

Novel heterocyclic α -amino acids with sulfur-containing side-chains

Raymond C. F. Jones,^{*a} Lisa J. Crumpling,^a and James N. Iley^b

^aDepartment of Chemistry, Loughborough University, Loughborough, Leics. LE11 3TU, UK

^bDepartment of Chemistry, Open University, Walton Hall, Milton Keynes MK7 6AA, UK

E-mail: r.c.f.jones@lboro.ac.uk

Dedicated to Professor Siegfried Blechert on the occasion of his 65th birthday

DOI: <http://dx.doi.org/10.3998/ark.5550190.0012.408>

Abstract

S-Alkylation of an N-protected cysteine ester with a range of ω -iodoalkyl heterocycles affords 11 novel non-proteinogenic heterocyclic amino acids having a sulfur atom in the backbone-heterocycle linker.

Keywords: Amino acid, heterocycle, cysteine, nucleobase

Introduction

Amongst the proteinogenic amino acids, there are only two examples of side chains carrying a heterocycle, namely histidine **1** and tryptophan **2** (Figure 1). On the other hand heterocyclic chemistry provides an enormous range of potential heterocyclic systems that might be exploited to replace the imidazole or indole moieties. As part of a programme to explore this possibility, we have previously reported on α -amino acids **3** ($n = 0, 2, 3$) carrying heterocycles in their side chains, including pyridine and isoxazole derivatives but in particular the pyrimidine and purine nucleobases tethered to the backbone with varying length carbon chains.^{1,2} These are potential α -PNA monomers,^{3,4} amongst many other possible applications, and are analogues of natural products such as discadenine.⁵ They are obtained by stereoselective conjugate radical addition to an optically active methylene oxazolidinone derived from *S*-methylcysteine,¹ or by stereoselective C-alkylation of an ephedrine-based glycinamide.² In pursuit of further variation in the heterocycle-backbone tether and the backbone-to-heterocycle linkage protocol, and to demonstrate diversity in the heterocycle, we report now the preparation of 11 novel heterocyclic amino acids **4** ($n = 1, 2, 3$) with sulfur-containing tethers, obtained by elaboration of cysteine (Figure 1).^{6,7} The strategy adopted (Scheme 1) was to alkylate the side chain sulfur atom of a cysteine derivative with haloalkyl heterocycles; the heterocycles employed were nucleobases

(uracil, thymine, adenine) and the simpler analogues indole, benzimidazole, benzotriazole and purine. This C–S bond formation is distinct from those employed by others, and extends the portfolio of heterocycles beyond the nucleobases.^{3,6,7} The product amino acids are possible components in novel non-natural amino acid sequences with potential biological properties.

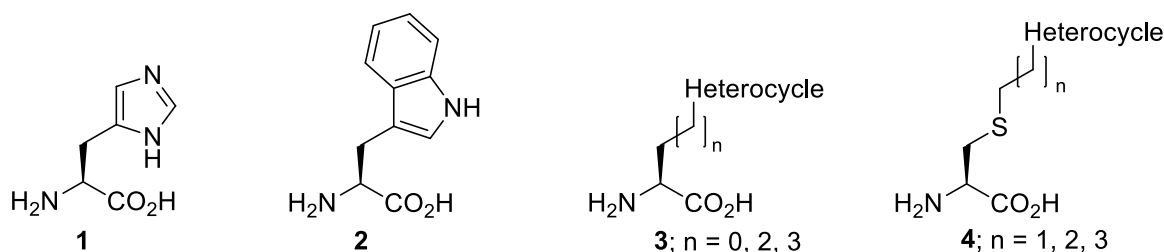
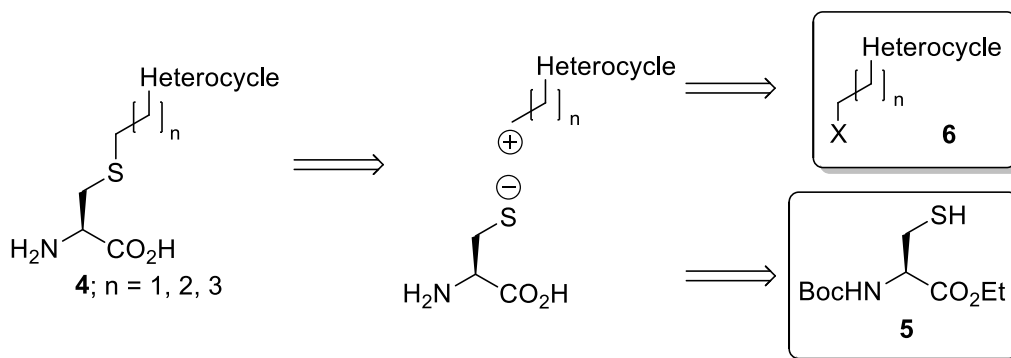


Figure 1. Natural and synthetic heterocyclic amino acids.

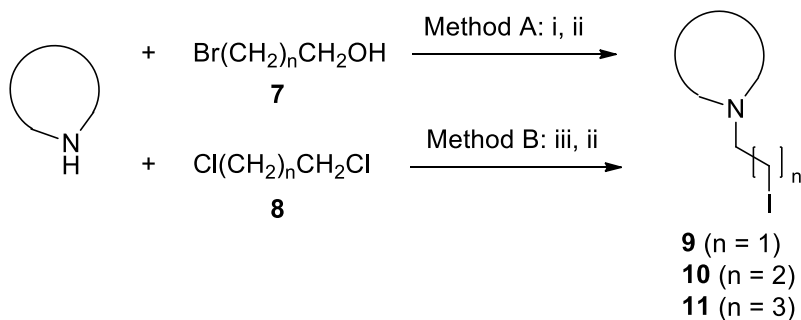


Scheme 1. Synthetic strategy for S-linked heterocyclic amino acids.

Results and Discussion

The scaffold used for the side chain elaboration was *N*-*tert*-butoxycarbonyl-(*R*)-cysteine ethyl ester **5**, easily prepared from commercial (*R*)-cysteine ethyl ester hydrochloride (Boc_2O , *i*- Pr_2NEt , CH_2Cl_2 , 2 h; 98%). Relevant electrophiles, the ω -iodoalkyl heterocycles **6** ($n = 1, 2, 3$), were prepared by two general methods (Scheme 2): (i) Method A, Mitsunobu coupling (*i*- $\text{PrO}_2\text{CN}=\text{NCO}_2i\text{-Pr}$ (DIAD), Ph_3P) of ω -bromoalcohols **7** (2-bromoethanol, 3-bromopropan-1-ol, 4-bromobutan-1-ol) with NH-heterocycles, followed by Finkelstein halide exchange with iodide ion (NaI , Me_2CO reflux);¹ or (ii) Method B, phase-transfer N-alkylation (Bu_4NI , KOH , K_2CO_3) of the heterocycle with an α,ω -dihalide **8** (1,2-dichloroethane, 1,3-dichloropropane, 1,4-dichlorobutane) used as solvent, again followed by halide exchange to generate the *N*-(ω -iodoalkyl) derivative. In this way, a range of *N*-iodoethyl derivatives **9a-h** were prepared from 3-benzoyluracil, 3-benzoylthymine, indole, benzimidazole, benzotriazole, *N*⁶-(*tert*-butoxycarbonyl)adenine and purine as shown in Figure 2 and Table 1. In the case of

benzotriazole, both N-1 and N-2 bromoethyl derivatives were formed using method A, and could be separated by column chromatography, whereas method B produced only the N-1 derivative. In the same manner, *N*-iodopropyl derivatives **10e,f** and *N*-iodobutyl derivatives **11e,f** were also prepared from benzotriazole. Method A provided only the N-1 bromopropyl variant (although this was not taken further), but using method B, both the N-1 and N-2 chloropropyl and chlorobutyl compounds were formed and were separated by column chromatography (Figure 1, Table 1). *N*-(3-Iodopropyl)indole **10c** could not be isolated pure from the corresponding chloroalkyl compound, and the *N*-(4-chlorobutyl) derivative could also not be purified from method B. When *N*-(3-chloropropyl)- or *N*-(3-bromopropyl)benzimidazole were subjected to our standard Finkelstein conditions with sodium iodide, the corresponding iodoalkylbenzimidazole **10d** was not observed but a very polar material was isolated, provisionally assigned on the basis of MS and ¹H NMR spectroscopic evidence as the cyclic double quaternary salt **12a** (Figure 3).^{8,9} Likewise, when formation of *N*-(4-chlorobutyl)benzimidazole was attempted according to method B, a polar product similarly assigned as the cyclic double salt **12b** was obtained.



Scheme 2. Haloalkylation of NH-heterocycles. *Reagents:* i, DIAD, Ph₃P; ii, NaI, Me₂CO reflux; iii, Bu₄NI, KOH, K₂CO₃.

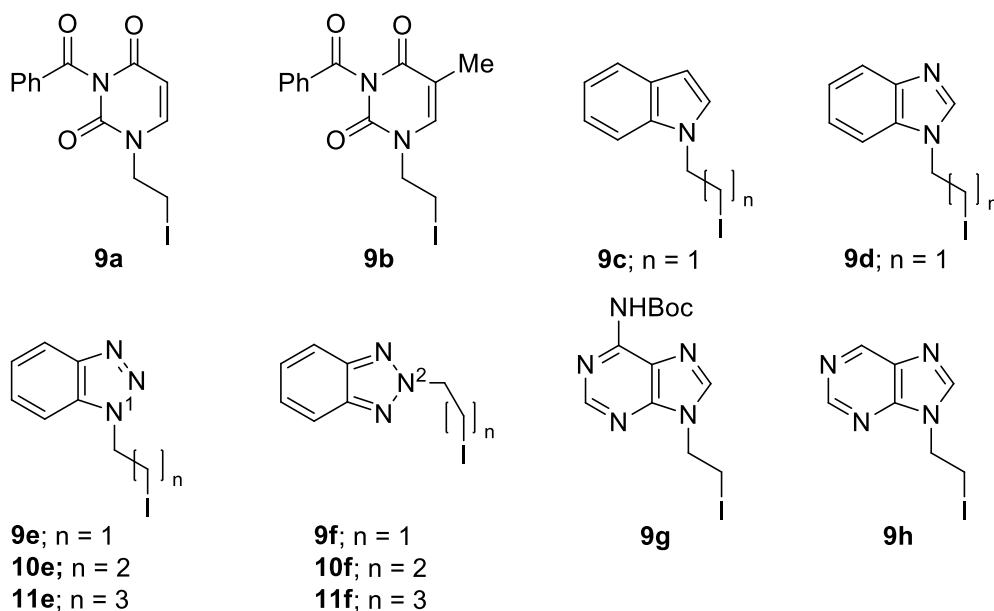


Figure 2. ω -Iodoalkyl heterocycles **9-11** (**9**: $n = 1$; **10**: $n = 2$; **11**: $n = 3$) prepared according to Scheme 2.

Table 1. Preparation of ω -iodoalkyl heterocycles **9-11**

Product	Method	Yield % (step 1)	Yield % (halide exchange)
9a	A	95%	36%
9b	A	56%	54%
9c	B	81%	64%
9d	A	98%	75%
	A*	80%	72%
	B	54%	72%
9e	B	66%	94%
	A	27%	95%
9f	A	56%	97%
9g	A	37%	51%
	B	59%	64%
9h	A	73%	79%
	B	59%	64%
10c	B	61%	†
10d	A	78%	‡
	B	62%	‡
10e	A	78%	–
	B	52%	77%
10f	B	24%	65%
11d	B	§	
11e	B	24%	70%
11f	B	21%	70%

*using 2-chloroethanol rather than 2-bromoethanol; †: impure product obtained; ‡: product assigned as **12a**; §: product assigned as **12b**.

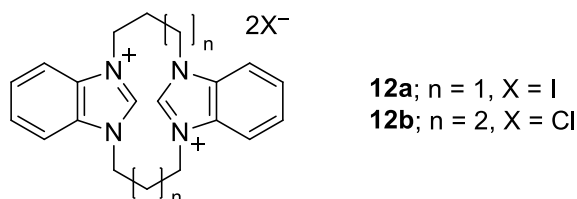
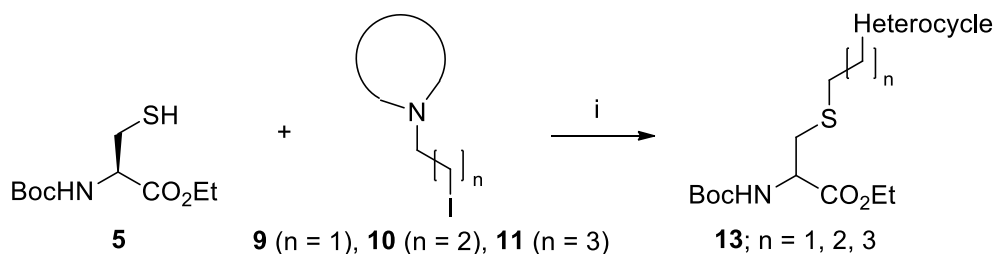


Figure 3. Products of benzimidazole haloalkylation.

With these electrophiles **9-11** in hand, we proceeded to successfully employ them in S-alkylation of the protected cysteine **5** under basic conditions (Scheme 3). After a number of investigations, our preferred protocol was determined to be treatment of N-protected cysteine ester **5** in dry THF under an atmosphere of nitrogen, with sodium hydride (2.5 mol eq) at 0 °C and the mixture allowed to warm to 20 °C over 5-20 min, followed by addition of the iodoalkyl heterocycle (4 mol eq). If the halide was added directly after the base, alkylation yields were lower and elimination product was observed; an optimum delay in each case was determined by experiment. After stirring the alkylation reaction mixture under reflux for an appropriate period, normally 16 h, the novel S-linked protected amino acids were isolated by conventional methods. In this way, 11 non-proteinogenic amino acid derivatives **13a-k** were prepared in the yields shown in Figure 4; compound **13g** was not fully characterised. Purine derivative **9h** did not afford significant alkylation product under our conditions. The optical purity of the new residues remains to be determined.



Scheme 3. S-alkylation of protected cysteine **5**. Reagents: **i**, **5** with NaH, THF, 0 to 20 °C, 5-20 min, then halide **9**, **10** or **11**.

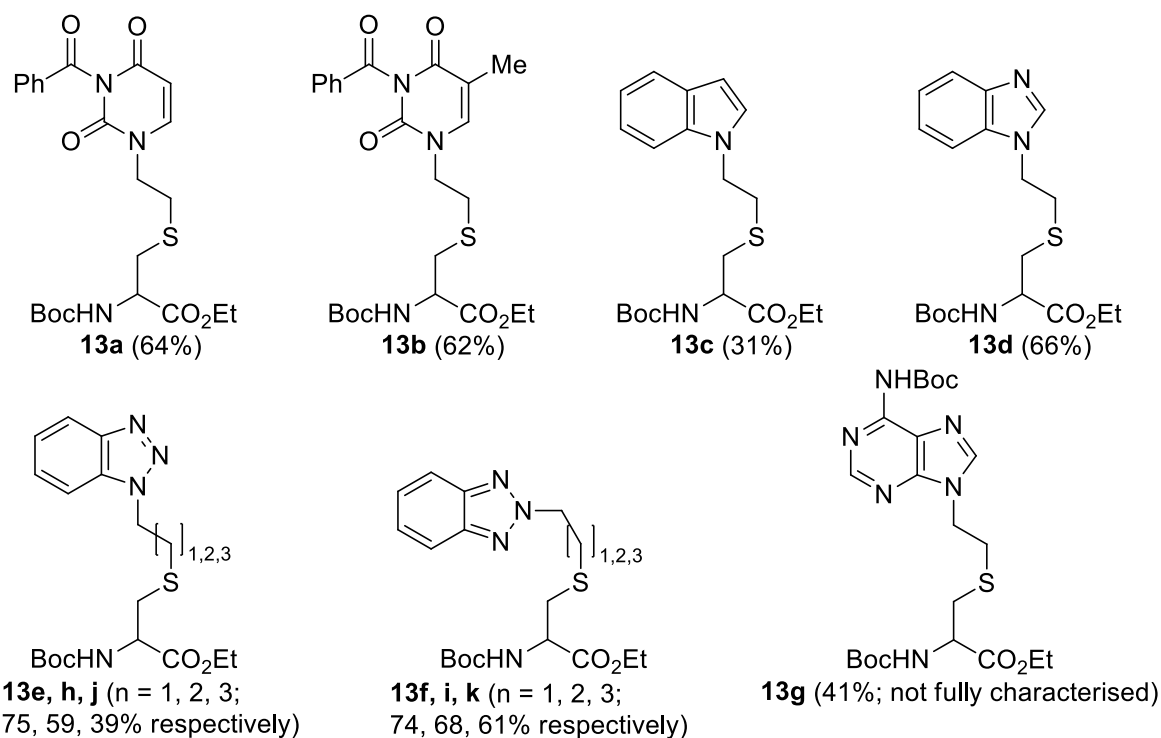
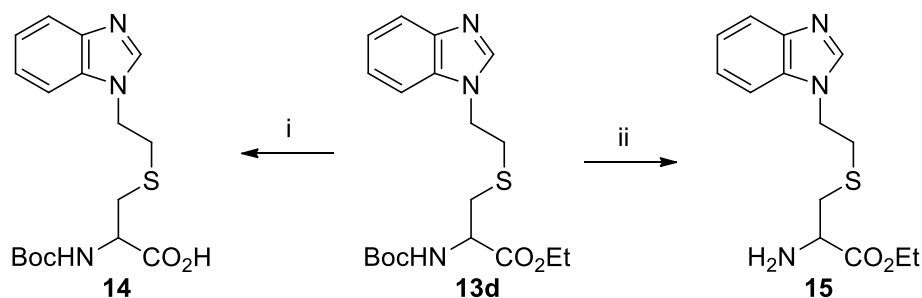


Figure 4. S-Linked heterocyclic amino acids prepared according to Scheme 3.

To demonstrate the potential for application of the new amino acids in organic synthesis, the benzimidazole-carrying derivative **13d** was orthogonally deprotected (Scheme 4) at either its C- or N-terminus. Thus, basic hydrolysis (1M LiOH aq, THF-MeOH, 20 °C, 8 h) afforded the N-protected acid **14** (97%). Alternatively, acid treatment (TFA, CH₂Cl₂, 20 °C, 2 h) and subsequent neutralization (*i*-Pr₂NEt) led to the free amine **15** (86%).



Scheme 4. Deprotection of S-linked amino acid derivative **13d**. Reagents: i, LiOH aq, THF-MeOH, 20 °C, 8 h; ii, TFA, CH₂Cl₂, 20 °C, 2 h, then *i*-Pr₂NEt.

A similar approach using *N*-benzyloxycarbonyl-(*S*)-serine methyl ester to give O-linked heterocyclic amino acids, was unsuccessful, as base treatment led to polymerisation rather than O-alkylation. Using a more hindered *tert*-butyl ester afforded no improvement. Alkylation using the free acid *N*-benzyloxycarbonyl-(*S*)-serine showed possible success that awaits further

study.¹⁰ Investigations with 2,3-diaminopropionic acid and 2,4-diaminobutyric acid were curtailed as a suitable protected amino acid alkylation substrate could not be easily obtained. On the other hand, we have indications that our approach will be successful using a protected homocysteine scaffold.¹¹

Conclusions

We have successfully prepared 11 novel heterocyclic α -amino acids by S-alkylation of an N-protected cysteine ester, and demonstrated the orthogonal deprotection that will enable their application in organic synthesis.

Experimental Section

General. Flash column chromatography was performed using silica gel 60 (40-63 μ , 230-400 mesh, 60 A), and thin layer chromatography (TLC) performed on Merck 0.2 mm silica 60 F₂₅₄ coated UV active aluminium sheets. Melting Points were measured on a Stuart Scientific Bibby SMP3 Melting Point Apparatus and IR spectra were recorded on a Perkin-Elmer Paragon 1000 FTIR spectrometer for neat samples on NaCl discs (CH₂Cl₂ used as a 'transport' solvent where necessary). ¹H NMR spectra were recorded at 400 MHz on a Bruker DPX-400 instrument or at 250 MHz on a Bruker AC-250 instrument; ¹³C NMR spectra were recorded at 100 MHz on the Bruker DPX-400 instrument. Chemical shifts are quoted as δ ppm with TMS as internal standard, and the coupling constants *J* in Hz. Signals are described as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br (broad). Mass spectra were recorded using either EI or FAB ionisation, on a JEOL SX102 spectrometer at Loughborough University or by the ESPRC National Mass Spectrometry Service Centre at Swansea University. Liquid Chromatography–Mass Spectrometry was performed using a Waters 600 Controller instrument, a Waters Symmetry C8 3.5 μ m 4.6 x 50 mm column and a Waters 996 photodiode array detector attached to a Micromass Platform mass spectrometer using electrospray ionization. Elemental Analyses for C, H and N were obtained using a Perkin-Elmer CHN-2400 Elemental Analyzer. 2-Bromoethanol, 2-chloroethanol, 3-bromopropan-1-ol, 4-bromobutan-1-ol, 1,2-dichloroethane, 1,3-dichloropropane, 1,4-dichlorobutane, uracil, thymine, indole, benzimidazole, benzotriazole, adenine, purine and (*R*)-cysteine ethyl ester were purchased from either Aldrich, Lancaster or Avocado and were used without purification unless otherwise stated. 3-Benzoyluracil and 3-benzoylthymine were prepared according to the method of Reese *et al.*¹² Solvents were dried where necessary, as follows: THF was pre-dried in the presence of solid K₂CO₃ then distilled from sodium metal and benzophenone under a positive atmosphere of nitrogen; EtOAc was distilled over calcium chloride; CH₂Cl₂ was distilled over anhydrous calcium hydride; and 1,4-

dioxane was shaken over potassium hydroxide pellets. Characterization data are given for pure compounds.

Method A. General procedure for heterocycle ω -bromoalkylation

NH-Heterocycle (1 mol equiv), triphenylphosphine (1.2 mol equiv) and the ω -bromoalcohol (1.2 mol equiv) were suspended in dry dioxane (50 mL per mmol of heterocycle) at 5 °C, to which was added DIAD (1.2 mol equiv) dropwise over 3 h. The mixture was stirred under an atmosphere of nitrogen at 20 °C to give a clear solution. The solvent was removed under reduced pressure and the residue purified by flash chromatography to give the pure title compound as a yellow oil, which was in some cases recrystallised to give the product.

Method B. General procedure for heterocycle ω -chloroalkylation

Heterocycle (1 mol equiv.) was added to a mixture of α,ω -dichloroalkane, which acted as both the solvent and the reagent, tetrabutylammonium iodide (5% w/w) as the phase-transfer catalyst, KOH (6 mol equiv.) and K₂CO₃ (2.5 mol equiv.). The reaction was stirred at 20 °C for 4 h. The inorganic material was filtered off and the organic portion washed with water (5 mL per mmol of heterocycle), dried (MgSO₄), filtered and the solvent removed under reduced pressure. The mixture was then purified by flash column chromatography to isolated the pure chloroalkyl compound.

General procedure C. Halide exchange to afford ω -iodoalkyl heterocycles (9-11)

N-(ω -Haloalkyl) heterocycle (1 mol equiv) and dry sodium iodide (5 mol equiv) were heated in dry acetone (50 mL per mmol of heterocycle) at reflux overnight in the dark under a nitrogen atmosphere. After cooling, the acetone was removed under reduced pressure and the residue taken up in EtOAc : water (1:1 v/v), the organic layer was separated and the aqueous layer was extracted twice more with EtOAc. The combined organic layers were washed twice with sodium thiosulfate solution (2% w/v), dried (MgSO₄), filtered and the solvent was removed under reduced pressure to yield the pure iodoalkyl compound.

3-Benzoyl-1-(2-bromoethyl)uracil. Prepared according to method A as previously reported, using 3-benzoyluracil (5.00 g, 23.04 mmol), triphenylphosphine (7.24 g, 27.63 mmol) and 2-bromoethanol (1.96 mL, 3.45 g, 27.62 mmol) in dry dioxane (100 mL) at 5 °C with DIAD (5.43 cm³, 5.58 g, 27.61 mmol) to give the title compound as a white solid (7.07 g, 95%), m.p. 177-179 °C (lit.,¹ m.p. 183-184 °C), IR (ν_{\max} , NaCl/cm⁻¹): 2920, 2850, 1747, 1704, 1662, 1437, 1346, 1241. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.69, 4.16 (each 2H, t, J = 5.6 Hz, CH₂CH₂), 5.83 (1H, d, J = 8.0 Hz, 5-CH), 7.33, 7.51 (each 2H, t, J = 8.0 Hz, Ar-H), 7.67 (1H, d, J = 8.0 Hz, Ar-H), 7.93 (1H, d, J = 8.0 Hz, 6-CH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 29.4 (CH₂Br), 51.3 (CH₂N), 101.9 (5-CH), 129.3, 130.5 (2 x Ar-CH), 132.1 (Ar-C), 135.2 (Ar-CH), 144.8 (6-CH), 149.7, 159.9, 168.1 (3 x C=O). MS, m/z = 324/322 (M⁺), 296, 294, 277, 254, 252, 215, 188, 105.

3-Benzoyl-1-(2-iodoethyl)uracil (9a). Prepared according to general procedure C as previously reported, using 3-benzoyl-1-(2-bromoethyl)uracil (5.00 g, 15.48 mmol) and dry NaI (11.71 g,

78.07 mmol) in dry acetone (100 mL) to yield the title compound as a pale yellow solid (2.11 g, 36%), m.p. 185-187 °C (lit.,¹ m.p. 190-191 °C), IR (ν_{\max} , NaCl/cm⁻¹) 2920, 2850, 1747, 1705, 1664, 1436, 1254. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.48, 4.11 (each 2H, t, $J = 6.4$ Hz, CH₂CH₂), 5.84 (1H, d, $J = 8.0$ Hz, 5-CH), 7.28, 7.51 (each 2H, t, $J = 8.0$ Hz, Ar-H), 7.66 (1H, d, $J = 8.0$ Hz, 6-CH), 7.93 (1H, d, $J = 8.0$ Hz, ArH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 22.9 (CH₂I), 51.8 (CH₂N), 102.0 (5-CH), 129.2, 130.5 (2 x Ar-CH), 133.3 (Ar-C), 135.9 (Ar-CH), 144.2 (6-CH), 149.1, 162.5, 171.1 (3 x C=O). MS, $m/z = 370$ (M⁺), 342, 277, 215, 188, 155, 105.

3-Benzoyl-1-(2-bromoethyl)thymine. Prepared according to method A as previously reported, using 3-benzoylthymine (5.00 g, 21.65 mmol), triphenylphosphine (6.84 g, 26.11 mmol) and 2-bromoethanol (1.85 mL, 3.36 g, 26.08 mmol) in dry dioxane (100 mL) at 5 °C with DIAD (5.13 cm³, 5.27 g, 26.08 mmol) to give the title compound as a white solid (7.32 g, 56%), m.p. 166-167 °C (lit.,¹ m.p. 183-184 °C), IR (ν_{\max} , NaCl/cm⁻¹) 2898, 2342, 1746, 1699, 1662, 1462, 1414, 1429, 1376. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.98 (3H, s, CH₃), 3.44, 4.07 (each 2H, t, $J = 6.6$ Hz, CH₂CH₂), 7.13 (1H, d, $J = 1.2$ Hz, 6-CH), 7.50 (2H, t, $J = 7.4$ Hz, Ar-H), 7.65 (1H, t, $J = 7.4$ Hz, Ar-H), 7.93 (2H, d, $J = 7.4$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 11.6 (CH₃), 36.7 (CH₂Br), 50.4 (CH₂N), 109.7 (5-C), 128.3, 129.6 (2 x Ar-CH), 130.6 (Ar-C), 134.2 (Ar-CH), 139.3 (6-CH), 148.7, 162.1, 167.8 (3 x C=O). MS, $m/z = 338/336$ (M⁺), 310, 308, 277, 254, 252, 230, 188, 105. HRMS: Calcd for C₁₄H₁₃⁷⁹BrN₂O₃: M⁺ 336.0110; found: M⁺ 336.0112.

3-Benzoyl-1-(2-iodoethyl)thymine (9b). Prepared according to general procedure C as previously reported, using 3-benzoyl-1-(2-bromoethyl)thymine (5.00 g, 14.84 mmol) and dry NaI (11.11 g, 74.07 mmol) in dry acetone (100 mL) to yield the title compound as a pale yellow solid (3.11 g, 54%), m.p. 145-147 °C (lit.,¹ m.p. 138-139 °C), IR (ν_{\max} (NaCl/cm⁻¹) 2954, 1743, 1702, 1640, 1451, 1411, 1363. ¹H NMR (400 MHz, CDCl₃), δ_{H} 2.04 (3H, s, CH₃), 3.57, 4.33 (each 2H, t, $J = 6.8$ Hz, CH₂CH₂), 7.47 (2H, t, $J = 7.4$ Hz, Ar-H), 7.62 (1H, t, $J = 7.4$ Hz, Ar-H), 7.66 (1H, d, $J = 2.6$ Hz, 6-CH), 7.75 (2H, d, $J = 7.4$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 13.2 (CH₃), 29.4 (CH₂I), 42.0 (CH₂N), 112.3 (5-C), 128.6, 129.3 (2 x Ar-CH), 132.8 (Ar-C), 134.0 (Ar-CH), 140.7 (6-CH), 149.9, 162.9, 169.7 (3 x C=O). MS, $m/z = 384$ (M⁺), 356, 277, 215, 230, 155, 105.

N-(2-Chloroethyl)indole. Prepared according to method B, using indole (2.00 g, 17.09 mmol), 1,2-dichloroethane (150 mL), tetrabutylammonium iodide (0.10 g, 5% w/w), KOH (5.75 g, 0.10 mol) and K₂CO₃ (5.90 g, 0.04 mol). After stirring for 24 h, work up gave the title compound as a brown oil (2.46 g, 81%), IR (ν_{\max} , NaCl/cm⁻¹) 3050, 2959, 1614, 1454, 1334, 1093. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.84, 4.48 (each 2H, t, $J = 6.6$ Hz, CH₂CH₂), 6.61 (1H, m, indole-CH), 7.18-7.28 (3H, m, 2 x Ar-H, indole-CH), 7.43, 7.71 (each 1H, d, $J = 7.9$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 42.5 (CH₂Cl), 48.0 (CH₂N), 102.0, 111.0 (2 x indole-CH), 120.0, 121.9, 124.2 (3 x Ar-CH), 127.9 (Ar C), 128.8 (Ar-CH), 135.8 (Ar-C). MS, $m/z = 180$ (MH⁺), 144, 116, 90. Anal. Calcd for C₁₀H₁₀ClN: C, 66.86; H, 5.61; N, 7.80%, Found: C, 66.84; H, 5.91; N, 7.63%.

N-(2-Iodoethyl)indole (9c). Prepared according to general procedure C, using N-(2-chloroethyl)indole (1.40 g, 7.80 mmol) and dry NaI (5.84 g, 38.93 mmol) in dry acetone (100

mL). After work up this gave the title compound as a dark brown oil (1.36 g, 64%), IR (ν_{\max} , NaCl/cm⁻¹) 3031, 2924, 1563, 1439, 1329, 1234, 1126. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.44, 4.55 (each 2H, t, $J = 7.7$ Hz, CH₂CH₂), 6.60 (1H, m, indole-CH), 7.41-7.92 (5H, m, 4 x Ar-H, indole-CH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 44.3 (CH₂I), 59.2 (CH₂N), 101.9, 106.7 (2 x indole-CH), 109.7 (Ar-C), 119.2, 127.3, 127.8, 134.1 (4 x Ar-CH), 145.5 (Ar-C). MS, $m/z = 272$ (MH⁺), 130, 117, 90. Anal. Calcd for C₁₀H₁₀IN: C, 44.30; H, 3.72; N, 5.17%, Found: C, 44.33; H, 3.90; N, 5.04%.

***N*¹-(2-Bromoethyl)benzimidazole.** Prepared according to method A, using benzimidazole (2.00 g, 16.95 mmol), triphenylphosphine (5.33 g, 20.34 mmol) and 2-bromoethanol (1.44 mL, 2.53 g, 20.28 mmol) in dry dioxane (100 mL) at 5 °C with DIAD (4.00 mL, 4.11 g, 20.34 mmol) to yield the title compound as colourless crystals (3.75 g, 98%), m.p. 77-79 °C, IR (ν_{\max} , NaCl/cm⁻¹) 2980, 2360, 1719, 1497, 1438, 1246. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.71, 4.61 (each 2H, t, $J = 6.0$ Hz, CH₂CH₂), 7.23-7.62 (4H, m, Ar-H), 7.90 (1H, s, 2-CH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 39.2 (CH₂Br), 45.1 (CH₂N), 109.2 (2-CH), 119.3, 123.1, 129.0, 130.4 (4 x Ar-CH), 133.7, 144.9 (2 x ArC). MS, $m/z = 226/224$ (M⁺), 145, 131, 118, 104, 90. HRMS: Calcd for C₉H₉⁷⁹BrN₂: M⁺ 223.9950; found: M⁺ 223.9950.

***N*¹-(2-Chloroethyl)benzimidazole.** Prepared according to method B, using benzimidazole (1.00 g, 8.47 mmol) in 1,2-dichloroethane (70 mL) with tetrabutylammonium iodide (0.05 g, 5% w/w), KOH (2.85 g, 50.89 mmol) and K₂CO₃ (2.93 g, 21.23 mmol). After stirring for 4 h, work up gave title compound (0.85 g, 54%), m.p. 76-78 °C (lit.,¹³ m.p. 84-86 °C), IR (ν_{\max} , NaCl/cm⁻¹) 3056, 2960, 2359, 1615, 1494, 1437, 1257. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.76, 4.42 (each 2H, t, $J = 6.0$ Hz, CH₂CH₂), 7.22-7.45 (3H, m, Ar-H), 7.59 (1H, m, Ar-H), 7.89 (1H, s, 2-CH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 42.5 (CH₂Cl), 47.0 (CH₂N), 109.6 (2-CH), 121.0, 122.9, 129.0, 132.4 (4 x Ar-CH), 133.3, 143.7 (2 x Ar-C). MS, $m/z = 180$ (M⁺), 144, 131, 117, 104, 90. HRMS: Calcd for C₉H₉³⁵ClN₂: M⁺ 180.0454; found: M⁺ 180.0458. Anal. Calcd for C₉H₉ClN₂: C, 59.84; H, 5.02; N, 15.51%, Found: C, 60.28; H, 5.22; N, 15.98%.

Also prepared by a modification to method A, using benzimidazole (1.00 g, 8.47 mmol), triphenylphosphine (2.66 g, 10.15 mmol) and 2-chloroethanol (0.82 g, 10.19 mmol) in dry dioxane (20 mL) at 5 °C with DIAD (2.00 mL, 2.05 g, 10.17 mmol). After stirring for 16 h, work up gave title compound (1.23 g, 80%). Data as for material prepared using method B.

***N*¹-(2-Iodoethyl)benzimidazole (9d).** Prepared according to general procedure C, using *N*¹-(2-chloroethyl)benzimidazole (2.00 g, 11.08 mmol) and dry NaI (8.12 g, 54.13 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow solid (2.12 g, 72%), IR (ν_{\max} (NaCl/cm⁻¹) 2980, 1653, 1506, 1374, 1240, 1108. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.51, 4.58 (each 2H, t, $J = 7.2$ Hz, CH₂CH₂), 7.34-7.53 (3H, m, Ar-H), 7.83 (1H, m, Ar-H), 7.99 (1H, s, 2-CH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 44.2 (CH₂I), 53.1 (CH₂N), 109.9 (2-CH), 121.3, 121.9, 124.9, 130.2 (4 x Ar-CH), 140.8, 146.4 (2 x Ar-C). MS, $m/z = 272$ (M⁺), 144, 131, 118, 104, 90. HRMS: Calcd for C₉H₉IN₂: M⁺ 271.9812; found: M⁺ 271.9808.

Also prepared according to general procedure C, but using *N*¹-(2-bromoethyl)benzimidazole (2.00 g, 8.89 mmol) and dry NaI (6.66 g, 44.40 mmol) in dry acetone (100 mL). Work up gave

the title compound as a pale yellow solid (1.82 g, 75%). Data as for material prepared using the 2-chloroethyl derivative.

***N*¹- and *N*²-(2-Bromoethyl)benzotriazole.** Prepared according to method A, using benzotriazole (2.00 g, 16.81 mmol), triphenylphosphine (5.28 g, 20.15 mmol) and 2-bromoethanol (1.43 mL, 2.52 g, 20.16 mmol) in dry dioxane (100 mL) with DIAD (3.97 mL, 4.08 g, 20.18 mmol). After workup and column chromatography this yielded the title compounds. *N*¹-(2-Bromoethyl)benzotriazole: yellow crystals (1.03 g, 27%), m.p. 107 °C (lit.,¹⁴ m.p. 119-120 °C), IR (ν_{\max} , NaCl/cm⁻¹) 2929, 1650, 1444, 1329, 1249. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.88, 5.03 (each 2H, t, $J = 6.6$ Hz, CH₂CH₂), 7.41 (1H, dd, $J = 7.4, 8.4$ Hz, Ar-H), 7.54 (1H, dd, $J = 6.8, 7.4$ Hz, Ar-H), 7.60 (1H, d, $J = 8.4$ Hz, Ar-H), 8.00 (1H, d, $J = 6.8$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 28.9 (CH₂Br), 49.3 (CH₂N), 109.2, 120.2, 124.2, 127.8 (4 x Ar-CH), 133.3, 145.8 (2 x Ar-C). MS, $m/z = 227/225$ (M⁺), 146, 132, 119. HRMS: Calcd for C₈H₈⁷⁹BrN₃: M⁺ 224.9902; found: M⁺ 224.9902. Anal. Calcd for C₈H₈BrN₃: C, 42.50; H, 3.57; N, 18.59%, Found: C, 42.23; H, 3.78; N, 18.22%. *N*²-(2-Bromoethyl)benzotriazole: yellow crystals (2.13 g, 56%), m.p. 62-64 °C (lit.,¹⁵ m.p. 59-60 °C), IR (ν_{\max} , NaCl/cm⁻¹) 2960, 1648, 1583, 1454, 1284. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.98, 5.12 (each 2H, t, $J = 6.7$ Hz, CH₂CH₂), 7.42, 7.88 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 28.1 (CH₂Br), 57.3 (CH₂N), 118.2, 126.7 (Ar-CH), 144.5 (Ar-C). MS, $m/z = 227/225$ (M⁺), 146, 132, 119. HRMS: Calcd for C₈H₈⁷⁹BrN₃: M⁺ 224.9902; found: M⁺ 224.9902. Anal. Calcd for C₈H₈BrN₃: C, 42.50; H, 3.57; N, 18.59%, Found: C, 42.97; H, 3.24; N, 18.62%.

***N*¹-(2-Chloroethyl)benzotriazole.** Prepared according to method B, using benzotriazole (2.00 g, 16.81 mmol), 1,2-dichloroethane (70 mL), tetrabutylammonium iodide (0.10 g, 5% w/w), KOH (5.65 g, 0.10 mol) and K₂CO₃ (5.80 g, 0.04 mol). Work up gave the title compound as white crystals (2.01 g, 66%), m.p. 118 °C (lit.,¹⁶ m.p. 108-109 °C), IR (ν_{\max} , NaCl/cm⁻¹) 2938, 1638, 1457, 1316, 1274, 1228, 1164. ¹H NMR (400 MHz, CDCl₃), δ_{H} 4.06, 4.96 (each 2H, t, $J = 6.2$ Hz, CH₂CH₂), 7.41 (1H, t, $J = 8.4$ Hz, Ar-H), 7.55 (1H, dd, $J = 6.8, 8.4$ Hz, Ar-H), 7.61 (1H, d, $J = 6.8$ Hz, Ar-H), 8.10 (1 H, d, $J = 8.4$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 42.3 (CH₂Cl), 49.5 (CH₂N), 109.3, 120.2, 124.1, 127.8 (4 x Ar-CH), 133.5, 145.9 (2 x Ar-C). MS, $m/z = 182$ (MH⁺), 146, 131, 119. HRMS: Calcd for C₈H₈³⁵ClN₃: MH⁺ 182.0479; found: MH⁺ 182.0478. None of the isomeric *N*²-(2-chloroethyl)benzotriazole was isolated.

***N*¹-(2-Iodoethyl)benzotriazole (9e).** Prepared according to general procedure C, using *N*¹-(2-bromoethyl)benzotriazole (1.00 g, 4.42 mmol) and dry NaI (3.32 g, 22.13 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow solid (1.15 g, 95%), m.p. 106 °C, IR (ν_{\max} , NaCl/cm⁻¹) 3018, 2350, 1722, 1613, 1435, 1311. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.67, 5.02 (each 2H, t, $J = 7.4$ Hz, CH₂CH₂), 7.36 (1H, d, $J = 6.6$ Hz, Ar-H), 7.48 (2H, m, Ar-H), 8.04 (1H, d, $J = 8.4$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 30.2 (CH₂I), 50.8 (CH₂N), 109.5, 120.8, 124.9, 128.3 (4 x Ar-CH), 133.8, 146.4 (2 x Ar-C). MS, $m/z = 274$ (MH⁺), 146, 132, 118. HRMS: Calcd for C₈H₈IN₃: MH⁺ 273.9838; found: MH⁺ 273.9836.

Also prepared according to general procedure C, but using *N*¹-(2-chloroethyl)benzotriazole (2.00 g, 11.02 mmol) and dry NaI (8.22 g, 54.80 mmol) in dry acetone (100 mL). Work up gave the

title compound as a pale yellow solid (2.82 g, 94%). Data as for material prepared using the bromoethyl derivative.

***N*²-(2-Iodoethyl)benzotriazole (9f).** Prepared according to general procedure C, using *N*²-(2-bromoethyl)benzotriazole (2.10 g, 9.29 mmol) and dry NaI (6.96 g, 46.40 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow solid (2.46 g, 97%), m.p. 66-69 °C, IR (ν_{\max} , NaCl/cm⁻¹) 3030, 1636, 1439, 1234. ¹H NMR (400 MHz, CDCl₃), δ_{H} 3.75, 5.09 (each 2H, t, *J* = 7.3 Hz, CH₂CH₂), 7.42 (2H, m, Ar-H), 7.88 (2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 31.3 (CH₂I), 59.3 (CH₂N), 119.2, 127.8 (2 x Ar-CH), 145.5 (Ar-C). MS, *m/z* = 273 (M⁺), 146, 131, 120. HRMS: Calcd for C₈H₈IN₃: M⁺ 272.9765; found: M⁺ 272.9757.

***N*⁶-(*tert*-Butoxycarbonyl)adenine (6-*tert*-butoxycarbonylaminopurine)**

To a dry round-bottomed flask were added adenine (2.00 g, 14.81 mmol), di-*tert*-butyl dicarbonate (3.23 g, 14.82 mmol) and diisopropylethylamine (3.84 mL, 2.85 g, 22.09 mmol) in dry CH₂Cl₂ (100 mL), and the mixture stirred for 2 h at 20 °C. The solution was washed with water (3 x 20 mL), dried (MgSO₄), filtered and the solvent removed under reduced pressure to give the title compound (3.46 g, quantitative yield), m.p. 280-282 °C (lit.,¹⁷ m.p. 280-300 °C), IR (ν_{\max} , NaCl/cm⁻¹) 3174, 2922, 1697, 1437, 1386. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.71 (9H, s, C(CH₃)₃), 1.85 (1H, br, 9-NH), 5.78 (1H, br, *NHBoc*), 8.51, 8.52 (each 1H, s, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 27.9 (C(CH₃)₃), 37.6 (C(CH₃)₃), 144.6 (Ar-CH), 147.9, 151.8 (2 x Ar-C), 155.4 (Ar-CH), 161.5 (Ar-C), 204.6 (C=O). MS, *m/z* = 235 (M⁺), 134, 119, 90. HRMS: Calcd for C₁₀H₁₃N₅O₂: M⁺ 235.1069; found: M⁺ 235.1069.

***N*⁹-(2-Bromoethyl)-6-*tert*-butoxycarbonylaminopurine**

Prepared according to method A, using 6-*tert*-butoxycarbonylaminopurine (3.45 g, 14.68 mmol), triphenylphosphine (4.62 g, 17.63 mmol) and 2-bromoethanol (1.24 mL, 2.18 g, 17.48 mmol) in dry dioxane (100 mL) with DIAD (3.47 mL, 3.56 g, 17.64 mmol). After workup this yielded the title compound as colourless crystals (1.88 g, 37%), m.p. 197 °C. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.63 (9H, s, C(CH₃)₃), 3.82, 4.41 (each 2H, t, *J* = 7.8 Hz, CH₂CH₂), 6.30 (1H, br, NH), 8.19, 8.47 (each 1H, s, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 31.9 (C(CH₃)₃), 38.2 (C(CH₃)₃), 38.6 (CH₂Br), 51.9 (CH₂N), 137.5 (Ar-CH), 147.2, 149.3 (2 x Ar-C), 157.8 (Ar-CH), 161.2 (Ar-C), 216.4 (C=O). MS, *m/z* = 343/341 (M⁺), 241, 133, 118, 90. LC-MS: Calcd for C₁₂H₁₆⁷⁹BrN₅O₂: M⁺ 341.05; found: M⁺ 341.05.

***N*⁹-(2-Iodoethyl)-6-*tert*-butoxycarbonylaminopurine (9g).** Prepared according to general procedure C, using *N*⁹-(2-bromoethyl)-6-*tert*-butoxycarbonylaminopurine (1.60 g, 4.68 mmol) and dry NaI (3.50 g, 23.33 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow oil (0.92 g, 51%), IR (ν_{\max} , NaCl/cm⁻¹) 3419, 1642, 1516, 1436, 1373, 1277. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.39 (9H, s, C(CH₃)₃), 3.57, 4.51 (each 2H, t, *J* = 6.6 Hz, CH₂CH₂), 5.99 (1H, br, NH), 7.80, 8.34 (each 1H, s, Ar-CH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 28.2 (C(CH₃)₃), 29.9 (CH₂I), 38.2 (C(CH₃)₃), 51.7 (CH₂N), 139.1 (Ar-CH), 148.7, 150.7 (2 x Ar-C),

155.8 (Ar-CH), 163.9 (Ar-C), 213.4 (C=O). MS, $m/z = 389$ (M^+), 288, 134, 118, 90. HRMS: Calcd for $C_{12}H_{16}IN_5O_2$: M^+ 389.0351; found: M^+ 389.0375.

***N*⁹-(2-Bromoethyl)purine.** Prepared according to method A, using purine (0.50 g, 4.16 mmol), triphenylphosphine (1.32 g, 5.04 mmol) and 2-bromoethanol (0.36 mL, 0.63 g, 5.07 mmol) in dry dioxane (100 mL) and DIAD (0.99 mL, 1.02 g, 5.03 mmol). After workup this yielded the title compound as a brown solid (0.62 g, 73%), m.p. 188-190 °C (lit.,¹⁸ m.p. 195 °C), IR (ν_{max} , NaCl/cm⁻¹) 3289, 2979, 1700, 1496, 1456, 1267, 1219. ¹H NMR (400 MHz, CDCl₃), δ_H 3.68, 4.58 (each 2H, t, $J = 5.5$ Hz, CH₂CH₂), 8.14, 8.74, 8.92 (each 1H, s, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 55.0 (CH₂Br), 60.3 (CH₂N), 138.8, 141.6 (2 x Ar-CH), 143.7 (Ar-C), 155.6 (Ar-CH), 160.8 (Ar-C). MS, $m/z = 228/226$ (M^+), 144, 117, 91. HRMS: Calcd for $C_7H_7^{81}BrN_4$ requires: M^+ 227.9834; found: M^+ 227.9851.

***N*⁹-(2-Chloroethyl)purine.** Prepared according to method B, using purine (0.50 g, 4.16 mmol) in 1,2-dichloroethane (50 mL) with tetrabutylammonium iodide (0.03 g, 5% w/w), KOH (1.41 g, 25.18 mmol) and K₂CO₃ (1.45 g, 10.51 mmol). Purification gave the title compound as a yellow oil (0.39 g, 59%), IR (ν_{max} , NaCl/cm⁻¹) 3082, 2962, 1594, 1502, 1408, 1305, 1200. ¹H NMR (400 MHz, CDCl₃), δ_H 3.91, 4.59 (each 2H, t, $J = 5.6$ Hz, CH₂CH₂), 8.16, 8.91, 9.09 (each 1H, s, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 42.2 (CH₂Cl), 45.6 (CH₂N), 134.1 (Ar-C), 145.8, 148.7 (2 x Ar-CH), 151.1 (Ar-C), 153.2 (Ar-CH). MS, $m/z = 183$ (MH⁺), 144, 117, 91. HRMS: Calcd for $C_7H_7^{35}ClN_4$: MH⁺ 183.0432; found: MH⁺ 183.0458.

***N*⁹-(2-Iodoethyl)purine (9h).** Prepared according to general procedure C, using *N*⁹-(2-bromoethyl)purine (0.25 g, 1.10 mmol) and dry NaI (0.92 g, 6.13 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow oil (0.24 g, 79%), IR (ν_{max} (NaCl/cm⁻¹) 2983, 1642, 1496, 1436, 1312. ¹H NMR (400 MHz, CDCl₃), δ_H 3.61, 4.66 (each 2H, t, $J = 6.5$ Hz, CH₂CH₂), 8.16, 8.95, 9.14 (each 1H, s, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 23.2 (CH₂I), 45.9 (CH₂N), 135.0 (Ar-C), 144.9, 148.8, (2 x Ar-CH), 152.0 (Ar-C), 152.6 (Ar-CH). MS, $m/z = 275$ (MH⁺), 145, 117, 90. HRMS: Calcd for $C_7H_7IN_4$: MH⁺ 274.9790; found MH⁺ 274.9768.

***N*-(3-Chloropropyl)indole.** Prepared according to method B, using indole (2.00 g, 17.09 mmol), 1,3-dichloropropane (150 cm³), tetrabutylammonium iodide (0.10 g, 5% w/w), KOH (5.75 g, 0.10 mol) and K₂CO₃ (5.90 g, 0.04 mol). Purification by flash column chromatography gave title compound as a brown oil (2.01 g, 61%), IR (ν_{max} , NaCl/cm⁻¹) 3054, 2970, 2862, 2218, 1718, 1589, 1437, 1191. ¹H NMR (400 MHz, CDCl₃), δ_H 2.35 (2H, q, $J = 6.4$ Hz, CH₂), 3.55 (2H, t, $J = 6.4$ Hz, CH₂Cl), 4.42 (2H, t, $J = 6.4$ Hz, CH₂N), 6.75, 7.28 (each 1H, d, $J = 3.1$ Hz, indole-CH), 7.38, 7.46 (each 1H, dd, $J = 7.1, 8.1$ Hz, Ar-H), 7.56, 7.80 (each 1H, d, $J = 8.1$ Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 30.4 (CH₂), 42.2 (CH₂Cl), 43.1 (CH₂N), 101.7, 109.5 (2 x indole-CH), 117.4, 119.7, 121.6, 122.1 (4 x Ar-CH), 133.7, 136.1 (2 x Ar-C). MS, $m/z = 194$ (MH⁺), 158, 118, 91. HRMS: Calcd for $C_{11}H_{12}^{35}ClN$: MH⁺ 194.0731; found: MH⁺ 194.0723.

***N*¹-(3-Bromopropyl)benzimidazole.** Prepared according to method A, using benzimidazole (1.00 g, 8.47 mmol), triphenylphosphine (2.66 g, 10.15 mmol) and 3-bromopropan-1-ol (0.92 mL, 1.41 g, 10.15 mmol) in dry dioxane (100 mL) at 5 °C with DIAD (4.00 mL, 4.11 g, 20.33 mmol) to yield the title compound as an oil (1.58 g, 78%). ¹H NMR (400 MHz, CDCl₃), δ_H 2.10

(2H, m, CH₂), 3.54 (2H, t, *J* = 6.3 Hz, CH₂Br), 4.20 (2H, t, *J* = 6.6 Hz, CH₂N), 7.23, 7.74 (each 2H, m, Ar-H.), 7.83 (1H, s, 2-CH). ¹³C NMR (100 MHz, CDCl₃), δ_C 27.0 (CH₂), 29.6 (CH₂Br), 34.0 (CH₂N), 109.6 (2-CH), 120.5, 122.1, 123.0, (3 x Ar-CH), 133.7 (Ar-C), 142.9 (Ar-CH), 143.9 (Ar-C). MS, *m/z* = 241/239 (MH⁺), 160, 118, 104, 90. HRMS: Calcd for C₁₀H₁₁⁷⁹BrN₂: MH⁺ 239.0179; found: MH⁺ 239.0150.

***N*¹-(3-Chloropropyl)benzimidazole.** Prepared according to method B, using benzimidazole (1.00 g, 8.47 mmol) in 1,3-dichloropropane (70 mL) with tetrabutylammonium iodide (0.05 g, 5% w/w), KOH (2.85 g, 50.89 mmol) and K₂CO₃ (2.93 g, 21.23 mmol). After stirring for 8 h, work up gave title compound as an oil (1.04 g, 62%). ¹H NMR (400 MHz, CDCl₃), δ_H 2.23 (2H, m, CH₂), 3.38 (2H, t, *J* = 6.5 Hz, CH₂Cl), 4.32 (2H, t, *J* = 6.5 Hz, CH₂N), 7.23, 7.74 (each 2H, m, Ar-H.), 7.83 (1H, s, 2-CH). ¹³C NMR (100 MHz, CDCl₃), δ_C 22.7 (CH₂), 32.2 (CH₂Cl), 46.4 (CH₂N), 102.8 (2-CH), 121.1, 128.9, 129.2, 133.6 (4 x Ar-CH), 138.2, 147.3 (2 x ArC). MS, *m/z* = 194 (M⁺), 159, 118, 104, 90. HRMS: Calcd for C₁₀H₁₁³⁵ClN₂: M⁺ 194.0611; found: M⁺ 194.0609.

Attempted preparation of *N*¹-(3-iodopropyl)benzimidazole (10c): Quaternary salt (12a)

Attempted according to general procedure C, using *N*¹-(3-bromopropyl)benzimidazole and (1.00 g, 4.18 mmol) and dry NaI (3.15 g, 21.00 mmol) in dry acetone (100 mL), and also by using *N*¹-(3-chloropropyl)benzimidazole (1.00 g, 5.14 mmol) and dry NaI (3.86 g, 25.73 mmol) in dry acetone (100 mL). Work up in both cases did not give the title compound, but gave what is provisionally assigned as the dimeric quaternary salt **12a**. ¹H NMR (250 MHz, CDCl₃), δ_H 1.90 (2H, m, CH₂), 4.90 (4H, t, *J* = 6.9 Hz, 2 x CH₂), 7.22-7.78 (4H, m, Ar-H), 7.95 (1H, s, 2-CH). LC-MS: Calcd for double quaternary cation C₂₀H₂₂N₄: M²⁺ 159.09, [M-H]⁺ 317.18; found: M²⁺ 158.93, [M-H]⁺ 317.30.

***N*¹-(3-Bromopropyl)benzotriazole.** Prepared according to method A, using benzotriazole (4.00 g, 33.61 mmol), triphenylphosphine (10.57 g, 40.34 mmol) and 3-bromopropan-1-ol (3.64 mL, 5.61 g, 40.33 mmol) in dry dioxane (100 mL) with DIAD (7.93 mL, 8.14 g, 40.32 mmol). After workup this yielded the title compound as a yellow oil (3.15 g, 78%), IR (ν_{max}, NaCl/cm⁻¹) 2900, 1660, 1542, 1447, 1287. ¹H NMR (400 MHz, CDCl₃), δ_H 2.33 (2H, tt, *J* = 6.8, 7.3 Hz, CH₂) 3.28 (2H, t, *J* = 6.8 Hz, CH₂Cl), 4.75 (2H, t, *J* 7.3 Hz, CH₂N), 7.37 (1H, m, Ar-H), 7.46-7.58 (2H, m, Ar-H), 8.01 (1H, d, *J* = 7.9 Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 27.8 (CH₂), 32.4 (CH₂Cl), 51.9 (CH₂N), 109.7, 121.2, 123.7, 126.5 (4 x Ar-CH), 136.3, 144.4 (2 x Ar-C). MS, *m/z* = 240 (MH⁺), 160, 132, 119. HRMS: Calcd for C₉H₁₀⁷⁹BrN₃: MH⁺ 240.0131; found: MH⁺ 240.0153. None of the isomeric *N*²-(3-bromopropyl)benzotriazole was isolated.

***N*¹- and *N*²-(3-Chloropropyl)benzotriazole.** Prepared according to method B, using benzotriazole (2.00 g, 16.81 mmol), 1,3-dichloropropane (100 cm³), tetrabutylammonium iodide (0.10 g, 5% w/w), KOH (5.65 g, 0.10 mol) and K₂CO₃ (5.80g, 0.04 mol). Work up and chromatography gave the title compounds. *N*¹-(3-Chloropropyl)benzotriazole: cream oil (1.59 g, 52%), IR (ν_{max}, NaCl/cm⁻¹) 2879, 1687, 1523, 1438, 1202. ¹H NMR (400 MHz, CDCl₃), δ_H 2.45 (2H, tt, *J* = 6.0, 6.5 Hz, CH₂) 3.49 (2H, t, *J* = 6.0 Hz, CH₂Cl), 4.79 (2H, t, *J* = 6.5 Hz, CH₂N),

7.34, 7.48 (each 1H, t, $J = 8.3$ Hz, Ar-H), 7.58, 8.03 (each 1H, d, $J = 8.3$ Hz, Ar-H). ^{13}C NMR (100 MHz, CDCl_3), δ_{C} 32.4 (CH_2), 41.5 (CH_2Cl), 44.7 (CH_2N), 109.2, 120.0, 124.0, 127.5 (4 x Ar-CH), 133.2, 145.9 (2 x ArC). MS, $m/z = 196$ (MH^+), 160, 133, 120. HRMS: Calcd for $\text{C}_9\text{H}_{10}^{35}\text{ClN}_3$: MH^+ 196.0636; found: MH^+ 196.0642. Anal. Calcd for $\text{C}_9\text{H}_{10}\text{ClN}_3$: C, 55.25; H, 5.15; N, 21.48%; Found: C, 55.86; H, 4.74; N, 21.82%. *N*²-(3-Chloropropyl)benzotriazole: yellow oil (0.72 g, 24%), IR (ν_{max} , $\text{NaCl}/\text{cm}^{-1}$) 2901, 1732, 1647, 1519, 1468, 1232. ^1H NMR (400 MHz, CDCl_3), δ_{H} 2.53 (2H, tt, $J = 6.2, 6.5$ Hz, CH_2), 3.56 (2H, t, $J = 6.2$ Hz, CH_2Cl), 4.89 (2H, t, $J = 6.5$ Hz, CH_2N), 7.35, 7.84 (each 2H, m, Ar-H). ^{13}C NMR (100 MHz, CDCl_3), δ_{C} 32.6 (CH_2), 41.4 (CH_2Cl), 53.4 (CH_2N), 117.9, 126.4 (2 x Ar-CH), 144.4 (Ar-C). MS, $m/z = 195$ (M^+), 146, 133, 119. HRMS: Calcd for $\text{C}_9\text{H}_{10}^{35}\text{ClN}_3$: M^+ 195.0563; found: M^+ 195.0560. Anal. Calcd for $\text{C}_9\text{H}_{10}\text{ClN}_3$: C, 55.25; H, 5.15; N, 21.48%; Found: C, 54.91; H, 5.24; N, 20.94%.

***N*¹-(3-Iodopropyl)benzotriazole (10e).** Prepared according to general procedure C, using *N*¹-(3-chloropropyl)benzotriazole (1.50 g, 7.67 mmol) and dry NaI (6.19 g, 41.27 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow solid (1.73 g, 77%), m.p. 98-99 °C, IR (ν_{max} , $\text{NaCl}/\text{cm}^{-1}$) 2958, 1610, 1464, 1357, 1205. ^1H NMR (400 MHz, CDCl_3), δ_{H} 2.44 (2H, tt, $J = 6.4, 6.5$ Hz, CH_2), 3.12 (2H, t, $J = 6.5$ Hz, CH_2I), 4.73 (2H, t, $J = 6.4$ Hz, CH_2N), 7.01, 7.39 (each 1H, t, $J = 7.0$ Hz, Ar-H), 7.44, 7.68 (each 1H, d, $J = 8.1$ Hz, Ar-H). ^{13}C NMR (100 MHz, CDCl_3), δ_{C} 3.5 (CH_2I), 33.3 (CH_2), 45.9 (CH_2N), 109.3, 118.3, 121.5, 124.5 (4 x Ar-CH), 127.0, 136.8 (2 x ArC). MS, $m/z = 286$ ($[\text{M-H}]^+$), 160, 132, 119. HRMS: Calcd for $\text{C}_9\text{H}_{10}\text{IN}_3$: $[\text{M-H}]^+$ 285.9849; found: $[\text{M-H}]^+$ 285.9887.

***N*²-(3-Iodopropyl)benzotriazole (10f).** Prepared according to general procedure C, using *N*²-(3-chloropropyl)benzimidazole (3.20 g, 16.37 mmol) and dry NaI (13.20 g, 88.00 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow oil (2.94 g, 65%), IR (ν_{max} , $\text{NaCl}/\text{cm}^{-1}$) 2962, 1615, 1448, 1278. ^1H NMR (400 MHz, CDCl_3), δ_{H} 2.58 (2H, tt, $J = 6.4, 6.7$ Hz, CH_2), 3.15 (2H, t, $J = 6.7$ Hz, CH_2I), 4.81 (2H, t, $J = 6.4$ Hz, CH_2N), 7.37, 7.82 (each 2H, m, Ar-H). ^{13}C NMR (100 MHz, CDCl_3), δ_{C} 4.2 (CH_2I), 30.6 (CH_2), 50.9 (CH_2N), 115.7, 128.2 (2 x Ar-CH), 138.1 (Ar-C). MS, $m/z = 286$, M^+ , 160, 133, 119. HRMS: Calcd for $\text{C}_9\text{H}_{10}\text{IN}_3$: M^+ 285.9849; found: M^+ 285.9856.

Attempted preparation of *N*¹-(3-chlorobutyl)benzimidazole: quaternary salt (12b)

Attempted according to method B, using benzimidazole (1.00 g, 8.47 mmol) in 1,4-dichlorobutane (70 mL) with tetrabutylammonium iodide (0.05 g, 5% w/w), KOH (2.85 g, 50.89 mmol) and K_2CO_3 (2.30 g, 16.67 mmol). After stirring for 8 h, work up did not give the title compound, but gave what is provisionally assigned as the dimeric quaternary salt **12b**. ^1H NMR (400 MHz, $(\text{CD}_3)_2\text{SO}$), δ_{H} 2.08 (4H, br, 2 x CH_2), 4.57 (4H, t, $J = 5.6$ Hz, 2 x CH_2), 7.67, 8.34 (each 2H, m, Ar-H), 9.77 (1H, s, 2-CH). ^{13}C NMR (100 MHz, $(\text{CD}_3)_2\text{SO}$), δ_{C} 24.9 (CH_2), 26.0 (CH_2), 29.7 (CH_2), 46.2 (CH_2N), 113.8 (2-CH), 126.4, 126.8 (2 x ArCH), 130.1 (Ar-C), 141.5, 142.1 (2 x Ar-CH), 144.6 (Ar-C). LC-MS: Calcd for double quaternary cation $\text{C}_{22}\text{H}_{26}\text{N}_4$: M^{2+} 173.11, $[\text{M-H}]^+$ 345.21; found: M^{2+} 173.04, $[\text{M-H}]^+$ 345.64.

***N*¹- and *N*²-(4-Chlorobutyl)benzotriazole.** Prepared according to method B, using benzotriazole (6.00 g, 50.42 mmol) 1,3-dichlorobutane (150 mL), tetrabutylammonium iodide (0.30 g, 5% w/w), KOH (16.96 g, 0.30 mol) and K₂CO₃ (17.40 g, 0.13 mol). Work up gave the title compounds. *N*¹-(4-Chlorobutyl)benzotriazole: brown crystals (2.56 g, 24%), m.p. 104-106 °C, IR (ν_{\max} , NaCl/cm⁻¹) 3051, 2958, 1566, 1445, 1327, 1264. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.85, 2.17 (each 2H, m, CH₂) 3.56 (2H, t, *J* = 6.4 Hz, CH₂Cl), 4.68 (2H, t, *J* = 6.9 Hz, CH₂N), 7.38 (1H, d, *J* = 7.1 Hz, Ar-H), 7.51 (2H, m, Ar-H), 8.05 (1H, d, *J* = 8.4 Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 26.8, 29.4 (2 x CH₂), 44.1 (CH₂I), 47.3 (CH₂N), 109.2, 120.1, 123.9, 127.4 (4 x Ar-CH), 132.8, 146.0 (2 x Ar-C). MS, *m/z* = 210 (MH⁺), 175, 137, 121. HRMS: Calcd for C₁₀H₁₂³⁵ClN₃: MH⁺ 210.0792; found: MH⁺ 210.0794. *N*²-(4-Chlorobutyl)benzotriazole: brown crystals (2.23 g, 21%), m.p. 87-88 °C, IR (ν_{\max} , NaCl/cm⁻¹) 2980, 1699, 1504, 1385, 1240. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.81, 2.28 (each 2H, m, CH₂), 3.55 (2H, t, *J* = 6.4 Hz, CH₂I), 4.76 (2H, t, *J* = 6.8 Hz, CH₂N), 7.36, 7.87 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 27.1, 29.3, (2 x CH₂), 44.1 (CH₂I), 55.6 (CH₂N), 118.0, 126.3 (2 x Ar-CH), 144.3 (Ar-C). MS, *m/z* = 210 (MH⁺), 175, 137, 121. HRMS: Calcd for C₁₀H₁₂³⁵ClN₃ requires: MH⁺ 210.0792; Found: MH⁺ 210.0791.

***N*¹-(4-Iodobutyl)benzotriazole (11e).** Prepared according to general procedure C, using *N*¹-(4-chlorobutyl)benzotriazole (2.00 g, 9.55 mmol) and dry NaI (7.15 g, 47.67 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow oil (2.01 g, 70%), IR (ν_{\max} , NaCl/cm⁻¹) 3030, 2934, 1601, 1462. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.83, 2.13 (each 2H, m, CH₂), 3.18 (2H, t, *J* = 6.7 Hz, CH₂I), 4.68 (2H, t, *J* = 6.9 Hz, CH₂N), 7.33 (1H, d, *J* = 6.2 Hz, Ar-H), 7.48 (2H, m, Ar-H), 8.06 (1H, d, *J* = 7.2 Hz, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 6.9 (CH₂I), 32.5, 35.4 (2 x CH₂), 55.5 (CH₂N), 119.2, 123.8, 124.1, 128.9 (4 x Ar-CH), 137.0, 142.7 (2 x ArC). MS, *m/z* = 310 (M⁺), 174, 146, 120. HRMS: Calcd for C₁₀H₁₂IN₃: M⁺ 301.0078; found: M⁺ 301.0073. Anal. Calcd for C₁₀H₁₂IN₃: C, 39.89; H, 4.02; N, 13.95%; Found: C, 39.52; H, 4.00; N, 13.63%.

***N*²-(4-Iodobutyl)benzotriazole (11f).** Prepared according to general procedure C, using *N*²-(4-chlorobutyl)benzimidazole (2.00 g, 9.55 mmol) and dry NaI (7.15 g, 47.67 mmol) in dry acetone (100 mL). Work up gave the title compound as a pale yellow solid (2.00 g, 70%), (ν_{\max} , NaCl/cm⁻¹) 3101, 2944, 1599, 1445, 1221. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.84 (2H, tt, *J* = 6.9, 7.3 Hz, CH₂), 2.16 (2H, dt, *J* = 6.8, 7.3 Hz, CH₂), 3.22 (2H, t, *J* = 6.9 Hz, CH₂I), 4.77 (2H, t, *J* = 6.8 Hz, CH₂N), 7.36, 7.48 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 5.8 (CH₂I), 30.1, 30.8 (2 x CH₂), 55.3 (CH₂N), 118.0, 126.4 (2 x Ar-CH), 144.3 (Ar-C). MS, *m/z* = 301 (M⁺), 174, 147, 120. HRMS: Calcd for C₁₀H₁₂IN₃: M⁺ 301.0078; found: M⁺ 301.0069. Anal. Calcd for C₁₀H₁₂IN₃: C, 39.89; H, 4.02; N, 13.95%; Found: C, 39.33; H, 4.14; N, 14.28%.

***N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester (5).** To a dry round-bottomed flask was added (*R*)-cysteine ethyl ester hydrochloride (10.00 g, 53.86 mmol), di-*tert*-butyl dicarbonate (11.76 g, 53.94 mmol) and diisopropylethylamine (11.18 mL, 8.30 g, 64.31 mmol) in dry CH₂Cl₂ (200 mL). The mixture was stirred for 2 h at 20 °C and then washed with hydrochloric acid (1M; 2 x

100 mL), NaHCO₃ solution (1M; 2 x 100 ml), and water (2 x 100 mL), dried (MgSO₄), and the solvent removed under reduced pressure. Work up gave the title compound as off white oil (12.62 g, 94%), (ν_{\max} (NaCl/cm⁻¹) 3367, 2979, 1712, 1507, 1368, 1165, 1029. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.22 (3H, t, $J = 7.1$ Hz, CH₂CH₃), 1.38 (9H, s, C(CH₃)₃), 2.90 (2H, m, CH₂S), 4.17 (2H, q, $J = 7.1$ Hz, CH₂CH₃), 4.50 (1H, m, CH), 5.43 (1H, br, NH). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 14.1 (CH₃), 27.2 (CH₂S), 28.2 (C(CH₃)₃), 56.2 (CH), 68.9 (OCH₂) 80.2 (C(CH₃)₃), 170.4 (C=O). MS, $m/z = 250$ (MH⁺), 211, 193, 135. HRMS: Calcd for C₁₀H₁₉NO₄S: MH⁺ 250.1107; found: MH⁺ 250.1108.

General procedure for preparation of S-linked heterocyclic amino acids, exemplified by *N*-*tert*-butoxycarbonyl-S-[2-(3-benzoyluracil-1-yl)ethyl]cysteine ethyl ester (13a)

To a flame-dried round-bottom flask under a positive pressure of nitrogen was added NaH (60% in mineral oil; 0.10 g, 2.50 mmol.), which was washed for 5 min with hexane, the hexane decanted, and then cooled to 0 °C. Dry THF (25 mL) was added to the reaction vessel followed by *N*-*tert*-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.50 g, 2.01 mmol) in dry THF (25 mL). The reaction mixture was allowed to warm to 20 °C over 15 min before the addition of 3-benzoyl-1-(2-iodoethyl)uracil **9a** (2.98 g, 8.05 mmol), after which the reaction was heated at reflux for 16 h. The THF was then removed under reduced pressure and the residue taken up in CH₂Cl₂-water (100 mL, 1:1 v/v). The CH₂Cl₂ layer was separated, washed with water (20 mL), dried (MgSO₄) and the solvent removed under reduced pressure. The residue was purified by column chromatography, eluting with ethyl acetate:hexane, to afford the title compound as a yellow oil (0.62 g, 64%), IR (ν_{\max} , NaCl/cm⁻¹) 3416, 2930 2857, 2357, 1713, 1650, 1555, 1453, 1249, 1162. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.19 (3H, t, $J = 7.1$ Hz, CH₂CH₃), 1.95 (9H, s, C(CH₃)₃), 3.13 (1H, apparent t, $J = 6.6$ Hz, CH₂S), 3.35 (2H, t, $J = 6.4$ Hz, CH₂S), 4.04 (2H, q, $J = 7.1$ Hz, CH₂CH₃), 4.86 (3H, m, CH₂N, CH), 5.72 (1H, d, $J = 8.0$ Hz, COCH=CH), 7.19-7.62 (6H, m, COCH=CH, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 7.1 (CH₂CH₃), 20.7 (C(CH₃)₃), 50.0, 52.4 (2 x CH₂S), 59.2 (CH₂N), 62.1 (OCH₂), 64.9 (C(CH₃)₃), 68.3 (CH), 100.5 (COCH=CH), 127.2 (2 x Ar-CH), 128.0 (COCH=CH), 130.5 (Ar-CH), 143.7 (Ar-C), 148.3, 155.4, 161.2, 167.4, 169.9 (5 x C=O). MS, $m/z = 272$ (MH⁺-NHBoc, C₂₃H₂₉N₃O₇S: C, 56.20; H, 5.95; N, 8.55%; Found C, 56.77; H, 5.82; N, 8.30%.

***N*-*tert*-Butoxycarbonyl-S-[2-(3-benzoylthymine-1-yl)ethyl]cysteine ethyl ester (13b).** Prepared according to the method used for **13a** but allowing the mixture to warm to 20 °C over 20 min before the addition of 3-benzoyl-1-(2-iodoethyl)thymine **9b** (3.09 g, 8.05 mmol). The title compound was isolated as a yellow oil (0.61 g, 62%), IR (ν_{\max} , NaCl/cm⁻¹) 3284, 2978, 2930, 1745, 1698, 1657, 1454, 1367, 1251, 1175. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.19 (3H, t, $J = 6.6$ Hz, CH₂CH₃), 1.75 (9H, s, C(CH₃)₃), 1.81 (3H, s, CH₃), 3.26 (1H, apparent t, $J = 7.8$ Hz, CH₂S), 3.39 (2 H, t, $J = 7.1$ Hz, CH₂S), 4.11 (2H, q, $J = 6.6$ Hz, CH₂CH₃), 4.16 (2H, t, $J = 7.1$ Hz, CH₂N), 4.68 (1H, t, $J = 7.8$ Hz, CH), 5.71 (1H, s, C=CH), 7.24-7.95 (5H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 11.2 (CH₂CH₃), 11.7 (CCH₃) 27.3 (C(CH₃)₃), 40.9, 48.3 (2 x CH₂S), 54.0 (C(CH₃)₃), 56.8 (CH₂N), 65.9 (OCH₂), 99.0 (CH), 106.5 (CH₃C=CH), 107.6 (Ar-CH), 135.3

(CCH₃=CH), 135.7, 140.3 (2 x Ar-CH), 150.1 (Ar-C), 160.0, 162.6, 163.1, 163.2, 163.4 (5 x C=O). MS, m/z = 388 (M⁺-H₂NBoc), 185, 127. Anal. Calcd for C₂₄H₃₁N₃O₇S: C, 57.02; H, 6.18; N, 8.31%. Found: C, 56.81; H, 6.40; N, 8.52%.

***N*-tert-Butoxycarbonyl-S-[2-(indol-1-yl)ethyl]cysteine ethyl ester (13c).** Prepared according to the method used for **13a** but allowing the mixture to warm to 20 °C over 10 min before the addition of *N*-(2-iodoethyl)indole **9c** (2.17 g, 8.01 mmol), and heating at reflux for 48 h. The title compound was isolated as a yellow oil (0.24 g, 31%). ¹H NMR (400 MHz, CDCl₃), δ_H 1.23 (3H, t, J = 7.1 Hz, CH₂CH₃), 1.36 (9H, s, C(CH₃)₃), 1.73, 2.92 (each 2H, m, CH₂S), 4.12 (2H, m, CH₂N), 4.14 (2H, q, J = 7.1 Hz, CH₂CH₃), 4.49 (1 H, m, CH), 5.29 (1H, br, NH), 7.09-7.26, 7.37-7.51, 7.58-7.63 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 14.1 (CH₂CH₃), 28.3 (C(CH₃)₃), 29.7 (CH₂S), 30.3 (C(CH₃)₃), 31.9 (CH₂S), 45.1 (CH₂N), 61.8 (OCH₂), 107.8 (CH), 126.4 (Ar-C), 128.4, 128.6, 129.2, 131.9, 132.2, 132.9 (6 x Ar-CH), 143.1 (Ar-C), 157.6, 179.9 (2 x C=O). MS, m/z = 347 (M⁺-OC₂H₅), 291, 249, 232, 203, 176, 146. HRMS: Calcd for C₂₀H₂₈N₂O₄S: [M-OC₂H₅]⁺ 347.1429; found: [M-OC₂H₅]⁺ 347.1431.

***N*-tert-Butoxycarbonyl-S-[2-(benzimidazol-1-yl)ethyl]cysteine ethyl ester (13d).** Prepared according to the method used for **13a** but allowing the mixture to warm to 20 °C over 20 min before the addition of *N*¹-(2-iodoethyl)benzimidazole **9d** (2.18 g, 8.01 mmol). The title compound was isolated as a yellow oil (0.44 g, 66%), IR (ν_{max}, NaCl/cm⁻¹) 3352, 3223, 2976, 1708, 1494, 1458, 1366, 1250, 1165. ¹H NMR (400 MHz, CDCl₃), δ_H 1.18 (3H, t, J = 7.2 Hz, CH₂CH₃), 1.36 (9H, s, C(CH₃)₃), 2.88 (2H, 2 x dd, J = 6.3, 6.6 Hz, CH₂S), 2.91 (2H, t, J = 7.0 Hz, SCH₂), 4.11 (2H, q, J = 7.2 Hz, CH₂CH₃), 4.29 (2H, t, J = 7.0 Hz, CH₂N), 4.41 (1H, m, CH), 5.23 (1H, br, NH), 7.20-7.75 (4H, m, Ar-H), 7.88 (1H, s, CH=N). ¹³C NMR (100 MHz, CDCl₃), δ_C 14.2 (CH₂CH₃), 28.3 (C(CH₃)₃), 32.3, 34.8 (2 x CH₂S), 42.9 (C(CH₃)₃), 44.9 (CH₂N), 61.9 (OCH₂), 109.4 (CH), 120.5, 122.4, 123.1, 128.6, 132.1 (5 x Ar-CH), 133.3, 143.7 (2 x Ar-C), 155.2, 170.7 (2 x C=O). MS, m/z = 394 (M⁺), 338, 279, 220, 192, 160, 146, 118. HRMS: Calcd for C₁₉H₂₇N₃O₄S: MH⁺ 394.1795; found: MH⁺ 394.1802. Anal. Calcd for C₁₉H₂₇N₃O₄S: C, 57.99; H, 6.92; N, 10.68%; Found C, 58.22; H, 6.80; N, 10.39%.

***N*-tert-Butoxycarbonyl-S-[2-(benzotriazol-1-yl)ethyl]cysteine ethyl ester (13e).** Prepared according to the method used for **13a** but using NaH (60% in mineral oil; 0.04 g, 1.00 mmol), and *N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.20 g, 0.80 mmol), and allowing the mixture to warm to 20 °C over 10 min before the addition of *N*¹-(2-iodoethyl)benzotriazole **9e** (0.88 g, 3.22 mmol). The title compound was isolated as a yellow oil (0.25 g, 75%), IR (ν_{max}, NaCl/cm⁻¹) 3444, 2915, 2065, 1634, 1495, 1393, 1247, 1162. ¹H NMR (400 MHz, CDCl₃), δ_H 1.24 (3H, t, J = 7.0 Hz, CH₂CH₃), 1.60 (9H, s, C(CH₃)₃), 3.09 (2H, br t, J = 6.9 Hz, CH₂S), 3.22 (2H, apparent t, J = 6.7 Hz, CH₂S), 4.08 (2H, q, J = 7.0 Hz, CH₂CH₃), 4.74 (2H, t, J = 6.9 Hz, CH₂N), 4.84 (1H, t, J = 6.7 Hz, CH), 5.23 (1H, br, NH), 7.26-7.48 (3H, m, Ar-H), 7.98 (1H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 12.3 (CH₂CH₃), 29.9 (C(CH₃)₃), 33.2, 35.2 (2 x CH₂S), 54.0 (CH₂N), 58.5 (OCH₂), 64.0 (C(CH₃)₃), 111.9 (CH), 128.6, 130.6, 131.1, 139.1 (4 x Ar-CH), 142.1, 146.9 (2 x Ar-C), 170.2, 184.6 (2 x C=O). MS, m/z = 395 (MH⁺), 360, 338, 294, 250, 191, 176. HRMS: Calcd for C₁₈H₂₆N₄O₄S: MH⁺ 395.1747; found: MH⁺ 395.1750.

***N*-tert-Butoxycarbonyl-S-[2-(benzotriazol-2-yl)ethyl]cysteine ethyl ester (13f).** Prepared according to the method used for **13a** but using NaH (60% in mineral oil; 0.04 g, 1.00 mmol) and *N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.20 g, 0.80 mmol), and allowing the mixture to warm to 20 °C over 10 min before the addition of *N*²-(2-iodoethyl)benzotriazole **9f** (0.88 g, 3.22 mmol). The title compound was isolated as a yellow oil (0.25 g, 74%), IR (ν_{\max} , NaCl/cm⁻¹) 3368, 2976, 1717, 1688, 1506, 1455, 1165, 1128. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.29 (3H, t, *J* = 7.1 Hz, CH₂CH₃), 1.47 (9H, s, C(CH₃)₃), 3.00 (2H, t, *J* = 6.9 Hz, CH₂S), 3.29 (2H, t, *J* = 7.1 Hz, CH₂S), 4.23 (2H, q, *J* = 7.1 Hz, CH₂CH₃), 4.55 (1H, m, CH), 4.93 (2H, t, *J* = 6.9 Hz, CH₂N), 5.44 (1H, br, NH), 7.40, 7.88 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 14.1 (CH₂CH₃), 28.3 (C(CH₃)₃), 32.1, 34.7 (2 x CH₂S), 44.9 (CH₂N), 55.8 (OCH₂), 61.8 (C(CH₃)₃), 118.0 (CH), 126.5, 133.3 (2 x Ar-CH), 144.4 (Ar-C), 177.9, 180.2 (2 x C=O). MS, *m/z* = 395 (MH⁺), 360, 338, 294, 250, 191, 176. HRMS: Calcd for C₁₈H₂₆N₄O₄S: MH⁺ 395.1747; found: MH⁺ 395.1751.

***N*-tert-Butoxycarbonyl-S-[2-(6-tert-butyloxycarbonylamino-purin-9-yl)ethyl]cysteine ethyl ester (13g).** Prepared according to the method used for **13a** but allowing the mixture to warm to 20 °C over 10 min before the addition of *N*⁹-(2-Iodoethyl)-6-*tert*-butoxycarbonylamino-purine **9g** (3.12 g, 8.02 mmol). The title compound was isolated as a yellow oil (0.42 g, 41%), IR (ν_{\max} , NaCl/cm⁻¹) 3419, 2979, 1738, 1701, 1657, 1455, 1367, 1252, 1175. ¹H NMR (250 MHz, CDCl₃), δ_{H} 1.14 (3H, t, *J* = 7.2 Hz, CH₂CH₃), 1.44, 1.47 (each 9H, s, C(CH₃)₃), 3.69-3.81 (4H, m, 2 x CH₂S), 4.14 (2H, q, *J* = 7.2 Hz, CH₂CH₃), 4.24 (2H, m, CH₂N), 6.00 (1H, m, CH), 6.48 (2H, br, NH), 7.50, 7.58 (each 1H, s, Ar-H); this material was not further characterized.

***N*-tert-Butoxycarbonyl-S-[3-(benzotriazol-1-yl)propyl]cysteine ethyl ester (13h).** Prepared according to the method used for **13a** but using NaH (60% in mineral oil; 0.04 g, 1.00 mmol) and *N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.20 g, 0.80 mmol), and allowing the mixture to warm to 20 °C over 10 min before the addition of *N*¹-(3-iodopropyl)benzotriazole **10e** (0.92 g, 3.21 mmol). The title compound was isolated as a yellow oil (0.20 g, 59%), IR (ν_{\max} , NaCl/cm⁻¹) 3447, 2990, 1636, 1556, 1432, 1102. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.22 (3H, m, CH₂CH₃), 1.43 (9H, s, C(CH₃)₃), 1.76 (2H, m, CH₂), 2.31, 2.74 (each 2H, t, *J* = 7.0 Hz, CH₂S), 4.22 (2H, m, CH₂CH₃), 4.66 (2H, t, *J* = 7.0 Hz, CH₂N), 5.23 (1H, m, CH), 5.97 (1H, br, NH), 7.31-7.47 (3H, m, Ar-H), 8.00 (1H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_{C} 12.5 (CH₃), 16.9 (CH₂), 26.1 (CH), 27.8 (C(CH₃)₃), 32.1, 37.1 (2 x CH₂S), 42.9 (CH₂N), 59.9 (OCH₂), 65.3 (C(CH₃)₃), 129.5, 130.2 (2 x Ar-CH), 131.3 (Ar-C), 135.4, 143.3 (2 x Ar-CH), 150.0 (Ar-C), 162.9, 170.0 (2 x C=O). MS, *m/z* = 408 (M⁺), 322, 250, 214, 198, 158, 144, 130, 118. HRMS: Calcd for C₁₉H₂₈N₄O₄S: M⁺ 408.1831; found: M⁺ 408.1832.

***N*-tert-Butoxycarbonyl-S-[3-(benzotriazol-2-yl)propyl]cysteine ethyl ester (13i).** Prepared according to the method used for **13a** but using NaH (60% in mineral oil; 0.04 g, 1.00 mmol) and *N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.20 g, 0.80 mmol), and allowing the mixture to warm to 20 °C over 10 min before the addition of *N*²-(3-iodopropyl)benzotriazole **10f** (0.92 g, 3.21 mmol). The title compound was isolated as a yellow oil (0.235g, 68%), IR (ν_{\max} , NaCl/cm⁻¹) 3447, 2987, 1653, 1559, 1457, 1372, 1164. ¹H NMR (400 MHz, CDCl₃), δ_{H} 1.24 (3H, t, *J* =

7.1 Hz, CH₂CH₃), 1.38 (9H, s, C(CH₃)₃), 1.70 (2H, m, CH₂), 2.50 (2H, t, *J* = 6.5 Hz, CH₂S), 3.09 (2H, t, *J* = 6.6 Hz, CH₂S), 4.12 (2H, q, *J* = 7.1 Hz, CH₂CH₃), 4.70 (2H, t, *J* = 6.5 Hz, CH₂N), 4.78 (1H, m, CH), 5.00 (1H, br, NH), 7.32, 8.00 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 12.3 (CH₃), 25.6 (C(CH₃)₃), 26.5 (CH), 27.9 (CH₂), 28.5 (CH₂N), 31.4 (CH₂S), 35.4 (C(CH₃)₃), 46.2 (CH₂S), 54.1 (OCH₂), 118.3 (Ar-C), 122.3, 125.8 (2 x Ar-CH), 164.8, 177.1 (2 x C=O). MS, *m/z* = 409 (MH⁺), 322, 250, 214, 184, 158, 144, 130, 118. HRMS: Calcd for C₁₉H₂₈N₄O₄S: MH⁺ 409.1904; found: MH⁺ 409.1913.

***N*-tert-Butoxycarbonyl-S-[4-(benzotriazol-1-yl)butyl]cysteine ethyl ester (13j).** Prepared according to the method used for **13a** but using NaH (60% in mineral oil; 0.04 g, 1.00 mmol) and *N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.20 g, 0.80 mmol), and allowing the mixture to warm to 20 °C over 10 min before the addition of *N*¹-(4-iodobutyl)benzotriazole **11e** (0.97 g, 3.22 mmol). The title compound was isolated as a yellow oil (0.13 g, 39%), IR (ν_{max}, NaCl/cm⁻¹) 3425, 2928, 1708, 1495, 1367, 1280, 1165. ¹H NMR (400 MHz, CDCl₃), δ_H 1.24 (3H, t, *J* = 7.2 Hz, CH₂CH₃), 1.43 (9H, s, C(CH₃)₃), 1.68, 2.17 (each 2H, m, CH₂), 2.58 (2H, t, *J* = 7.3 Hz, CH₂S), 2.92 (2H, t, *J* = 4.9 Hz, CH₂S), 4.18 (2H, q, *J* = 7.2 Hz CH₂CH₃), 4.67 (3H, m, CH₂N, CH), 5.34 (1H, br, NH), 7.37-7.53 (2H, m, Ar-H), 7.85, 8.04 (each 1H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 11.8 (CH₃), 19.6, 20.0 (2 x CH₂), 25.2 (C(CH₃)₃), 28.2 (CH), 31.2, 35.4 (2 x CH₂S), 46.8 (CH₂N), 66.3 (OCH₂), 67.2 (C(CH₃)₃), 133.4, 133.9 (2 x Ar-CH), 134.1, 140.7 (2 x Ar-CH), 146.7, 150.0 (2 x Ar-C), 168.8, 176.5 (2 x C=O). MS, *m/z* = 422 (M⁺), 340, 214, 198, 173, 158, 130. HRMS: Calcd for C₂₀H₃₀N₄O₄S requires: M⁺ 422.1988; found: M⁺ 422.2013.

***N*-tert-Butoxycarbonyl-S-[4-(benzotriazol-2-yl)butyl]cysteine ethyl ester (13k).** Prepared according to the method used for **13a** but using NaH (60% in mineral oil; 0.04 g, 1.00 mmol) and *N*-tert-butoxycarbonyl-(*R*)-cysteine ethyl ester **5** (0.20 g, 0.80 mmol), and allowing the mixture to warm to 20 °C over 10 min before the addition of *N*²-(4-iodobutyl)benzotriazole **11f** (0.97 g, 3.22 mmol). The title compound was isolated as a yellow oil (0.20 g, 61%), IR (ν_{max}, NaCl/cm⁻¹) 3418, 2928, 1710, 1510, 1321, 1157. ¹H NMR (400 MHz, CDCl₃), δ_H 1.23 (3H, t, *J* = 7.1 Hz, CH₂CH₃), 1.43 (9H, s, C(CH₃)₃), 1.66, 2.21 (each 2H, m, CH₂), 2.49 (2H, t, *J* = 7.2 Hz, CH₂S), 2.58 (2H, m, CH₂S), 4.15 (2H, q, *J* = 7.1 Hz, CH₂CH₃), 4.47 (1H, m, CH), 4.73 (2H, t, *J* = 6.2 Hz, CH₂N), 5.28 (1H, br, NH), 7.35-7.84 (each 2H, m, Ar-H). ¹³C NMR (100 MHz, CDCl₃), δ_C 9.9 (CH₃), 24.9 (C(CH₃)₃), 27.0, 29.1 (2 x CH₂), 30.3 (CH), 31.0, 40.2 (2 x CH₂S), 45.4 (CH₂N), 63.2 (OCH₂), 70.6 (C(CH₃)₃), 122.7, 126.5 (2 x Ar-CH), 133.1 (Ar-C), 171.4, 173.6 (2 x C=O). MS, *m/z* = 423 (MH⁺), 340, 214, 184, 173, 158, 130. HRMS: Calcd for C₂₀H₃₀N₄O₄S: MH⁺ 423.2060; found: MH⁺ 423.2063.

***N*-tert-Butoxycarbonyl-S-[2-(benzimidazol-1-yl)ethyl]cysteine (14)**

N-tert-Butoxycarbonyl-S-[2-(benzimidazol-1-yl)ethyl]cysteine ethyl ester **13d** (0.5 g, 1.28 mmol) in THF:methanol (50 mL, 3:1 v/v) was cooled in an ice bath, then LiOH solution (1M; 250 mL) was then added, the mixture allowed to warm to 20 °C and stirred for a further 8 h. The organic solvents were removed under reduced pressure, the water layer acidified by hydrochloric acid (2M) to pH 4 and the product extracted with EtOAc (3 x 20 mL). The combined extracts

were dried (MgSO_4) and the solvent removed under reduced pressure to leave the title compound as a white solid (0.45 g, 97%), (ν_{max} , $\text{NaCl}/\text{cm}^{-1}$) 3412 (br), 2925, 1711, 1514, 1368, 1256, 1160. ^1H NMR (250 MHz, CDCl_3), δ_{H} 1.43 (9H, s, $\text{C}(\text{CH}_3)_3$), 3.04-3.08 (4H, m, CH_2SCH_2), 4.52 (1H, m, CH), 4.67 (2H, t, $J = 7.2$ Hz, CH_2N), 5.64 (1H, m, NH), 7.50-7.52 (4H, m, Ar-H), 7.96 (1H, s, $\text{CH}=\text{N}$). HRMS: Calcd for $\text{C}_{17}\text{H}_{23}\text{N}_3\text{O}_4\text{S}$ requires: MH^+ 366.1482; found: MH^+ 366.1458.

S-[2-(Benzimidazol-1-yl)ethyl]cysteine ethyl ester (15)

N-tert-Butoxycarbonyl-S-[2-(benzimidazol-1-yl)ethyl]cysteine ethyl ester **13d** (0.5 g, 1.28 mmol) was added to TFA: CH_2Cl_2 (50 mL, 1:1 v/v) and the mixture stirred for 2 h. Diisopropylethylamine: CH_2Cl_2 (50 mL, 5:95 v/v) was then added to neutralise the acid. The solvents were removed under reduced pressure to give the crude amine that was chromatographed through a short silica column to give the title compound (0.37g, 86%), IR (ν_{max} , $\text{NaCl}/\text{cm}^{-1}$) 3415, 2981, 1766, 1475, 1371. ^1H NMR (250 MHz, CDCl_3), δ_{H} 1.22 (3H, t, $J = 7.1$ Hz, CH_2CH_3), 3.07-3.09 (4H, m, CH_2SCH_2), 4.18 (2H, q, $J = 7.1$ Hz, CH_2CH_3) 4.35 (1H, t, $J = 5.6$ Hz, CH), 4.66 (1H, t, $J = 6.5$ Hz, CH_2N), 6.11 (2H, br, NH_2), 7.49-7.63 (4H, m, Ar-H), 7.95 (1H, s, $\text{CH}=\text{N}$). HRMS: Calcd for $\text{C}_{14}\text{H}_{19}\text{N}_3\text{O}_2\text{S}$: MH^+ 294.1271; found: MH^+ 294.1247.

Acknowledgements

We acknowledge a studentship from Loughborough University (L. J. C.), and some MS data from the National Mass Spectrometry Service Centre, Swansea University.

References and Notes

1. Jones, R. C. F.; Berthelot, D. J. C.; Iley, J. N. *Tetrahedron* **2001**, *57*, 6539.
2. Jones, R. C. F.; Berthelot D. J. C.; Iley, J. N. *Arkivoc* **2007**, (xi), 73.
3. Howarth, N. M.; Wakelin, L. P. G. *J. Org. Chem.* **1997**, *62*, 5441.
4. For an introduction to PNAs, see for example: *Peptide Nucleic Acids: Protocols and Applications*; Nielsen, P. E. Ed; Horizon Bioscience: Wymondham, 2004.
5. See, for example: Adlington, R. M.; Baldwin, J. E.; Catterick, D.; Pritchard, G. J. *J. Chem. Soc., Perkin Trans. 1* **1999**, 855.
6. For O-linked (serine-based) heterocyclic amino acids assembled via a different strategy, see, for example: Garner, P.; Dey, S.; Huang, Y. in *Pseudo-Peptides in Drug Discovery*; Nielsen, P. E. Ed; Wiley-VCH: Weinheim, 2004; p 193.
7. For heterarylalanine derivatives from serine via Mitsunobu coupling to nucleobases, see: Guo, H.-M.; Wu, Y.-Y.; Niu, H.-Y.; Wang, D.-C.; Qu, G.-R. *J. Org. Chem.* **2010**, *75*, 3863.
8. Compound **12a** showed just two methylene signals in the ^1H NMR spectrum, integral ratio 2:1 with the more intense signal as a downfield shifted triplet, δ_{H} 4.90, suggesting a

- symmetrical (quaternary salt) structure. Likewise, compound **12b** showed two methylene signals, ratio 1:1, with one downfield triplet, δ_{H} 4.57. In both cases, MS data indicated a dimeric salt with ions observed corresponding to M^{2+} and to $[\text{M}-\text{H}]^+$.
9. For a related example of a double salt from alkylation of benzimidazole with a 1,3-dihalopropane, see: Murphy, J. A.; Khan, T. A.; Zhou, S.-z.; Thomson, D. W.; Mahesh, M. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 1356.
 10. Sugano, H.; Miyoshi, M. *J. Org. Chem.* **1976**, *41*, 2352.
 11. Commercial *N-tert*-butoxycarbonyl-(*RS*)-homocysteine methyl ester and *N*¹-(2-iodoethyl)benzimidazole **9d** reacted under the NaH protocol to give the corresponding amino acid derivative in 57% yield.
 12. Cruikshank, K. A.; Jiricny, J.; Reese, C. B. *Tetrahedron Lett.* **1984**, *25*, 681.
 13. Díez-Barra, E.; de la Hoz, A.; Sánchez-Migallón, A.; Tejada, J. *Heterocycles* **1992**, *34*, 1365.
 14. Märky, M.; Schmid, H.; Hansen, H. J. *Helv. Chim. Acta* **1979**, *62*, 2129.
 15. Krollpfeiffer, F.; Pötz, H.; Rosenberg, A. *Chem. Ber.* **1938**, *71*, 596.
 16. Hirschberg, E.; Gellhorn, A.; Gump, W. S. *Cancer Res.* **1957**, *17*, 904.
 17. Davies, R. G.; Gibson, V. C.; Hursthouse, M. B.; Light, M. E.; Marshall, E. L.; North, M.; Robson, D. A.; Thompson, I.; White, A. J. P.; Williams, D. J.; Williams, P. J. *J. Chem. Soc., Perkin Trans. 1* **2001**, 3365.
 18. Shimidzu, T.; Kanou, T.; Murakami, A.; Yamana, K.; Inagaki, H.; Donkai, N. *J. Chem. Research (S)* **1980**, 376, (*M*) 4668.