Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron Letters 49 (2008) 2795–2798
Tetrahedron Letters

The facile synthesis of a series of tryptophan derivatives

Georg Blaser, John M. Sanderson *, Andrei S. Batsanov, Judith A. K. Howard

Durham University, Department of Chemistry, University Science Laboratories, South Road, Durham, DH1 3LE, UK

Received 11 January 2008; revised 4 February 2008; accepted 22 February 2008
Available online 4 March 2008

Abstract

This study reports a facile method for the synthesis of a variety of 5- and 6-substituted tryptophan derivatives that are difficult to prepare using alternative enzymatic approaches. Acylation of an activated amino acid, derived from serine in situ, is coupled with an enzymatic resolution step to furnish enantiopure analogues bearing a range of electron withdrawing and releasing substituents. Isolation of a dehydroalanine derivative as a by-product from some reactions provides some insights into the likely mechanism of the reaction.

Keywords: Tryptophan; Acylase; L-Serine; Indole; Unnatural amino acids

As a building block of proteins and nonribosomal peptides, tryptophan plays important roles in protein function, for which the amphiphilic and electrostatic properties of the indole group are of key importance. A number of processes, such as the anchoring of membrane proteins in cell membranes and host–guest binding interactions, are dependent on combinations of dipolar effects and cation–π, hydrogen bonding and aromatic interactions.[1] Due to their luminescence properties, tryptophan derivatives are commonly used as probes in these systems for the determination of protein environment and conformational dynamics through the measurement of inter- and intramolecular distances in fluorescence resonance energy transfer (FRET) experiments and dielectric-dependent changes in emission spectra.[2] Tryptophan is a substrate for the biosyntheses of numerous 2° metabolites, many of which exhibit a range of substituents at various positions on the indole ring. Halogenated intermediates in particular, frequently have significantly modified bioactivity and bioavailability[3] and are desirable precursors for in vitro syntheses, as well as potential substrates for further selective functionalisation through reactions such as the Suzuki and Sonogashira couplings. Methods for the efficient synthesis of tryptophan analogues with a wide range of functionality and electronic properties are therefore highly desirable.

Tryptophan can be synthesised using a variety of chemical and enzymatic methods; amongst the latter, tryptophan synthase has received much attention.[4] This enzyme uses indole-3-glycerol phosphate and L-serine as substrates to produce L-tryptophan in two sequential reactions: indole must first enter the α-subunit and pass through a tunnel 2.5 nm in length before it reaches the active site in the β-subunit.[5] The accessibility of the enzyme active site restricts the range of tryptophan derivatives that may be formed, with substitutions on the 4- or 7-positions normally being poor substrates.[4] Some substitutions at the 5- and 6-positions are accepted, but large groups such as iodo or nitro are usually poorly tolerated.

Chemical synthesis can be achieved by classical electrophilic substitution at the C-3 indole carbon. Substitution at this position is generally favoured on resonance and electrostatic grounds. If a serine precursor is used to generate the electrophile, under the appropriate conditions the resulting product will be a tryptophan analogue. This chemical approach has the advantage that there are fewer restrictions in terms of the indole derivates that may be used, giving access to a wide range of products. In this Letter, we describe the synthesis of a range of 5- and

* Corresponding author. Tel.: +44 191 334 2107; fax: +44 191 384 4737.
E-mail address: j.m.sanderson@durham.ac.uk (J. M. Sanderson).

0040-4039/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2008.02.120

6-substituted tryptophan derivatives by electrophilic substitution and their subsequent resolution using enzymatic methods.

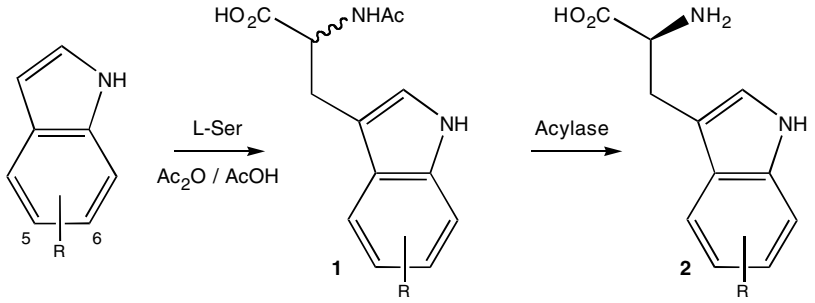
A series of racemic N^{α}-acetyl tryptophan derivatives (**1**) was produced by the reaction of the corresponding indoles with L-serine in a mixture of acetic acid and acetic anhydride.[6] The products were isolated in good yield (Table 1) and cover a broad spectrum of electron density on the indole ring, from electron rich (Table 1, entries e and k; R = OMe) to electron poor (Table 1, entries f and l; R = NO_2), demonstrating the applicability of this reaction for the preparation of a broad range of analogues.

Enzymatic resolution of the racemates using Acylase 'Amano' made both the free L-amino acid (**2**) and the corresponding N^{α}-acetyl D-enantiomer readily available.[7] The L-tryptophan derivatives were used for further functionalisation, either through the introduction of an Fmoc protecting group, or conversion to the corresponding N^{α}-acetyl *N*-ethyl amides (Scheme 1).

Confirmation of the substitution position and stereochemistry was obtained through the collection of X-ray crystallographic data for a number of compounds (Fig. 1) and comparison with literature data for the optical activities of the free amino acids (see Supplementary data). All amino acids were obtained with enantiomeric excesses $\geqslant$90%. Isolation of the N^{α}-acetyl D-enantiomer provides a means for preparing similar analogues of the naturally less abundant stereoisomer. In addition, this enantiomer could potentially be racemised and recycled to produce more L-enantiomer following enzymatic resolution, although this was not performed in our case.

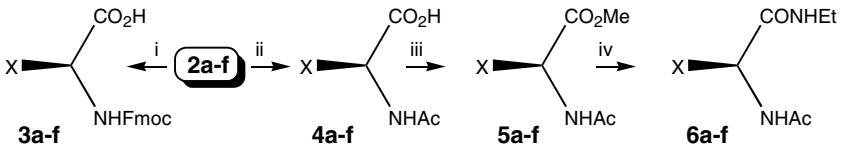
Compound **3e** crystallised as a 1:1 solvate with deuterochloroform (the presence of which allowed the direct determination of the absolute configuration) linked to the host molecule via a C–D···O hydrogen bond (Fig. 1). Otherwise, hydrogen bonding in this structure is surprisingly inefficient, notwithstanding the abundance of potential acceptor sites (oxygen atoms). The N2–H group forms an

Table 1
Data for the reaction of monosubstituted indoles with L-serine and isolation of free L-tryptophan analogues following resolution



#	R	Pos.	[Ser][a]	Yield **1** (%)	Yield **2**[b] (%)	ee[c] (%)
a	F	C-5	2.0	82 (**1a**)	97 (**2a**)	100
b	Cl	C-5	2.0	86 (**1b**)	80 (**2b**)	100
c	Br	C-5	2.0	77 (**1c**)	94 (**2c**)	100
d	I	C-5	1.5	74 (**1d**)	79 (**2d**)	—
e	OMe	C-5	1.0	80 (**1e**)	97 (**2e**)	100
f	NO_2	C-5	2.5	44 (**1f**)	68 (**2f**)	92
g	F	C-6	2.0	98 (**1g**)	65 (**2g**)	91
h	Cl	C-6	2.0	86 (**1h**)	52 (**2h**)	91
i	Br	C-6	2.0	73 (**1i**)	97 (**2i**)	93
j	Me	C-6	1.5	73 (**1j**)	92 (**2j**)	93
k	OMe	C-6	1.0	77 (**1k**)	63 (**2k**)	91
l	NO_2	C-6	2.5	58 (**1l**)	61 (**2l**)	91

[a] Molar equivalents of L-serine with respect to indole.
[b] The yield calculated based on the maximum theoretical recovery from the resolution of a racemic mixture.
[c] Enantiomeric excesses are estimated from optical rotation measurements and literature data and have an error of ±5%. No suitable literature data are available for **2d**.



Scheme 1. Preparation of tryptophan derivatives. Reagents and conditions: (i) Fmoc-OSu, 1,4-dioxane, Na_2CO_3; (ii) Ac_2O, Et_3N; (iii) MeOH, concd H_2SO_4 (cat.), Δ; (iv) $EtNH_2$, MeOH. Key: X = 3-indolylmethyl.

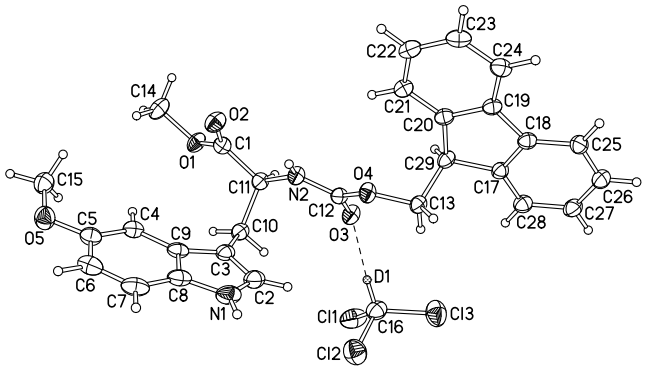


Figure 1. Molecular structure of the methyl ester of **3e**·$CDCl_3$. Thermal ellipsoids are drawn at the 50% probability level (CCDC Ref. 670184).

extremely long intermolecular N–H···O bond (N···O distance of 3.522(2) Å), which can be explained by competition from the intramolecular interaction N2–H···O2. The N1–H group, which has no such hindrance, forms only a weak N–H···π bond with the indole C-8 atom of an adjacent molecule. The planar C14–O1–O2–C1–C11 ester group (i) and the N2–C12–O3–O4 moiety (ii) form an interplanar angle of 15.8° between them and are inclined by 52.4° and 45.8° to the indole plane, which forms a 70.3° angle with the fluorene plane.

Compounds **1b** (Fig. 2) and **6b** both crystallise in centrosymmetric space groups (i.e., as racemates). Bond distances are unexceptional (see Supplementary data, Table S3), although the C–Cl distances (1.754(2) and 1.757(4) Å, respectively) fall within the longest decile of C(sp^2)–Cl bonds present in the Cambridge Structural Database (cf. the mean of 1.742(7) Å and the median value of 1.741 Å for low-temperature structures).[8] In either structure all the NH and OH groups donate intermolecular hydrogen bonds to terminal (carbonyl) O atoms, giving rise to a 3-dimensional network, in which the chlorine atom plays no role. In other respects, the tryptophan conformations are similar to those of the parent compounds.[9,10]

The relative success of the electrophilic substitution depends on the electron density distribution on the indole ring. The less electron rich nitro-substituted indoles were

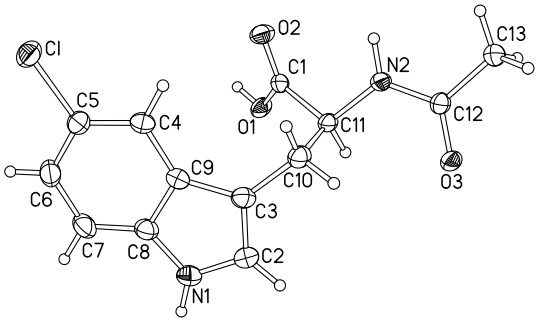


Figure 2. Molecular structure of **1b** (L molecule in the racemic crystal; CCDC Ref. 670183). Thermal ellipsoids are drawn at the 50% probability level.

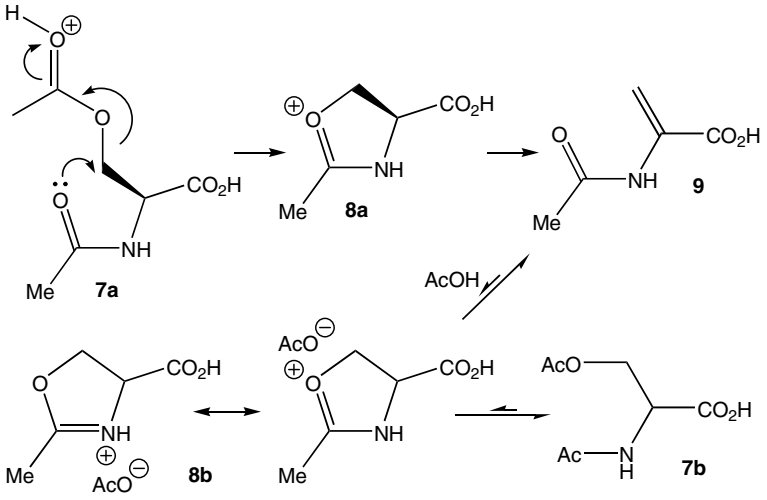
of lower reactivity and gave moderate yields. More electron rich indoles were found to be highly reactive towards electrophilic attack, leading to double-substitution at positions C-2 and C-3. Reaction yields were therefore maximised by adjusting the stoichiometry accordingly, either by increasing or decreasing the amount of indole for electron deficient and electron rich compounds, respectively. The highest yields were observed for substituents with intermediate electron releasing ability (Table 1, entries a, b, g and h; R = Cl, F). The total conversion of indole to product could be further improved by recycling unreacted indole recovered from the reaction.

A drawback of the substitution reaction is the racemisation of the amino acid stereocentre. As the reaction between indole and L-serine is conducted under acidic conditions, it is probable that the observed racemisation proceeds neither via oxazolone formation, nor via a classical direct enolisation, both of which require the presence of a base. On the other hand, azlactone formation is promoted under the conditions used in this synthesis, which will undoubtedly contribute to the formation of racemates.[11]

It also seems likely that racemisation can occur during the formation of reactive intermediates from serine. Interestingly, *N*-acetyl dehydroalanine (**9**, Scheme 2) was isolated as a by-product in reactions with less reactive indoles.

The above observation is consistent with the known reaction between indole and **9** to produce tryptophan, which occurs under acidic conditions.[12,13] Formation of **9** from either L-serine or D,L-serine by heating with acetic anhydride in acetic acid has been described.[13] However, on repeating this reaction we found the major component of the mixture to be *N*,*O*-diacetyl D,L-serine (**7b**).

A plausible mechanism for the formation of **9** is through acid-promoted loss of acetic acid from the acylated L-serine intermediate **7a** (Scheme 1). This will be assisted by neighbouring group participation from the *N*-acetyl group, leading to the formation of the cationic species **8a** as a transient intermediate, before the loss of a proton to form **9**. These kinds of neighbouring group effects have been observed for serine in the gas phase[14] and are likely to persist to a lesser extent in solution. Under the conditions used, *N*-acetyl dehydroalanine would be susceptible to the addition of acetic acid across the olefinic double bond to form **7b** (the racemic counterpart of **7a**). Protonation of the olefinic bond of **9** as a precursor to addition of acetic acid will form **8b** (the racemic counterpart of **8a**), with neighbouring group participation again helping to stabilise the cationic species. We anticipate that the reaction of indoles with electrophile **8b** is the product forming step in the conversion to tryptophan analogues. Under the experimental conditions used, the steady-state concentration of electrophile **8b** is expected to be extremely low, but nevertheless sufficient for product formation. Direct reaction of *N*-acetyl dehydroalanine (**9**) with indoles in the presence of acetic anhydride and acetic acid tends to give significantly lower yields than the corresponding reaction with serine,[13] which suggests that the equilibrium between **9** and **8b** lies in



Scheme 2. Proposed equilibria established by the reaction of *N*-acetyl L-serine with acetic anhydride in acetic acid.

favour of the former. This would also suggest that a significant amount of product from the reaction arises from reaction with **8a** and is subjected to subsequent racemisation via an azlactone route.

In summary, by using a combined chemical and enzymatic resolution strategy, a range of substituted tryptophan analogues have been prepared with high enantiomeric purity. Both indole reactivity and enzyme activity are not adversely affected by substituents at either the 5- or 6-position of the indole ring.

Acknowledgements

EPSRC and Durham University are gratefully acknowledged for funding.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tetlet.2008.02.120.

References and notes

1. Sanderson, J. M. *Org. Biomol. Chem.* **2005**, *3*, 201–212.
2. Royer, C. A. *Chem. Rev.* **2006**, *106*, 1769–1784.
3. Neidleman, N. L.; Geigert, J. *Biohalogenation: Principles, Basic roles and Applications*; Ellis Horwood Ltd: Chichester, 1986.
4. Goss, R. J. M.; Newill, P. L. A. *Chem. Commun.* **2006**, *47*, 4924–4925.
5. Hyde, C. C.; Ahmed, S. A.; Padlun, E. A.; Miles, E. W.; Davies, D. R. *J. Biol. Chem.* **1988**, *263*, 17587–17871.
6. *Typical procedure*: L-serine (0.2 mmol) was added to a solution of the substituted indole (0.1 mmol) and acetic anhydride (0.9 mmol) in acetic acid (0.24 ml). The mixture was stirred under argon at 73 °C for 2 h. Upon cooling to room temperature, the solution was adjusted to pH 11 by careful addition of 30% aqueous sodium hydroxide and extracted with ether (5 ml). Following the addition of sodium thiosulfate (1 M aq, two drops), the solution was adjusted to pH 7 using concentrated hydrochloric acid and the volume reduced by 50% in vacuo. If precipitation occurred, samples were stored at 4 °C for 24 h to allow further precipitation. Precipitates were filtered and dried in vacuo. The filtrate was further concentrated to half its original volume and any further precipitates were collected. If no precipitation occurred following concentration, the filtrates were acidified to pH 3 with 5% hydrochloric acid, extracted with ethyl acetate (3 × 7 ml) and the combined organic extracts were dried ($MgSO_4$). After filtration, the solvent was removed in vacuo and the residue treated with benzene, before re-evaporation, to yield the crude product. The ethereal solution was washed with 1 N sodium hydroxide solution (2 × 2 ml) and dried ($MgSO_4$). Removal of the solvent under vacuum (CAUTION) allowed the recovery of unreacted indole.
7. *Typical procedure*: N^{α}-acetyl-D,L-tryptophan (**2**, 0.40 mmol) was dissolved in borate buffer solution (50 mM, 5 ml) containing cobalt(II) chloride hexahydrate (0.125 mM) at pH 8.0. A solution of amano acylase (Amano Enzyme Inc., Nagoya, Japan; 100 mg) in borate buffer (50 mM, 10 ml) was added with stirring. The mixture was stirred at 37 °C for 48 h and then quenched by adjusting the pH to 5 with 10% aqueous hydrochloric acid. The mixture was filtered through a Celite pad and the filtrate extracted with ethyl acetate (3 × 25 ml). The combined organic extracts were dried over magnesium sulfate, filtered and concentrated in vacuo to afford the crude N^{α}-acetyl-D-tryptophan. The filtrate was passed through Dowex® 50X2-200 resin (H^+ form) and the resin washed with water (150 ml). The product was eluted using 10% ammonia in methanol (200 ml) and the solvent removed in vacuo to afford the crude L-enantiomer. Products were purified by HPLC using a Hypersil C_{18} column (250 × 4.6 mm, 10 mm particle size) at 25 °C running linear gradients of water/MeCN containing 0.1% TFA.
8. Allen, F. H.; Taylor, R. *Chem. Soc. Rev.* **2004**, *33*, 463–475.
9. Sun, H.; Oldfield, E. *J. Am. Chem. Soc.* **2004**, *126*, 4726–4734. The structure is mistakenly referred to as L-tryptophan derivative, although the space group (*Pc*) includes mirror transformations.
10. Harada, Y.; Iitaka, Y. *Acta Crystallogr., Sect. B* **1977**, *33*, 244–247.
11. Du Vigneaud, V.; Meyer, C. E. *J. Biol. Chem.* **1933**, *99*, 143–151.
12. Synder, H. R.; MacDonald, J. A. *J. Am. Chem. Soc.* **1955**, *77*, 1257–1259.
13. Yokoyama, Y.; Hikawa, H.; Mitsuhashi, M.; Uyama, A.; Hiroki, Y.; Murakami, Y. *Eur. J. Org. Chem.* **2004**, *6*, 1244–1253.
14. Reid, G. E.; Simpson, R. J.; O'Hair, R. A. J. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 1047–1060.