCM (5 mL) was added an aqueous solution of HCl (2 M, 5 mL). The reaction was stirred at rt for 3 h. Water was added and the organics were extracted with DCM, combined, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. Pure **17e** was obtained as a light orange gum (31 mg, 0.20 mmol, 49%).

IR ν_{Max} cm^{-1} (Thin Film) 2928 (C-H stretch), 1694 (C=O stretch), 1603 (C=C stretch), 1479 (C-H bend), 1248 (C-F stretch), 1094, 917, 783; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 10.56 (s, 1H, H1''), 10.49 (s, 1H, H1'), 7.75-7.69 (m, 2H, H5, H6), 7.45-7.41 (m, 1H, H4); **^{13}C NMR**

(125 MHz, $CDCl_3$) δ_C 191.3 (d, *J* = 2.7 Hz, C1'), 188.6 (d, *J* = 8.3 Hz, C1''), 164.4 (d, *J* = 260.1 Hz, C3), 138.3 (C1), 135.8 (d, *J* = 9.5 Hz, C5), 125.0 (d, *J* = 3.3 Hz, C6), 124.3 (d, *J* = 7.8 Hz, C2), 121.6 (d, *J* = 21.8 Hz, C4); **HRMS** ($ASAP^+$) *m/z* calculated for $C_8H_6O_2F$ [M+H]: 153.0352; found: 153.0365. See SI2 for ^1H and ^{13}C NMR spectra.

2-(4-Fluoro-1-oxoisoindolin-2-yl)propanoic acid (18e) with 2-(7-fluoro-1-oxoisoindolin-2-yl)propanoic acid (19e)

A mixture of **18e** and **19e** was synthesised according to general procedure **B** using 3-fluorophthalaldehyde (**17e**, 1.0 equiv., 15 mg, 0.10 mmol) and alanine (**16**, 1.2 equiv., 11 mg, 0.12 mmol). A pure regioisomeric mixture of **18e** and **19e** was obtained as a yellow solid (13 mg, 0.06 mmol, 60%). The ^1H and ^{13}C NMR signals integration are given assuming a mixture of **18e**:**19e** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2918 (O-H stretch), 2359 (C-H stretch), 1749 (C=O stretch), 1653 (C=O stretch), 1558 (C=C stretch), 1489 (C-H bend), 1456 (O-H bend), 1396, 1173 (C-F stretch); **^{1}H NMR** (700 MHz, CD_3OD) δ_H 7.64-7.61 (m, 2H, H7 in **18e**, H5 in **19e**), 7.55 (td, *J* = 8.0, 4.5 Hz, 1H, H6 in **18e**), 7.39 (d, *J* = 7.5 Hz, 1H, H4 in **19e**), 7.36 (t, *J* = 8.0 Hz, 1H, H5 in **18e**), 7.16 (t, *J* = 9.0 Hz, 1H, H6 in **19e**), 5.00-4.94 (m, 2H, CH in **19e**, CH in **18e**), 4.69 (d, *J* = 17.5 Hz, 1H, H3 in **18e**), 4.65-4.63 (m, 2H, H3 in **19e**, H3 in **18e**), 4.58 (d, *J* = 17.5 Hz, 1H, H3 in **19e**), 1.63 (d, *J* = 7.5 Hz, 3H, CH_3 in **18e**), 1.61 (d, *J* = 7.5 Hz, 3H, CH_3 in **19e**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 175.1 (2C, COOH in **18e**, COOH in **19e**), 169.8 (C1 in **18e**), 167.8 (C1 in **19e**), 160.1 (d, *J* = 257.8 Hz, C7 in **19e**), 159.0 (d, *J* = 249.4 Hz, C4 in **18e**), 146.7 (d, *J* = 2.5 Hz, C3a in **19e**), 136.3 (d, *J* = 4.8 Hz, C7a in **18e**), 135.3 (d, *J* = 7.5 Hz, C5 in **19e**), 131.7 (d, *J* = 6.5 Hz, C6 in **18e**), 129.5 (d, *J* = 18.6 Hz, C3a in **18e**), 120.54 (d, *J* = 4.1 Hz, C4 in **19e** or C7 in **18e**), 120.47 (d, *J* = 3.7 Hz, C4 in **19e** or C7 in **18e**), 120.3 (d, *J* = 13.2 Hz, C7a in **19e**) 119.6 (d, *J* = 19.7 Hz, C5 in **18e**), 116.0 (d, *J* = 19.4 Hz, C6 in **19e**), 51.5 (CH in **18e**), 51.1 (CH in **19e**), 48.4 (C3 in **19e**), 45.4 (C3 in **18e**), 15.95 (CH_3 in

18e or **19e**), 15.92 (CH_3 in **18e** or **19e**); **HRMS** (ES^-) *m/z* calculated for $C_{11}H_9NO_3F$ $[M-H]^-$: 222.0572; found: 222.0572. See SI2 for ^1H and ^{13}C NMR spectra.

4-Methoxyphthalaldehyde (20a)[14]

2-(1,3-Dioxolan-2-yl)-4-methoxybenzaldehyde (**24a**, 1 equiv., 2.00 g, 9.6 mmol) was stirred in an aqueous solution of HCl (3 M, 330 mL) at rt for 16 h. The organics were extracted with Et_2O, combined, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo* to afford pure **20a** as an orange solid (1.35 g, 8.2 mmol, 86%).

Mp: 78–81 °C (lit. 75–78 °C)[14]; **^{1}H NMR** (400 MHz, $CDCl_3$) δ_H 10.66 (s, 1H, CHO), 10.33 (s, 1H, CHO), 7.93 (d, *J* = 8.5 Hz, 1H, H6), 7.45 (d, *J* = 3.0 Hz, 1H, H3), 7.22 (dd, *J* = 8.5, 3.0 Hz, 1H, H5), 3.95 (s, 3H, OCH_3). Spectral data in accordance with those reported in the literature.[14]

2-(6-Methoxy-1-oxoisoindolin-2-yl)propanoic acid (21a) with 2-(5-methoxy-1-oxoisoindolin-2-yl)propanoic acid (22a)

A mixture of **21a** and **22a** was synthesised according to general procedure **B** using 4-methoxyphthalaldehyde (**20a**, 1.0 equiv., 30 mg, 0.18 mmol) and alanine (**16**, 1.2 equiv., 19 mg, 0.22 mmol). A pure regioisomeric mixture of **21a** and **22a** was obtained as a light orange solid (38 mg, 0.16 mmol, 90%). Each regioisomer was synthesised in pure form *via* a different method (SI1 part III.2.) and their characterisation is reported here.

21a: **Mp:** 194–196 °C; **IR** ν_{max} cm^{-1} (thin film) 3370 (O-H stretch), 1730 (C=O stretch), 1636 (C=O stretch), 1493 (C=C stretch), 1456 (C-H bend), 1447 (C-H bend), 1196 (C-O stretch), 1022, 770; **^{1}H NMR** (500 MHz, CD_3OD) δ_H 7.48 (d, *J* = 8.5 Hz, 1H, H4), 7.29 (d, *J* = 2.5 Hz, 1H, H7), 7.19 (dd, *J* = 8.5, 2.5 Hz, 1H, H5), 5.00 (q, *J* = 7.5 Hz, 1H, CH), 4.54 (d, *J* = 17.0 Hz, 1H, H3), 4.48 (d, *J* = 17.0 Hz, 1H, H3), 3.86 (s, 3H, OCH_3), 1.61 (d, *J* = 7.5 Hz, 3H, CH_3); **^{13}C NMR** (125 MHz, CD_3OD) δ_c 174.8 (COOH), 171.0 (C1), 161.6 (C6), 135.8 (C3a), 134.3 (C7a), 125.2 (C4), 121.0 (C5), 107.4 (C7), 56.1 (OCH_3), 51.2 (CH), 48.1 (C3), 15.9

(CH_3); **HRMS** (ES^+) *m/z* calculated for $C_{12}H_{14}NO_4$ $[M+H]^+$: 236.0917; found: 236.0921. See SI1 part IX for experimental procedure and SI2 for ^1H and ^{13}C NMR spectra.

22a: **Mp:** 191–193 °C; **IR** ν_{max} cm^{-1} (thin film) 3229 (O-H stretch), 1717 (C=O stretch), 1628 (C=O stretch), 1611 (C=C stretch), 1558 (C=C stretch), 1506 (C=C stretch), 1447 (C-H bend), 1435 (C-H bend), 1298 (C-N stretch), 1206 (C-O stretch), 1086, 1026, 845, 775; **^{1}H NMR** (500 MHz, CD_3OD) δ_H 7.68 (d, *J* = 8.5 Hz, 1H, H7), 7.13 (d, *J* = 2.0 Hz, 1H, H4), 7.04 (dd, *J* = 8.5, 2.2 Hz, 1H, H6), 4.98 (q, *J* = 7.5 Hz, 1H, CH), 4.57 (d, *J* = 17.0 Hz, 1H, H3), 4.50 (d, *J* = 17.0 Hz, 1H, H3), 3.88 (s, 3H, OCH_3), 1.60 (d, *J* = 7.5 Hz, 3H, CH_3); **^{13}C NMR** (125 MHz, CD_3OD) δ_c 174.8 (COOH), 171.0 (C1),164.9 (C5), 146.2 (C3a), 125.6 (C7), 125.4 (C7a), 116.3 (C6), 108.8 (C4), 56.2 (OCH_3), 50.9 (CH), 48.4 (C3), 15.9 (CH_3); **HRMS** (ES^+) *m/z* calculated for $C_{12}H_{14}NO_4$ $[M+H]^+$: 236.0917; found: 236.0921. See SI1 part IX for experimental procedure and SI2 for ^1H and ^{13}C NMR spectra.

4-Fluorophthalaldehyde (20b)[14]

20b was synthesised according to general procedure **A** using (4-fluoro-1,2-phenylene)dimethanol (**26b**, 1.0 equiv., 200 mg, 1.28 mmol), $(COCl)_2$ (2.6 equiv., 0.3 mL, 3.33 mmol), DMSO (5.2 equiv., 0.5 mL, 6.66 mmol). The reaction was stirred at rt for 17 h. Pure **20b** was obtained as a yellow gum (150 mg, 0.99 mmol, 77%).

^{1}H NMR (400 MHz, $CDCl_3$) δ_H 10.60 (d, *J* = 2.0 Hz, 1H, CHO), 10.42 (s, 1H, CHO), 8.02 (dd, *J* = 8.5, 5.2 Hz, 1H, H6), 7.67 (dd, *J* = 8.5, 2.5 Hz, 1H, H3), 7.45 (ddd, *J* = 8.5, 7.5, 2.5 Hz, 1H, H5). Spectral data in accordance with those reported in the literature.[14]

2-(6-Fluoro-1-oxoisoindolin-2-yl)propanoic acid (21b) with 2-(5-fluoro-1-oxoisoindolin-2-yl)propanoic acid (22b)

A mixture of **21b** and **22b** was synthesised according to general procedure **B** using 4-fluorophthalaldehyde (**20b**, 1.0 equiv., 22 mg, 0.14 mmol) and alanine (**16**, 1.2 equiv., 15 mg, 0.17 mmol). A pure regioisomeric mixture of **21b** and **22b** was obtained as an orange oil (30

mg, 0.13 mmol, 93%). The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture of **21b**:**22b** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2924 (O-H stretch), 1732 (C=O stretch), 1647 (br. C=O stretch), 1487 (C-H bend), 1456 (O-H bend), 1207 (C-F stretch), 1179 (C-O stretch), 772; **^{1}H NMR** (700 MHz, CD_3OD) δ_H 7.79 (dd, *J* = 8.5, 4.9 Hz, 1H, H7 in **22b**), 7.60 (dd, *J* = 8.5, 4.5 Hz, 1H, H4 in **21b**), 7.46 (d, *J* = 8.0 Hz, 1H, H7 in **21b**), 7.38-7.34 (m, 2H, H4 in **22b**, H5 in **21b**), 7.25 (t, *J* = 8.5 Hz, 1H, H6 in **22b**), 5.01-4.96 (m, 2H, CH in **21b**, CH in **22b**), 4.65-4.53 (m, 4H, H3 in **21b**, H3 in **22b**), 1.612 (d, *J* = 7.5 Hz, 3H, CH_3 in **21b**), 1.607 (d, *J* = 7.5 Hz, 3H, CH_3 in **22b**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 174.9 (COOH in **22b**), 174.8 (COOH in **21b**), 169.9 (C1 in **21b** or **22b**), 169.8 (C1 in **21b** or **22b**), 166.8 (d, *J* = 250.1 Hz, C5 in **22b**), 164.2 (d, *J* = 245.5 Hz, C6 in **21b**), 146.5 (d, *J* = 10.6 Hz, C3a in **22b**), 139.4 (C3a in **21b**), 135.1 (d, *J* = 8.6 Hz, C7a in **21b**), 129.3 (C7a in **22b**), 126.5 (d, *J* = 9.8 Hz, C7 in **22b**), 126.3 (d, *J* = 8.4 Hz, C4 in **21b**), 120.5 (d, *J* = 23.9 Hz, C5 in **21b**), 116.8 (d, *J* = 23.9 Hz, C6 in **22b**), 111.6 (d, *J* = 24.7 Hz, C4 in **22b**), 110.6 (d, *J* = 23.8 Hz, C7 in **21b**), 51.3 (CH in **21b**), 51.2 (CH in **22b**), 48.3 (C3 in **22b**), 48.2 (C3 in **21b**), 15.9 (2C, CH_3 in **21b**, CH_3 in **22b**); **HRMS** (ES^-) *m/z* calculated for $C_{11}H_9NO_3F$ $[M-H]^-$: 222.0572; found: 222.0567. See SI2 for ^{1}H and ^{13}C NMR spectra.

4-Bromophthalaldehyde (20c)[15]

20c was synthesised according to general procedure **A** using (4-bromo-1,2-phenylene)dimethanol (**26c**, 1.0 equiv., 100 mg, 0.46 mmol), $(COCl)_2$ (2.6 equiv., 0.10 mL, 1.20 mmol), DMSO (5.2 equiv., 0.17 mL, 2.40 mmol) and additional DMSO (0.15 mL) was required to help **26c** solubilisation. The reaction was stirred at rt for 14 h. Pure **20c** was obtained as a light yellow solid (86 mg, 0.41 mmol, 88%).

Mp: 95–98 °C (lit. 97–100 °C)[16]; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 10.51 (s, 1H, CHO), 10.45 (s, 1H, CHO), 8.10 (d, *J* = 2.0 Hz, 1H, H3), 7.91 (dd, *J* = 8.0, 2.0 Hz, 1H, H5), 7.84 (d, *J* = 8.0 Hz, 1H, H6). Spectral data in accordance with those reported in the literature.[15]

2-(6-Bromo-1-oxoisoindolin-2-yl)propanoic acid (21c) with 2-(5-bromo-1-oxoisoindolin-2-yl)propanoic acid (22c)

A mixture of **21c** and **22c** was synthesised according to general procedure **B** using 4-bromophthalaldehyde (**20c**, 1.0 equiv., 15 mg, 0.070 mmol) and alanine (**16**, 1.2 equiv., 8 mg, 0.084 mmol). A pure regioisomeric mixture of **21c** and **22c** was obtained as a light orange oil (18 mg, 0.063 mmol, 90%). The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture of **21c**:**22c** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2920 (O-H stretch), 2359 (C-H stretch), 1732 (C=O stretch), 1653 (br. C=O stretch), 1456 (O-H bend), 1418 (C-H bend), 1204 (C-O stretch), 768; **^{1}H NMR** (700 MHz, CD_3OD) δ_H 7.90 (s, 1H, H7 in **21c**), 7.80 (s, 1H, H4 in **22c**), 7.76 (d, *J* = 8.0 Hz, 1H, H5 in **21c**), 7.68-7.67 (m, 2H, H6 in **22c**, H7 in **22c**), 7.53 (d, *J* = 8.0 Hz, 1H, H4 in **21c**), 4.989 (q, *J* = 7.5 Hz, 1H, CH in **21c**), 4.985 (q, *J* = 7.5 Hz, 1H, CH in **22c**), 4.64-4.52 (m, 4H, H3 in **21c**, H3 in **22c**), 1.613 (d, *J* = 7.5 Hz, 3H, CH_3 in **21c**), 1.609 (d, *J* = 7.5 Hz, 3H, CH_3 in **22c**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 174.8 (2C, COOH in **21c**, COOH in **22c**), 169.9 (C1 in **22c**), 169.4 (C1 in **21c**), 145.8 (C3a in **22c**), 142.7 (C3a in **21c**), 136.0 (C5 in **21c**), 135.3 (C7a in **21c**), 132.6 (C6 in **22c**), 132.2 (C7a in **22c**), 127.8 (C4 in **22c**), 127.6 (C5 in **22c**), 127.2 (C7 in **21c**), 126.3 (C4 in **21c**), 125.8 (C7 in **22c**), 122.9 (C6 in **21c**), 51.3 (CH in **21c**), 51.2 (CH in **22c**), 48.4 (C3 in **21c**), 48.2 (C3 in **22c**), 16.0 (2C, CH_3 in **21c**, CH_3 in **22c**); **HRMS** ($ASAP^+$) *m/z* calculated for $C_{11}H_{11}NO_3{}^{79}Br$ [M+H]: 283.9922; found: 283.9919. See SI2 for ^{1}H and ^{13}C NMR spectra.

4-Methylphthalaldehyde (20d)[17]

20d was synthesised according to general procedure **A** using (4-methyl-1,2-phenylene)dimethanol (**26d**, 1.0 equiv., 200 mg, 1.31 mmol), $(COCl)_2$ (2.6 equiv., 0.29 mL, 3.4 mmol), DMSO (5.2 equiv., 0.48 mL, 6.8 mmol). The reaction was stirred at rt for 17 h. Pure **20d** was obtained as an orange gum (174 mg, 1.18 mmol, 90%).

^{1}H NMR (400 MHz, $CDCl_3$) δ_H 10.53 (s, 1H, CHO) 10.44 (s, 1H, CHO), 7.86 (d, *J* = 8.0 Hz, 1H, ArH), 7.75 (s, 1H, ArH), 7.56 (d, *J* = 8.0 Hz, 1H, ArH), 2.50 (s, 3H, CH_3). Spectral data in accordance with those reported in the literature.[17]

2-(6-Methyl-1-oxoisoindolin-2-yl)propanoic acid (21d) with 2-(5-methyl-1-oxoisoindolin-2-yl)propanoic acid (22d)

A mixture of **21d** and **22d** was synthesised according to general procedure **B** using 4-methylphthalaldehyde (**20d**, 1.0 equiv., 26 mg, 0.18 mmol) and alanine (**16**, 1.2 equiv., 19 mg, 0.22 mmol). A pure regioisomeric mixture of **21d** and **22d** was obtained as a yellow solid (29 mg, 0.13 mmol, 73%). The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture of **21d**:**22d** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2914 (O-H stretch), 1734 (C=O stretch), 1638 (C=O stretch), 1456 (O-H bend), 1198 (C-O stretch), 1173 (C-O stretch), 772; **^{1}H NMR** (700 MHz, CD_3OD) δ_H 7.64 (d, *J* = 8.0 Hz, 1H, H7 in **22d**), 7.57 (s, 1H, H7 in **21d**), 7.45 (d, *J* = 8.0 Hz, 1H, H4 in **21d**), 7.43 (d, *J* = 8.0 Hz, 1H, H5 in **21d**), 7.38 (s, 1H, H4 in **22d**), 7.31 (d, *J* = 8.0 Hz, 1H, H6 in **22d**), 5.00-4.96 (m, 2H, CH in **22d**, CH in **21d**), 4.57 (d, *J* = 17.0 Hz, 2H, H3 in **21d**, H3 in **22d**), 4.48 (d, *J* = 17.0 Hz, 2H, H3 in **21d**, H3 in **22d**), 2.45 (s, 3H, H1' in **22d**), 2.43 (s, 3H, H1' in **21d**), 1.59 (d, *J* = 7.5 Hz, 6H, CH_3 in **21d**, CH_3 in **22d**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 175.4 (2C, COOH in **21d**, COOH in **22d**), 171.1 (2C, C1 in **21d**, C1 in **22d**), 144.2 (C3a or C5 in **22d**), 144.1 (C3a or C5 in **22d**), 141.0 (C3a in **21d**), 139.4 (C6 in **21d**), 134.1 (C5 in **21d**), 133.2 (C7a in **21d**), 130.5 (C7a in **22d**), 130.1 (C6 in **22d**), 124.7 (C4 in **22d**), 124.3 (C7 in **21d**), 124.03 (C7 in **22d** or C4 in **21d**), 124.02 (C7 in **22d** or C4 in **21d**),

51.24 (CH in **21d**), 51.17 (CH in **22d**), 48.32 (C3 in **22d**), 48.29 (C3 in **21d**), 21.9 (C1' in **22d**), 21.3 (C1' in **21d**), 16.1 (2C, CH_3 in **21d**, CH_3 in **22d**); **HRMS** ($ASAP^+$) *m/z* calculated for $C_{12}H_{14}NO_3$ [M+H]: 220.0974; found: 220.0973. See SI2 for ^1H and ^{13}C NMR spectra.

2-(6-Trifluoromethyl-1-oxoisoindolin-2-yl)propanoic acid (21e) with 2-(5-trifluoromethyl-1-oxoisoindolin-2-yl)propanoic acid (22e) *via* 4-(trifluoromethyl)phtalaldehyde (20e)

20e was synthesised according to general procedure **A** using (4-(trifluoromethyl)-1,2-phenylene)dimethanol (**26e**, 1.0 equiv., 100 mg, 0.49 mmol), $(COCl)_2$ (2.6 equiv., 0.11 mL, 1.27 mmol), DMSO (5.2 equiv., 0.18 mL, 2.55 mmol). The reaction was stirred at rt for 15 h. Crude **20e** was obtained with small amount of impurities (74 mg, 0.37 mmol, 75% assuming pure **20e**). Characterisation was performed on this slightly impure mixture with the ^1H and ^{13}C NMR signals reported for **20e** only. The small intensity of the $\underline{C}4$ and $\underline{C}F_3$ carbon signals on the ^{13}C NMR spectrum revealed doublets instead of the expected quartets.

IR ν_{Max} cm^{-1} (Thin Film) 2932 (C-H stretch), 1697 (C=O stretch), 1327 (C-F stretch), 1165, 1121 (C-F stretch), 1057, 889, 831; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 10.58 (s, 1H, H1'), 10.55 (s, 1H, H1''), 8.23 (s, 1H, H3), 8.11 (d, *J* = 8.0 Hz, 1H, H6), 8.03 (d, *J* = 8.0 Hz, 1H, H5); **^{13}C NMR** (125 MHz, $CDCl_3$) δ_C 191.3 (C1'), 190.8 (C1''), 138.8 (C1), 136.9 (C2), 135.5 (d, *J* = 33.8 Hz, C4), 131.7 (C6), 130.5 (q, *J* = 3.4 Hz, C5), 128.2 (q, *J* = 3.4 Hz, C3), 123.0 (d, *J* = 273.2 Hz, CF_3); **HRMS** (ES^-) *m/z* calculated for $C_9H_5O_2F_3$ $[M]^-$: 202.0247; found: 202.0247. See SI2 for ^1H and ^{13}C NMR spectra.

A portion of this impure mixture (1.0 equiv., 30 mg, 0.15 mmol assuming pure **20e**) was treated according to general procedure **B** using alanine (**16**, 1.2 equiv., 16 mg, 0.18 mmol). A pure regioisomeric mixture of **21e** and **22e** was obtained as an orange oil (30 mg, 0.11 mmol, 55% extrapolated yield over 2 steps). The ^1H and ^{13}C NMR signals integration are given assuming a mixture of **21e**:**22e** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2924 (O-H stretch), 2359 (C-H stretch), 1717 (C=O stretch), 1653 (br. C=O stretch), 1456 (O-H bend), 1325 (C-F stretch), 1161 (C-O stretch), 1117 (C-F stretch), 1057 (C-O stretch); **^{1}H NMR** (400 MHz, CD_3OD) δ_H 8.05 (s, 1H, H7 in **21e**), 7.96-7.91 (m, 3H, H4 in **22e**, H7 in **22e**, H5 in **21e**), 7.83-7.79 (m, 2H, H6 in **22e**, H4 in **21e**), 5.01 (q, *J* = 7.5 Hz, 2H, CH in **21e**, CH in **22e**), 4.75 (d, *J* = 18.0 Hz, 1H, H3 in **21e**), 4.73 (d, *J* = 18.0 Hz, 1H, H3 in **22e**), 4.66 (d, *J* = 18.0 Hz, 1H, H3 in **21e**), 4.65 (d, *J* = 18.0 Hz, 1H, H3 in **22e**), 1.63 (d, *J* = 7.5 Hz, 6H, CH_3 in **21e**, CH_3 in **22e**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 174.5 (2C, COOH in **21e**, COOH in **22e**), 169.4 (2C, C1 in **21e**, C1 in **22e**), 147.6 (C3a in **21e**), 144.4 (C3a in **22e**), 136.6 (C7a in **22e**), 134.6 (q, *J* = 32.2 Hz, C5 in **22e**), 134.0 (C7a in **21e**), 131.7 (q, *J* = 32.7 Hz, C6 in **21e**), 129.8 (q, *J* = 4.0 Hz, C5 in **21e**), 126.3 (q, *J* = 3.6 Hz, C6 in **22e**), 125.6 (C4 in **21e**), 125.38 (q, *J* = 271.5 Hz, CF_3 in **21e** or **22e**), 125.37 (q, *J* = 272.0 Hz, CF_3 in **21e** or **22e**), 125.1 (C7 in **22e**), 121.8 (q, *J* = 4.2 Hz, C4 in **22e**), 121.2 (q, *J* = 4.2 Hz, C7 in **21e**), 51.24 (CH in **21e** or **22e**), 51.20 (CH in **21e** or **22e**), 48.7 (2C, C3 in **21e**, C3 in **22e**), 15.9 (2C, CH_3 in **21e**, CH_3 in **22e**); **HRMS** (ES^-) *m/z* calculated for $C_{12}H_9NO_3F_3$ $[M-H]^-$: 272.0540; found: 272.0540. See SI2 for ^{1}H and ^{13}C NMR spectra.

2-(6-Nitro-1-oxoisoindolin-2-yl)propanoic acid (21f) with 2-(5-nitro-1-oxoisoindolin-2-yl)propanoic acid (22f) *via* 4-nitrophtalaldehyde (20f)

20f was synthesised according to general procedure **A** using (4-nitro-1,2-phenylene)dimethanol (**26f**, 1.0 equiv., 100 mg, 0.55 mmol), $(COCl)_2$ (2.6 equiv., 0.12 mL, 1.43 mmol) and DMSO (5.2 equiv., 0.20 mL, 2.84 mmol). The reaction was stirred at rt for 18 h. Crude **20f** was obtained impure (110 mg, 0.61 mmol, >100% due to impurities). The considerable amount of impurities prevented any characterisation. Instead, a portion of this impure mixture (1.0 equiv., 30 mg, 0.17 mmol assuming pure **20f**) was directly treated according to general procedure **B** using alanine (**16**, 1.2 equiv., 18 mg, 0.20 mmol). A pure regioisomeric mixture of **21f** and **22f** was obtained as an orange oil (13 mg, 0.05 mmol, 35%

extrapolated yield over 2 steps). The ^{1}H and ^{13}C NMR signal integration are given assuming a mixture of **21f**:**22f** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2922 (O-H stretch), 2357 (C-H stretch), 1732 (C=O stretch), 1651 (br. C=O stretch), 1526 (N-O stretch), 1449 (O-H bend), 1341 (N-O stretch), 1196 (C-O stretch), 816; **^{1}H NMR** (400 MHz, CD_3OD) δ_H 8.57 (d, *J* = 2.0 Hz, 1H, H7 in **21f**), 8.51-8.49 (m, 2H, H4 in **22f**, H5 in **21f**), 8.39 (dd, *J* = 8.5, 2.0 Hz, 1H, H6 in **22f**), 7.98 (d, *J* = 8.5 Hz, 1H, H7 in **22f**), 7.84 (d, *J* = 8.5 Hz, 1H, H4 in **21f**), 5.06-5.00 (m, 2H, CH in **21f**, CH in **22f**), 4.79-4.68 (m, 4H, H3 in **21f**, H3 in **22f**), 1.65 (d, *J* = 7.5 Hz, 3H, CH_3 in **22f**), 1.64 (d, *J* = 7.5 Hz, 3H, CH_3 in **21f**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 174.62 (COOH in **21f** or **22f**), 174.56 (COOH in **21f** or **22f**), 168.8 (C1 in **21f** or **22f**), 168.7 (C1 in **21f** or **22f**), 151.8 (C5 in **22f**), 149.82 (C3a or C6 in **21f**), 149.78 (C3a or C6 in **21f**), 144.9 (C3a in **22f**), 138.4 (C7a in **22f**), 134.7 (C7a in **21f**), 127.8 (C5 in **21f**), 126.0 (C4 in **21f**), 125.4 (C7 in **22f**), 124.7 (C6 in **22f**), 120.1 (C4 in **22f**), 119.5 (C7 in **21f**), 51.6 (CH in **21f** or **22f**), 51.4 (CH in **21f** or **22f**), 48.7 (2C, C3 in **21f**, C3 in **22f**), 15.9 (2C, CH_3 in **21f**, CH_3 in **22f**); **HRMS** ($ASAP^+$) *m/z* calculated for $C_{11}H_{11}N_2O_5$ [M+H]: 251.0668; found: 251.0662. See SI2 for ^{1}H and ^{13}C NMR spectra.

2-(4,6-Dimethoxy-1-oxoisoindolin-2-yl)propanoic acid (28) with 2-(5,7-dimethoxy-1-oxoisoindolin-2-yl)propanoic acid (29)

Small scale condensation reaction at reflux: A mixture of **28** and **29** was synthesised according to general procedure **B** using 3,5-dimethoxyphthalaldehyde (**27**, 1.0 equiv., 30 mg, 0.15 mmol) and alanine (**16**, 1.2 equiv., 17 mg, 0.19 mmol). A pure regioisomeric mixture of **28** and **29** was obtained as a yellow gum (37 mg, 0.14 mmol, 93%).

Small scale condensation reaction at rt: To a solution of 3,5-dimethoxyphthalaldehyde (**27**, 1.0 equiv., 30 mg, 0.15 mmol) in anhydrous MeCN (0.55 mL) was added alanine (**16**, 1.2 equiv., 17 mg, 0.19 mmol) under a nitrogen atmosphere. The reaction was stirred at rt for 29

h before being concentrated *in vacuo* to afford the crude mixture of regioisomers. A quantitative ^{1}H NMR spectrum was acquired on this crude reaction mixture and processed as explained in general procedure **B**. A pure regioisomeric mixture of **28** and **29** was then obtained after purification by column chromatography (0–10% MeOH in DCM then 0–10% MeOH in DCM with 1% CH_3COOH) as a yellow gum (40 mg, 0.15 mmol, quant.).

Gram-scale condensation reaction at reflux: A mixture of **28** and **29** was synthesised according to general procedure **B** using 3,5-dimethoxyphthalaldehyde (**27**, 1.0 equiv., 0.97 g, 5.0 mmol) and alanine (**16**, 1.2 equiv., 0.53 g, 6.0 mmol). A pure regioisomeric mixture of **28** and **29** was obtained as a yellow gum (1.20 g, 4.52 mmol, 90%).

The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture of **28**:**29** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2922 (O-H stretch), 1732 (C=O stretch), 1601 (br. C=O stretch), 1504 (C=C stretch), 1456 (O-H bend), 1202 (C-O stretch), 1146 (C-O stretch), 772; **^{1}H NMR** (700 MHz, CD_3OD) δ_H 6.88 (d, *J* = 2.0 Hz, 1H, H7 in **28**), 6.74 (d, *J* = 2.0 Hz, 1H, H5 in **28**), 6.69 (d, *J* = 2.0 Hz, 1H, H4 in **29**), 6.52 (d, *J* = 2.0 Hz, 1H, H6 in **29**), 4.97 (q, *J* = 7.5 Hz, 2H, CH in **28**, CH in **29**), 4.49 (d, *J* = 17.0 Hz, 1H, H3 in **29**), 4.44 (d, *J* = 17.0 Hz, 1H, H3 in **28**), 4.42 (d, *J* = 17.0 Hz, 1H, H3 in **29**), 4.39 (d, *J* = 17.0 Hz, 1H, H3 in **28**), 3.89 (s, 3H, H1’ in **28**), 3.88 (s, 3H, H1’ in **29**), 3.87 (s, 3H, H1’’ in **29**), 3.85 (s, 3H, H1’’ in **28**), 1.60 (d, *J* = 7.5 Hz, 3H, CH_3 in **28**), 1.57 (d, *J* = 7.5 Hz, 3H, CH_3 in **29**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 174.7 (2C, COOH in **28**, COOH in **29**), 171.0 (C1 in **28**), 170.0 (C1 in **29**), 166.6 (C5 in **29**), 163.5 (C6 in **28**), 159.8 (C7 in **29**), 156.9 (C4 in **28**), 148.3 (C3a in **29**), 135.0 (C7a in **28**), 124.3 (C3a in **28**), 113.3 (C7a in **29**), 103.5 (C5 in **28**), 100.5 (C4 in **29**), 99.1 (C6 in **29**), 98.8 (C7 in **28**), 56.3 (C1’ or C1’’ in **29**), 56.23 (C1’ or C1’’ in **28**), 56.17 (C1’ or C1’’ in **28**), 56.0 (C1’ or C1’’ in **29**), 51.2 (CH in **28**), 50.5 (CH in **29**), 48.1 (C3 in **29**), 46.0 (C3 in

28), 15.92 (CH_3 in **28**), 15.88 (CH_3 in **29**); **HRMS** (ES^+) *m/z* calculated for $C_{13}H_{15}NO_5Na$ $[M+Na]^+$: 288.0842; found: 288.0841. See SI2 for ^1H and ^{13}C NMR spectra.

Methyl 2-(4,6-dimethoxy-1-oxoisoindolin-2-yl)propanoate (30) and methyl 2-(5,7-dimethoxy-1-oxoisoindolin-2-yl)propanoate (31)

To a solution of 2-(4,6-dimethoxy-1-oxoisoindolin-2-yl)propanoic acid (**28**) and 2-(5,7-dimethoxy-1-oxoisoindolin-2-yl)propanoic acid (**29**) (1 equiv., 1.20 g, 4.52 mmol) in anhydrous MeOH (82 mL) was added freshly distilled $SOCl_2$ (8.2 mL) dropwise at 0 °C under a nitrogen atmosphere. The reaction was stirred at rt for 12 h under a nitrogen atmosphere before being concentrated *in vacuo* within a fume cupboard. Separation of the two isomers **30** and **31** was achieved by column chromatography (0–100% EtOAc in petroleum ether). Pure **30** was obtained as a yellow solid (1.06 g, 3.80 mmol, 84%) and pure **31** was obtained as an oil (0.20 g, 0.70 mmol, 16%).

30: Mp: 98–100 °C; **IR** ν_{Max} cm^{-1} (Thin Film) 2949 (C-H stretch), 1740 (C=O stretch), 1682 (C=O stretch), 1605 (C=C stretch), 1503 (C=C stretch), 1329, 1209 (C-O stretch), 1144 (C-O stretch), 1110, 1061 (C-O stretch), 1020, 941, 839, 775; **^{1}H NMR** (400 MHz, $CDCl_3$) δ_H 6.94 (d, *J* = 2.0 Hz, 1H, H7), 6.59 (d, *J* = 2.0 Hz, 1H, H5), 5.16 (q, *J* = 7.5 Hz, 1H, CH), 4.41 (d, *J* = 16.5 Hz, 1H, H3), 4.29 (d, *J* = 16.5 Hz, 1H, H3), 3.853 (s, 3H, H1' or H1''), 3.848 (s, 3H, H1' or H1''), 3.71 (s, 3H, OCH_3), 1.56 (d, *J* = 7.5 Hz, 3H, CH_3); **^{13}C NMR** (125 MHz, $CDCl_3$) δ_C 172.4 (C=O), 168.9 (C1), 161.9 (C6), 155.3 (C4), 134.3 (C7a), 123.0 (C3a), 102.7 (C5), 98.0 (C7), 56.0 (C1' or C1''), 55.6 (C1' or C1''), 52.6 (OCH_3), 49.4 (CH), 44.6 (C3), 16.0 (CH_3); **HRMS** (ES^+) *m/z* calculated for $C_{14}H_{17}NO_5Na$ $[M+Na]^+$: 302.0999; found: 302.0997. See SI2 for ^1H and ^{13}C NMR spectra.

31: IR ν_{Max} cm^{-1} (Thin Film) 2947 (C-H stretch), 1738 (C=O stretch), 1678 (C=O stretch), 1603 (C=C stretch), 1452 (C-H bend), 1433 (C-H bend),1327, 1213 (C-O stretch), 1148 (C-O stretch), 1076 (C-O stretch), 835; **^{1}H NMR** (400 MHz, $CDCl_3$) δ_H 6.51 (d, *J* = 1.9 Hz, 1H,

H4), 6.41 (d, *J* = 1.9 Hz, 1H, H6), 5.12 (q, *J* = 7.4 Hz, 1H, CH), 4.45 (d, *J* = 16.6 Hz, 1H, H3), 4.30 (d, *J* = 16.6 Hz, 1H, H3), 3.91 (s, 3H, H1'), 3.85 (s, 3H, H1''), 3.69 (s, 3H, OCH_3), 1.52 (d, *J* = 7.4 Hz, 3H, CH_3); **^{13}C NMR** (175 MHz, $CDCl_3$) δ_C 172.8 (C=O), 167.6 (C1), 164.6 (C5), 158.7 (C7), 146.4 (C3a), 113.1 (C7a), 99.2 (C4), 98.3 (C6), 56.0 (C1' or C1''), 55.9 (C1' or C1''), 52.4 (OCH_3), 48.9 (CH), 46.7 (C3), 15.9 (CH_3); **HRMS** ($ASAP^+$) *m/z* calculated for $C_{14}H_{18}NO_5$ [M+H]: 280.1185; found: 280.1182. See SI2 for ^{1}H and ^{13}C NMR spectra.

(2-(Hydroxymethyl)-3-methoxyphenyl)methan-*d2*-ol (42)

42 was synthesised according to general procedure **C** using 4-methoxyisobenzofuran-1(*3H*)-one (**40**, 1 equiv., 2.00 g, 12.2 mmol), $NaBD_4$ (2 equiv., 1.02 g, 24.4 mmol), $ZnCl_2$ (1 equiv., 1.66 g, 12.2 mmol) and *N*,*N*-dimethylaniline (1 equiv., 1.55 mL, 12.2 mmol). Pure **42** was obtained as a white solid (1.29 g, 7.6 mmol, 62%). The small intensity of the CD_2 carbon on the ^{13}C NMR spectrum revealed a multiplet.

Mp: 92–94 °C; **IR** ν_{Max} cm^{-1} (Thin Film) 3256 (O-H stretch), 1584 (C=C stretch), 1472 (O-H bend), 1261 (C-O stretch), 1186, 1094, 1042 (C-O stretch), 1007, 982, 775; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 7.29-7.26 (m, 1H, H5), 6.97 (d, *J* = 7.5 Hz, 1H, H6), 6.90 (d, *J* = 8.5 Hz, 1H, H4), 4.84 (s, 2H, H1''), 3.86 (s, 3H, OCH_3); **^{13}C NMR** (125 MHz, $CDCl_3$) δ_C 158.0 (C3), 141.1 (C1), 129.4 (C5), 127.8 (C2), 122.0 (C6), 111.0 (C4), 63.8-63.2 (m, C1'), 56.4 (C1''), 55.9 (OCH_3); **HRMS** (ES^+) *m/z* calculated for $C_9H_{10}O_3{}^2H_2Na$ $[M+Na]^+$: 193.0804; found: 193.0803. See SI2 for ^{1}H and ^{13}C NMR spectra.

2-(Formyl-*d*)-6-methoxybenzaldehyde (35)

35 was synthesised according to general procedure **A** using (2-(hydroxymethyl)-3-methoxyphenyl)methan-*d2*-ol (**42**, 1.0 equiv., 200 mg, 1.18 mmol), $(COCl)_2$ (2.6 equiv., 0.26 mL, 3.06 mmol), DMSO (5.2 equiv., 0.44 mL, 6.14 mmol). Additional DMSO (0.3 mL) was

required to help **42** solubilisation. The reaction was stirred at rt for 17 h. Pure **35** was obtained as a yellow solid (130 mg, 0.79 mmol, 67%).

Mp: 78–81 °C; **IR** ν_{Max} cm^{-1} (Thin Film) 2928 (C-H stretch), 1674 (C=O stretch), 1582 (C=C stretch), 1472 (C-H bend), 1271 (C-O stretch), 1090, 1028 (C-O stretch), 795; **^{1}H NMR** (500 MHz, $CDCl_3$) δ_H 10.64 (s, 1H, CHO), 7.66-7.63 (m, 1H, H5), 7.45 (d, *J* = 7.5 Hz, 1H, H6), 7.24 (d, *J* = 8.5 Hz, 1H, H4), 3.98 (s, 3H, OCH_3); **^{13}C NMR** (125 MHz, $CDCl_3$) δ_C 192.6 (t, *J* = 28.8 Hz, CDO), 191.4 (CHO), 162.0 (C3), 138.6 (t, *J* = 3.7 Hz, C1), 135.4 (C5), 125.1 (C2), 120.3 (C6), 116.3 (C4), 56.4 (OCH_3); **HRMS** (ES^+) *m/z* calculated for $C_9H_7O_3{}^2HNa$ $[M+Na]^+$: 188.0428; found: 188.0429. See SI2 for ^{1}H and ^{13}C NMR spectra.

2-(4-Methoxy-1-oxoisoindolin-2-yl)propanoic acid (18a) with 2-(7-methoxy-1-oxoisoindolin-2-yl-3-*d*)propanoic acid (19a')

A mixture of **18a** and **19a'** was synthesised according to general procedure **B** using 2-(formyl-*d*)-6-methoxybenzaldehyde (**35**, 1.0 equiv., 30 mg, 0.18 mmol) and alanine (**16**, 1.2 equiv., 19 mg, 0.22 mmol). A pure regioisomeric mixture of **18a** and **19a'** was obtained as a light yellow solid (26 mg, 0.11 mmol, 61%). The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture of **18a**:**19a'** with a ratio of 1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2920 (O-H stretch), 2359 (C-H stretch), 1734 (C=O stretch), 1624 (C=O stretch), 1605 (C=C stretch), 1558 (C=C stretch), 1497 (C-H bend), 1489 (C-H bend), 1456 (O-H bend), 1273 (C-O stretch), 1192, 1061 (C-O stretch); **^{1}H NMR** (700 MHz, CD_3OD) δ_H 7.54 (t, *J* = 8.0 Hz, 1H, H5 in **19a'**), 7.47 (t, *J* = 8.0 Hz, 1H, H6 in **18a**), 7.36 (d, *J* = 8.0 Hz, 1H, H7 in **18a**), 7.18 (d, *J* = 8.0 Hz, 1H, H5 in **18a**), 7.10 (d, *J* = 8.0 Hz, 1H, H4 in **19a'**), 7.01 (d, *J* = 8.0 Hz, 1H, H6 in **19a'**), 4.98 (q, *J* = 7.5 Hz, 2H, CH in **18a**, CH in **19a'**), 4.54-4.45 (m, 3H, H3 in **18a**, H3 in **19a'**), 3.93 (s, 3H, OCH_3 in **18a**), 3.91 (s, 3H, OCH_3 in **19a'**), 1.61 (d, *J* = 7.5 Hz, 3H, CH_3 in **18a**), 1.58 (d, *J* = 7.5 Hz, 3H, CH_3 in **19a'**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 175.2 (COOH in **19a'**), 175.0 (COOH in **18a**), 171.0 (C1

in **18a**), 170.0 (C1 in **19a'**), 158.8 (C7 in **19a'**), 156.2 (C4 in **18a**), 146.3 (C3a in **19a'**), 135.0 (C5 in **19a'**), 134.5 (C7a in **18a**), 131.4 (C3a in **18a**), 131.1 (C6 in **18a**), 120.1 (C7a in **19a'**), 116.3 (C4 in **19a'**), 116.2 (C7 in **18a**), 114.5 (C5 in **18a**), 111.2 (C6 in **19a'**), 56.1 (OCH_3 in **18a**), 56.0 (OCH_3 in **19a'**), 51.2 (CH in **18a**), 50.8 (CH in **19a'**), 47.7 (t, *J* = 21.6 Hz, C3 in **19a'**), 46.3 (C3 in **18a**), 16.0 (CH_3 in **18a**), 15.9 (CH_3 in **19a'**); **HRMS** (ES^-) *m/z* calculated for $C_{12}H_{11}{}^2H_1NO_4$ $[M-H]^-$: 235.0835; found: 235.0829 for **19a'**; *m/z* calculated for $C_{12}H_{12}NO_4$ $[M-H]^-$: 234.0772; found: 234.0773 for **18a**. See SI2 for ^1H and ^{13}C NMR spectra.

(2-(Hydroxymethyl)-6-methoxyphenyl)methan-*d2*-ol (43)

43 was synthesised according to general procedure **C** using 7-methoxyisobenzofuran-1(*3H*)-one (**41**, 1 equiv., 145 mg, 0.88 mmol), $NaBD_4$ (2 equiv., 74 mg, 1.77 mmol), $ZnCl_2$ (1 equiv., 120 mg, 0.88 mmol) and *N*,*N*-dimethylaniline (1 equiv., 0.11 mL, 0.88 mmol). Additional THF (1 mL) was required to help **41** solubilisation. Pure **43** was obtained as a white gum (73 mg, 0.43 mmol, 49%). The small intensity of the CD_2 carbon on the ^{13}C NMR spectrum revealed a multiplet.

IR ν_{Max} cm^{-1} (Thin Film) 3248 (O-H stretch), 1584 (C=C stretch), 1470 (O-H bend), 1449 (C-H bend), 1439 (C-H bend), 1260 (C-O stretch), 1086, 1038 (C-O stretch), 959, 783; **^{1}H NMR** (500 MHz, CD_3OD) δ_H 7.27 (t, *J* = 8.0 Hz, 1H, H5), 7.02 (d, *J* = 8.0 Hz, 1H, H6), 6.94 (d, *J* = 8.0 Hz, 1H, H4), 4.72 (s, 2H, H1'), 3.83 (s, 3H, CH_3); **^{13}C NMR** (125 MHz, CD_3OD) δ_C 159.3 (C3), 142.7 (C1), 130.1 (C5), 127.8 (C2), 121.8 (C6), 111.3 (C4), 63.2 (C1'), 56.1 (CH_3), 55.5-54.8 (m, C1''); **HRMS** (ES^+) *m/z* calculated for $C_9H_{10}O_3{}^2H_2Na$ $[M+Na]^+$: 193.0804; found: 193.0801. See SI2 for ^1H and ^{13}C NMR spectra.

2-(Formyl-*d*)-3-methoxybenzaldehyde (39)

39 was synthesised according to general procedure **A** using (2-(hydroxymethyl)-6-methoxyphenyl)methan-*d2*-ol (**45**, 1.0 equiv., 50 mg, 0.29 mmol), $(COCl)_2$ (2.6 equiv., 0.07 mL, 0.76 mmol), DMSO (5.2 equiv., 0.11 mL, 1.51 mmol). Additional DMSO (0.1 mL) was

required to help **43** solubilisation. The reaction was stirred at rt for 15 h. Pure **39** was obtained as a light yellow gum (42 mg, 0.25 mmol, 86%).

IR ν_{Max} cm^{-1} (Thin Film) 2924 (C-H stretch), 1697 (C=O stretch), 1657 (C=O stretch), 1607 (C=C stretch), 1584 (C=C stretch), 1487 (C-H bend), 1472 (C-H bend), 1437 (C-H bend), 1269 (C-O stretch), 1242 (C-O stretch), 1067, 966, 916, 868, 768; **^{1}H NMR** (700 MHz, $CDCl_3$) δ_H 10.42 (s, 1H, CHO), 7.62 (t, *J* = 8.0 Hz, 1H, H5), 7.42 (d, *J* = 8.0 Hz, 1H, H6), 7.22 (d, *J* = 8.0 Hz, 1H, H4), 3.95 (s, 3H, OCH_3); **^{13}C NMR** (175 MHz, $CDCl_3$) δ_C 192.9 (CHO), 191.4-190.9 (m*, CDO), 162.0 (C3), 138.7 (C1), 135.4 (C5), 124.9 (C2), 120.2 (C6), 116.3 (C4), 56.4 (OCH_3).The rapid decomposition of **41** prevented mass spectrometric analysis. *Multiplet signal consistent with potential traces of undeuterated **17a**. See SI2 for ^{1}H and ^{13}C NMR spectra.

2-(4-Methoxy-1-oxoisoindolin-2-yl-3-*d*)propanoic acid (18a') with 2-(7-methoxy-1-oxoisoindolin-2-yl)propanoic acid (19a) and with traces of 2-(4-methoxy-1-oxoisoindolin-2-yl)propanoic acid (18a)

A mixture of **18a'** and **19a** with traces of **18a** was synthesised according to general procedure **B** using 2-(formyl-*d*)-3-methoxybenzaldehyde (**39**, 1.0 equiv., 42 mg, 0.25 mmol) and alanine (**16**, 1.2 equiv., 27 mg, 0.31 mmol). A pure regioisomeric mixture of **18a'** and **19a** with traces of **18a** was obtained as a brown solid (46 mg, 0.20 mmol, 80%). The ^{1}H and ^{13}C NMR signals integration are given assuming a mixture **18a**:**18a'**:**19a** with a ratio of 1:1:1 for clarity.

IR ν_{Max} cm^{-1} (Thin Film) 2922 (O-H bend), 2359 (C-H stretch), 1734 (C=O stretch), 1636 (C=O stretch), 1603 (C=C stretch), 1491 (C-H bend), 1418 (O-H bend), 1269 (C-O stretch), 1065 (C-O stretch), 953; **^{1}H NMR** (700 MHz, CD_3OD) δ_H 7.51 (t, *J* = 8.0 Hz, 1H, H5 in **19a**), 7.44 (t, *J* = 8.0 Hz, 2H, H6 in **18a**, H6 in **18a'**), 7.34 (d, *J* = 8.0 Hz, 2H, H7 in **18a**, H7 in **18a'**), 7.15 (d, *J* = 8.0 Hz, 2H, H5 in **18a**, H5 in **18a'**), 7.08 (d, *J* = 8.0 Hz, 1H, H4 in **19a**),

6.98 (d, *J* = 8.0 Hz, 1H, H6 in **19a**), 4.96-4.89 (m, 3H, CH in **18a**, CH in **18a'**, CH in **19a**), 4.56-4.41 (m, 5H, H3 in **18a**, H3 in **18a'**, H3 in **19a**), 3.90 (s, 6H, OCH_3 in **18a**, OCH_3 in **18a'**), 3.89 (s, 3H, OCH_3 in **19a**), 1.58 (d, *J* = 7.5 Hz, 6H, CH_3 in **18a**, CH_3 in **18a'**), 1.55 (d, *J* = 7.5 Hz, 3H, CH_3 in **19a**); **^{13}C NMR** (175 MHz, CD_3OD) δ_C 176.1 (3C, COOH in **18a**, COOH in **18a'**, COOH in **19a**), 171.0 (2C, C1 in **18a,** C1 in **18a'**), 169.9 (C1 in **19a**), 158.6 (C7 in **19a**), 156.1 (2C, C4 in **18a**, C4 in **18a'**), 146.3 (C3a in **19a**), 134.8 (C5 in **19a**), 134.7 (2C, C7a in **18a**, C7a in **18a'**), 131.4 (2C, C3a in **18a**, C3a in **18a'**), 131.0 (2C, C6 in **18a**, C6 in **18a**'), 120.2 (C7a in **19a**), 116.23 (C4 in **19a**), 116.15 (2C, C7 in **18a**, C7 in **18a'**), 114.4 (2C, C5 in **18a**, C5 in **18a'**), 111.1 (C6 in **19a**), 56.1 (2C, OCH_3 in **18a**, OCH_3 in **18a'**), 56.0 (OCH_3 in **19a**), 51.8 (2C, CH in **18a**, 1C, CH in **18a'**), 51.3 (CH in **19a**), 48.0 (C3 in **19a**), 46.4 (C3 in **18a**), 46.1 (t, *J* = 21.5 Hz, C3 in **18a'**), 16.2 (2C, CH_3 in **18a**, CH_3 in **18a'**), 16.1 (CH_3 in **19a**); **HRMS** (ES^-) *m/z* calculated for $C_{12}H_{11}{}^2H_1NO_4$ $[M-H]^-$: 235.0835; found: 235.0834 for **18a'**; *m/z* calculated for $C_{12}H_{12}NO_4$ $[M-H]^-$: 234.0772; found: 234.0775 for **18a** and **19a**. See SI2 for ^{1}H and ^{13}C NMR spectra.

Acknowledgments

Funding: We would like to thank the EPSRC DTG at the University of St Andrews for funding. We also acknowledge the EPSRC UK Mass Spectrometry Facility at Swansea University for mass spectrometric analysis.

Conflicts of interest: None.

References and notes

1. (a) Hussein Z, Mulford DJ, Bopp BA, Granneman GR. Br J Clin Pharmac. 1993; 36: 357-361. (b) Huskisson EC, Scott J. Rheumatol Rehabil. 1979; 18: 49-52. (c) Sagarad SV, Kerure SB, Kumar SC, Mr R. J Clin Diagn Res. 2013; 7: 687-690.

2. (a) Das S, Addis D, Knöpke LR, Bentrup U, Junge K, Brückner A, Beller M. Angew Chem Int Ed. 2011; 50: 9180-9184. (b) Luzzio FA, Mayorov AV, Ng SSW, Kruger EA, Figg WD. J Med Chem. 2003; 46: 3793-3799.

3. (a) Lee HJ, Lim SJ, Oh SJ, Moon DH, Kim DJ, Tae J, Yoo KH. Bioorg Med Chem Lett. 2008; 18: 1628-1631. (b) De Lucca GV, Shi Q, Liu Q, Batt DG, Beaudoin Bertrand M, Rampulla R, Mathur A, Discenza L, D'Arienzo C, Dai J, Obermeier M, Vickery R, Zhang Y, Yang Z, Marathe P, Tebben AJ, Muckelbauer JK, Chang CJ, Zhang H, Gillooly K, Taylor T, Pattoli MA, Skala S, Kukral DW, McIntyre KW, Salter-Cid L, Fura A, Burke JR, Barrish JC, Carter PH, Tino JA. J Med Chem. 2016; 59: 7915-7935.

4. (a) Ogiwara Y, Uchiyama T, Sakai N. Angew Chem Int Ed. 2016; 55: 1864-1867. (b) Shi L, Hu L, Wang J, Cao X, Gu H. Org Lett. 2012; 14: 1876-1879.

5. (a) Wan J, Wu B, Pan Y. Tetrahedron. 2007; 63: 9338-9344. (b) DoMinh T, Johnson AL, Jones JE, Senise PPJr. J Org Chem. 1977; 42: 4217-4221. (c) Luo W, Yu Q-S, Salcedo I, Holloway HW, Lahiri DK, Brossi A, Tweedie D, Greig NH. Bioorg Med Chem. 2011; 19: 3965-3972. (d) Hussein SH, Ahmed BA, Al-Sharook MM, Al-Rawi JMA. Asian J Chem. 1991; 3: 30-37. (e) Aubert T, Farnier M, Guilard R. Can. J. Chem. 1990; 68: 842-851. (f) Aubert T, Farnier M, Hanquet B, Guilard R. Synth. Commun. 1987; 17: 1831-1837 – during the review process it was identified that the [1,3]- hydride/deuteride shift proposed in references 5(a), 5(c), 5(e) and 5(f) is only symmetry allowed if the hydride transfer is antarafacial (as 4π electrons are involved). This is not sterically feasible in this system and we thank the Reviewer for pointing this out.

6. (a) Allin SM, Hodkinson CC, Taj N. Synlett. 1996: 781-782. (b) Nefkens GHL, Zwanenburg B. Tetrahedron. 1985; 41: 6063-6066. (c) Alajarín M, Sánchez-Andrada P, López-Leonardo C, Álvarez Á. J Org Chem. 2005; 70: 7617-7623.
7. Guertin KR. Pat WO 0248106. 2002.
8. Jolidon S, Rodriguez-Sarmiento RM, Thomas AW, Wyler R. Pat WO 2004014856. 2004.
9. (a) Paget S, Hlasta D. Pat WO 0142242. 2001. (b) Yu W, Tong L, Chen L, Kozlowski JA, Lavey BJ, Shih N-Y, Madison VS, Zhou G, Orth P, Guo Z, Wong MKC, Yang D-Y, Kim SH, Shankar BB, Siddiqui MA, Rosner KE, Dai C, Popovici-Muller J, Girijavallabhan VM, Li D, Rizvi R, Micula AM, Feltz R. Pat US 20070219218. 2007. (c) Taylor EC, Jennings LD, Mao Z, Hu B, Jun J-G, Zhou P. J Org Chem. 1997; 62: 5392-5403. (d) Tung CL, Wong CTT, Fung EYM, Li X. Org Lett. 2016; 18: 2600-2603.
10. (a) Bunce RA, Harrison T, Nammalwar B. Heterocycl Commun. 2012; 18: 123-126. (b) Lima HM, Sivappa R, Yousufuddin M, Lovely CJ. J Org Chem. 2014; 79: 2481-2490.
11. (a) Wolfenden R, Jencks WP. J Am Chem Soc. 1961; 83: 2763-2768. (b) Collett CJ, Massey RS, Taylor JE, Maguire OR, O'Donoghue AC, Smith AD. Angew Chem Int Ed. 2015; 54: 6887-6892. (c) Kool ET, Park D-H, Crisalli P. J Am Chem Soc. 2013; 135: 17663-17666.
12. Yamakawa T, Masaki M, Nohira H. Bull Chem Soc Jpn. 1991; 64: 2730-2734.
13. Wagner J, Van Eis M, Von Matt P, Evenou J-P, Schuler W. Pat WO 2007006533. 2007.
14. Sánchez-Larios E, Holmes JM, Daschner CL, Gravel M. Org Lett. 2010; 12: 5772-5775.

15. Dąbrowski M, Kubicka J, Luliński S, Serwatowski J. Tetrahedron. 2005; 61: 6590-6595.

16. Sivasankaran R, Zimmermann K. Pat WO 2008008821. 2008.

17. Farooq O. Synthesis. 1994; 1994: 1035-1036.

Supplementary Material

Additional results, discussion, synthesis and experimental procedures are provided in SI1. ^{1}H and ^{13}C NMR spectra are given in SI2.