Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2019, **17**, 354

A diversity of alkylation/acylation products of uracil and its derivatives: synthesis and a structural study†

Olga Michalak, *[a] Piotr Cmoch,[b] Piotr Krzeczyński,[a] Marcin Cybulski[a] and Andrzej Leś[a,c]

Received 15th October 2018,
Accepted 10th December 2018

DOI: 10.1039/c8ob02552e

rsc.li/obc

tert-Butyl dicarbonate (Boc_2O) and ethyl iodide (EtI) reactions with uracil (U), thymine (T) and 6-methyluracil (6-MU) were performed following routine procedures in pyridine/DMF solvents and with DMAP as the catalyst. Among 20 synthesized compounds, a derivative of 6-methyluracil substituted by the Boc-pyridine moiety at the C5 position appeared unexpectedly. The NMR spectra confirmed the molecular structure of all uracil derivatives. Parallel quantum mechanical DFT calculations supported the experimental findings.

Introduction

The uracil moiety is one of the most important structures encountered in life sciences. As nucleobases, uracil (U) and thymine (T) occur in the RNA and DNA of nucleic acids and other natural products, *i.e.* willardiine, sparsomycin, and polyoxins (Fig. 1).[1] Various derivatives of the uracil family serve as the building blocks in the synthesis of nucleoside analogues with potential antiviral, anti-inflammatory or anticancer properties.[2]

For an effective synthetic strategy, protection of the N^1- and/or N^3-amine position of the uracil core by various groups, such as SEM,[3] acyl,[4] benzyl,[5] benzoyl,[6] benzyloxymethyl,[7] tetrahydropyranyl,[8] can be achieved. While working on our recent research project, we selected the Boc-protecting group as the most efficient for the thymine and 6-methyluracil intermediates, as it was stable in comparison with most nucleophiles and bases[9] and could be removed under neutral conditions in a clean and selective manner.[10]

Surprisingly, only three teams have so far described the synthesis of the Boc-protected thymine under basic conditions. Jaime-Figueroa *et al.*[11] obtained an N^1-Boc derivative using Boc_2O at a stoichiometric ratio in relation to thymine, while Gothelf *et al.*[12] obtained N^3-Boc protected thymine by the selective removal of N^1-Boc from the previously synthesized N^1,N^3-di-Boc derivative. This reaction was carried out in the presence of K_2CO_3 in dioxane with 31% yield. Bessières *et al.*[13] developed the synthesis of the Boc derivative at the N^3-position of thymine with a high yield, comprising two phases: complete pyrimidine protection by the microwave radiation, followed by the selective N^1 deprotection by SiO_2 in dimethoxymethane/ethanol (9 : 1).

N^1-Boc protected uracil was synthesized by Jaime-Figueroa *et al.*[11] according to the procedure developed for thymine, using 1.0 eq. Boc_2O. Likewise, Lipani *et al.*[14] obtained N^1-Boc uracil by the regioselective derivatization, using 1.1 eq. Boc_2O and pyridine as the solvent.

There are no available literature data for the Boc protection of 6-methyluracil. However, Lee[15] obtained 6-methyl-2,4-pyrimidyldiesters as a result of the 2,4-dihydroxy-6-methylpyrimidine reaction with the acyl chlorides in dichloromethane in the presence of TEA at room temperature. Active esters, *i.e.* O^2- and O^4-acetylation products, were obtained with good yields and they were thermally stable.

Due to the incomplete and inconsistent data concerning the Boc protection of uracil, thymine and 6-methyluracil, we

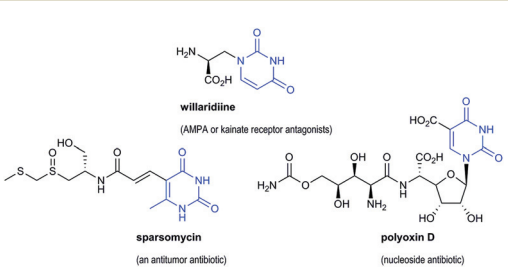


Fig. 1 Bioactive molecules based on the uracil moiety.

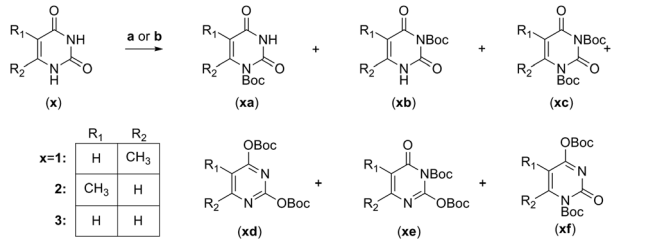
[a]*Pharmaceutical Research Institute, 8 Rydygiera Str., 01-793 Warsaw, Poland. E-mail: o.michalak@ifarm.eu; Tel: +48 22 456 3925*
[b]*Institute of Organic Chemistry Polish Academy of Sciences, 44/52 Kasprzaka street, 01-224 Warsaw, Poland*
[c]*University of Warsaw, Faculty of Chemistry, 1 Pasteura Str., 02-093 Warsaw, Poland*
† Electronic supplementary information (ESI) available: DFT data, comparison of DFT and experimental results. See DOI: 10.1039/c8ob02552e

have decided to validate the hypothesis of different chemical affinities of these structurally similar nucleobases to the acylation reagent. As a continuation of this effort, we have also tested the course of the alkylation reaction with EtI by comprehensively analyzing the separated products. All the obtained compounds were isolated and fully characterized by various NMR and HRMS techniques. Numerous quantum mechanical calculations within the density functional theory (DFT/B3LYP) were performed to support the experimental results and discuss product structures.

Results and discussion

Chemistry and NMR spectroscopic analysis

Acylation of uracil (U), thymine (T) and 6-methyluracil (6-MU) with Boc_2O. According to the literature, the reaction of thymine with Boc_2O in the presence of DMAP in acetonitrile leads to only one product: N^1-Boc-thymine (**2a**) with a good yield,[11] see Scheme 1.

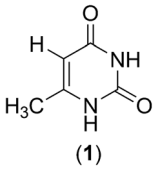
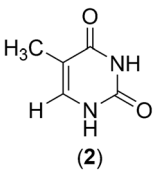
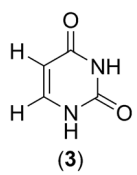


Scheme 1 Reagents and conditions: (a) Boc_2O (3 eq.), DMAP, MeCN, r.t., (b) Boc_2O (3 eq.), pyridine, MeCN, temp. 55 °C.

We have decided to apply this procedure to obtain the N^1-Boc protected analogue of other pyrimidine bases, *i.e.* 6-methyluracil, using 1.5 mol eq. Boc_2O. The result was unexpected. Our experiments afforded not one, but four different products, see Table 1.

The compounds were isolated by column chromatography and their structures were characterized by the NMR data to determine the acylation position. It turned out that one of the unexpectedly identified products was 1,3-di-*O-tert*-butyloxycarbonyl-(1*H*,3*H*)-6-methylpyrimidine (**1d**). A careful analysis of the NMR data for **1d**, especially ^{15}N NMR chemical shifts, led to the confirmation of this structure. The comparison of ^{15}N NMR data for **1d** with the data collected for the unprotected 6-methyluracil showed an essential change in the nitrogen nuclei character of diester **1d**. In **1**, two nitrogen nuclei are of the "amide" type, whereas in **1d** both nitrogens are of the "pyridine" type. This transformation of the nitrogen character was related to the following changes: nitrogen/carbon atom hybridization and the aromatization of the ring, resulting in the nitrogen/carbon shielding changes. Other reaction products, identified on the basis of the analysis of the ^{1}H-^{13}C/^{15}N HMBC spectra, were as follows: N^1/N^3-di-Boc derivative (**1c**) and two mono–Boc derivatives at the N^1 or N^3 positions (**1a** and **1b**) (Scheme 1). The comparison of the ^{13}C and ^{15}N NMR chemical shifts between **1** and **1a**/**1b**/**1c** did not reveal such spectacular differences as in the case of **1**/**1d**, but several significant effects

Table 1 Distribution of the products in the alkylation and acylation reactions of the uracil family under different conditions

		Acylation conditions				Alkylation conditions[c]		
		a[a]		**b**[b]				
Nucleobase substrate	Product	1.5 eq. Boc_2O	3 eq. Boc_2O	3 eq. Boc_2O	Product	1.0 eq. EtI	1.5 eq. EtI	3 eq. EtI
 (**1**)	**1a**	5[d]	3	9	**1h**	—	—	—
	1b	5	3	15	**1i**	17[d]	5	—
	1c	5	11	20	**1j**	30	75	99
	1d	12	32	—	**1k**	—	—	—
	1e	—	—	—	**1l**	—	3	—
	1g	—	—	20				
 (**2**)	**2a**	33	59	—	**2h**	19	22	33
	2b	—	—	4	**2i**	—	—	—
	2c	—	39	89	**2j**	26	30	32
 (**3**)	**3a**	—	12	—	**3h**	—	12	16
	3b	—	12	23	**3i**	—	1	1
	3c	—	42	77	**3j**	—	31	29
	3f	—	—	—	**3m**	—	1	1

[a] **a**: Boc_2O (3 eq.), DMAP, MeCN, r.t;. [b] **b**: Boc_2O (3 eq.), pyridine, MeCN, temp. 55 °C. [c] EtI, K_2CO_3, TBABr, DMF, r.t. [d] Yields [%].

were also noticed. An important one refers to the change of the ^{15}N shielding after the replacement of hydrogen by the Boc group at the nitrogen atom. When both N^1 and N^3 protons were replaced with the Boc group (**1c**), the ^{15}N shielding decreased by *ca.* 30 ppm, respectively (Table 2).

After proton N^3 or N^1 substitution with Boc, a decrease in the ^{15}N shielding was still evident (*ca.* 30 ppm), while the unsubstituted nitrogen nucleus was only insignificantly shielded by *ca.* 2–3 ppm. Additionally, the above mentioned replacement caused small, but noteworthy ^{13}C shielding effects. *ortho*-Carbon nuclei relative to nitrogens always experienced a shielding increase by *ca.* 2–3 ppm (Table 2).

These interesting results prompted us to investigate in greater detail the regioselectivity of the acylation of uracil and its isomeric methyl derivatives: thymine and 6-methyluracil. We began by examining the effects of the Boc_2O reagent excess on the reaction product specificity. Then, the influence of other reaction conditions, such as the pyridine addition or the increase of the temperature on the reaction's regioselectivity, was studied.

In the first type of the experiment (path a, Scheme 1) thymine, 6-methyluracil and uracil were reacted with 1.5 or 3 eq. of Boc_2O in the presence of DMAP in acetonitrile. As a result, two products were obtained and isolated for thymine. Based on the analysis of the multinuclear NMR data, especially the $^1H/^{13}C$ HMBC experiments, the structures of these products were confirmed. The first was Boc-di-substituted thymine at the positions N^1 and N^3 (**2c**) and the second one was mono-substituted thymine at the N^3 position (**2a**) (see Table 2). In the case of uracil, like for thymine, the di-substituted compound at the N^1/N^3 position (**3c**) was the main reaction product. Moreover, apart from the Boc-mono-substituted product at the N^3 position (**3a**), a similar amount of the N^1-Boc product was formed (**3b**). The ^1H, ^{13}C and ^{15}N NMR data for the Boc-substituted-6-methyluracil are presented in Table 2. For 6-methyluracil, the main product proved to be di-substituted at the O^2/O^4 di-*O*-*tert*-butyloxycarbonyl-pyrimidine derivative (**1d**).

Additionally, after the reaction of Boc_2O with 6-methyluracil three other products were found: di-N^1/N^3-Boc (**1c**) and small quantities of the mono-substituted N^1-Boc and N^3-Boc (**1a**, **1b**), see Table 2.

While changing the conditions of the acylation by replacing DMAP with pyridine and increasing the temperature to 55 °C (path b), practically only one di-N^1/N^3-Boc product (**2c**) was obtained for thymine. This result was consistent with literature data;[12] however, a small amount of an N^3-substituted product (**2b**) was also isolated. For uracil, two products: di-N^1/N^3-Boc (**3c**, yield 77%) and N^3-Boc (**3b**, yield 23%) were obtained. In the case of 6-methyluracil, three previously described products were obtained again: **1a**/**1b**/**1c**. In contrast to the reaction with DMAP (path a), no *O*-acylated derivatives of 6-methyluracil (at the O^2 and O^4 positions) were observed under these conditions.

Unexpected product of the 6-methyluracil (6-MU) acylation with Boc_2O and the mechanism of its formation. During the

Table 2 The ^1H, ^{13}C and ^{15}N NMR chemical shifts measured in $CDCl_3$ solutions at T = 298 K

	1 [a]	**1a**	**1b**	**1c**	**1d**	**2** [a]	**2a**	**2b**	**2c**	**3** [a]	**3a** [b]	**3b** [b]	**3c** [b]
N1/H1	−243.5/10.48	−215.4	−246.8/9.82	−216.4	−114.8	−252.2/10.16	−225.6	−255.6/9.95	−226.5	−249.7/10.40	−223.9	−253.4/10.04	−224.8
C2	151.6	148.9[d]	150.5	147.8[d]	160.0	151.4	148.5[d]	150.7	145.7[d]	151.3	147.1	150.5	148.0[d]
N3/H3	−225.0/9.58	−227.3/8.73	−193.1	−194.0	−136.3	−223.5/9.98	−225.1/8.67	−191.6	−185[c]	−221.3/10.16	−223.6/8.76	−189.7	−190.0
C4	164.3	162.2	160.8	159.5	166.7	164.9	163.3	161.9	161.1	164.3	162.5	161.0	159.9
C5	—	—	—	—	—	108.3	112.0	110.5	111.7	—	—	—	—
C5/H5	99.1/5.21	101.6/5.55	100.3/5.56	101.2/5.54	108.7/7.01	—	—	—	—	100.4/5.28	103.7/5.80 [8.5]	102.0/5.76 [7.9]	103.4/5.81 [8.6]
C6	152.4	149.7	151.3	148.9	173.4	—	—	—	—		—	—	—
C6/H6	—	—	—	—	—	136.5/6.74	135.3/7.71	135.9/7.04	134.6/7.68	140.6/6.87	139.8/7.88 [8.5]	140.1/7.22 [7.9]	138.9/7.87 [8.6]
CH_3 at C5	—	—	—	—	—	11.6/1.61	12.4/1.96	12.3/1.91	12.5/1.95	—	—	—	—
CH_3 at C6	18.4/1.94	18.6/2.21	18.9/2.14	18.6/2.18	24.1/2.56	—	—	—	—	—	—	—	—
$\underline{C}OO$–$C(CH_3)_3$	—	148.2[d]	147.6	147.3[d]||147.1[d]	149.4[d]||148.7[d]	—	147.3[d]	147.8	148.4[d]||147.4[d]	—	148.1	147.6	147.2[d]||145.5[d]
$CO\underline{O}$–$\underline{C}(CH_3)_3$	—	87.7	86.7	88.1[d]||87.1[d]	85.4[d]||84.7[d]	—	86.9	87.0	87.3[d]||87.0[d]	—	87.4	87.1	87.1[d]||87.6[d]
COO–$C(\underline{C}H_3)_3$	—	27.4/1.60	27.4/1.60	27.4[d]/1.60[d]|| 27.3[d]/1.59[d]	27.6[d]/1.54[d]|| 27.5[d]/1.55[d]	—	27.8/1.61	27.5/1.60	27.8[d]/1.59[d]|| 27.4[d]/1.59[d]	—	27.8/1.61	27.5/1.61	27.7[d]/1.61[d]|| 27.4[d]/1.60[d]

[a] The $^1H/^{13}C$ and ^{15}N chemical shifts measured for **1**, **2** and **3** in the $CDCl_3$/DMSO solution (10 : 1). [b] In square brackets the ^3J(H–H) spin–spin coupling constants, in Hz. [c] Predicted nitrogen signals (based on the linear regression line) of all compounds presented in Table 2, unrecorded in the ^1H–^{15}N HMBC experiment. [d] Assignments of the carbon/proton signals of the $\underline{C}OO$–$C(CH_3)_3$, $CO\underline{O}$–$\underline{C}(CH_3)_3$ COO–$C(\underline{C}H_3)_3$ groups and C2 atom for individual/respective compounds could be reversed.

column chromatography purification of the 6-methyluracil reaction mixture, a new entity was isolated. The NMR spectra of this compound (**1g**) did not correspond to those of any previously identified Boc-6-methyluracil derivatives. Very broad signals at δ *ca.* 4.5 and 4.7 ppm were observed at ambient temperature in the ^{1}H NMR spectrum. Lowering the temperature to −20 °C sharpened the ^{1}H NMR signals and allowed full analysis of the NMR spectra and assignment of the structure of this new compound. In the ^{1}H NMR spectrum of **1g** recorded at low temperature there was no H5 signal at δ *ca.* 5.7 ppm which had been typically detected for uracils (Table 2). Moreover, the correlation spot in the 2D ^{1}H–^{13}C HSQC spectrum for this C5–H5 pair was not observed in the low temperature spectrum. Instead, five "aromatic/double bond" proton signals, which are mutually coupled in the COSY spectrum, were noticed. The ^{1}H NMR signal at δ = 4.77 ppm in the 2D ^{1}H–^{13}C HMBC spectrum correlates with five carbon atoms at δ = 161.1, 151.7, 123.4, 122.9 and 113.8 ppm, respectively. As it was proved by analyzing the ^{1}H–^{13}C HMBC correlation for the methyl group at δ = 2.24 ppm, the ^{13}C NMR chemical shifts (161.1, 151.7 and 113.8 ppm) are related to the 6-methyluracil ring and this is why two remaining δ_C values at 123.4 and 122.9 ppm are related to the pyridine ring. Additionally, the 2D ^{1}H–^{15}N HMBC experiment revealed correlation spots related to three ^{15}N NMR chemical shifts of pyrimidine (δ = −243.3 and −194.4 ppm) and pyridine rings. The additional nitrogen signal at δ = −254.2 ppm comes from the pyridine ring as evidenced by the ^{1}H–^{15}N correlations of the pyridine protons at δ = 4.64, 4.74, 6.76 and 6.91 ppm (Table 3). This allowed us to state that the compound **1g** was not the product of a simple *N*- or *O*-acylation, but of another reaction mechanism. The NMR studies, including theoretical DFT calculations, suggested the chemical structure of an unknown **1g** compound as shown in Table 3.

Results of the modelling of the **1g** formation with the DFT B3LYP/6-31G(d,p) calculations are given in the ESI.†

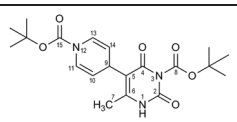
Although the details of the mechanism of the **1g** formation were unknown, it is reasonable to assume that it may be analogous to the known reactions of the structurally similar compounds. It is possible that the reaction follows a route comparable to the coupling of an indole with Boc-pyridine with the consecutive proton elimination from the indole ring.[16] A plausible pathway for the **1g** formation is presented in Scheme 2. First, the pyridine reaction with Boc_2O afforded the salt **5**. This compound was coupled with N^3-Boc-6-methyluracil (**1b**), previously formed in the reaction environment.

This process led to the intermediate complex **6** which was transformed to **1g** through the removal of the C5 proton from the pyrimidine ring by the Boc anionic form. The modeling of the **1g** formation with the DFT B3LYP/6-31G(d,p) calculations predicts the final output of the reaction free energy of about 14 kcal mol^{-1} (Table S9, ESI†).

The molecular model of **1g** is presented in Fig. 2.

A rough estimation of the top of the energy barrier corresponding to the transition structure **6** is about 63 kcal mol^{-1} (the B3LYP/6-31G(d) calculations). This result was obtained under the assumption that $(Boc)_2O$ acts as an acylation

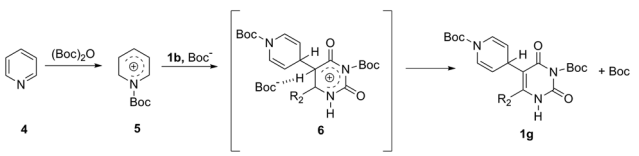
Table 3 The NMR chemical shifts for compound **1g**



1g

Atom	δ_{ppm} experimental	δ_{ppm} DFT
H1 (N1)	10.81 (−243.3)	6.69 (−258.4)
C2	147.6	154.9
N3	−194.4	−199.9
C4	161.1	168.5
C5	113.8	118.2
C6	151.7	159.0
H(CH_3) at C6 (C7)	2.24 (16.7)	2.19 (21.6)
C8	150.0	159.9
H9 (C9)	4.77 (27.8)	4.28 (39.7)
H10 (C10)	4.74 (106.3)*	4.91 (114.6)
H11 (C11)	6.91 (122.9)**	6.93 (130.9)
N12	−254.2	−259.5
H13 (C13)	6.76 (123.4)**	6.86 (131.5)
H14 (C14)	4.64 (105.6)*	4.78 (113.7)
C15	149.9	157.3
C_{IV}(N12)/C_{IV}(N3)	82.4/87.0	92.2/97.0
CH_3(C_{IV}N12)/CH_3(C_{IV}N3)	27.8/28.0	28.3/28.9

*, ** – ^{1}H/^{13}C NMR signals can be assigned inversely.



4 5 6 1g

Scheme 2 A plausible pathway for the formation of **1g**.

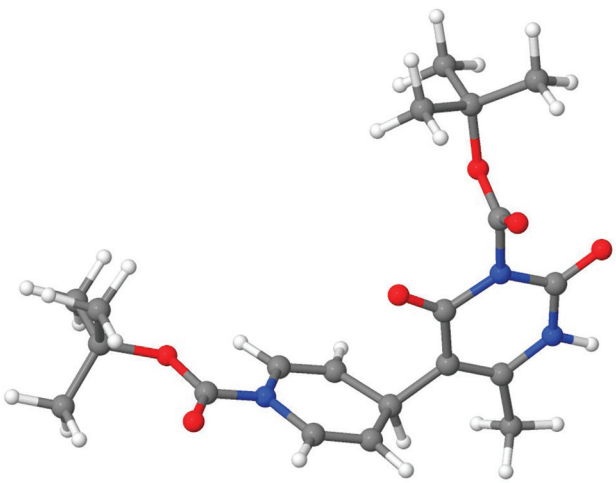


Fig. 2 The molecular model of **1g** obtained with the theoretical DFT/6-311G(d) calculations.

reagent and simultaneously as the medium, *i.e.* Boc^-, abstracting the hydrogen atom from the pyrimidine ring.

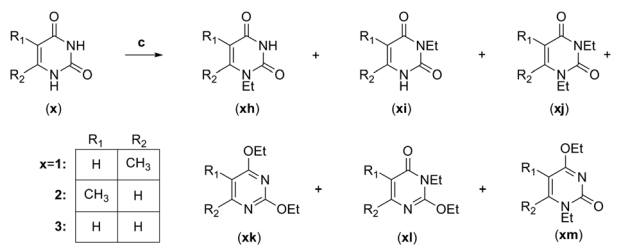
The molecular mechanism of the acylation reactions for other uracil derivatives under our complex and multi-component conditions (DMAP, MeCN, pyridine, variable concentration of $(Boc)_2O$, temperature r.t. and 55 °C) demands an extended study which is considered as the next step of our investigation. As long as the details of the molecular mechanism are not determined, it is hard to predict the corresponding kinetic barriers and provide a reasonable explanation of a diversity in the yield of the uracil derivatives obtained here.

Alkylation of uracil, thymine and 6-methyluracil with EtI. A variety of derivatives obtained in the course of pyrimidine acylation with Boc_2O encouraged us to test the regioselectivity of the ethylation reaction with EtI. The ethylation reactions of thymine, uracil and 6-methyluracil were carried out in a polar aprotic solvent (DMF) in the presence of K_2CO_3 and catalytic amounts of tetrabutylammonium bromide (TBABr). All products were identified in the course of 2D NMR experiments including ^1H–^{13}C and ^1H–^{15}N HSQC and HMBC correlations. The use of primary halogenoalkane resulted in the monoalkylation of thymine and uracil at the position N^1 (**2h**/**3h**) and dialkylation at N^1/N^3 atoms (**2j**/**3j**),[17,18] see Scheme 3.

In the case of uracil we also observed the formation of small quantities of two products: N^1/O^4- (**3m**) and N^3- (**3i**). This was consistent with the observations of Gambacorta *et al.*[19] who had also described the formation of the N^1/O^4-uracil product while studying the ethylation reaction of the lithium or potassium salts of 4-methoxy-2(1*H*)-pyrimidinone conjugated bases. However, these N^1/O^4-derivatives had not been completely characterized in the quoted paper.

For 6-methyluracil, a small amount of the O^2- and N^3-dialkylation product **1l** was identified in addition to the N^3/N^1-dialkylation (**1j**) and N^3-monoalkylation (**1i**) products. In contrast to the observations of Gambacorta *et al.*,[19] who had identified small quantities of the N^1/O^4-diethyl derivative of thymine after the ethylation procedure, we did not observe any thymine *O*-alkylation in our experiments.

The NMR studies showed that the replacement of the proton with the ethyl group in compounds **1–3** caused minor changes in the ^{15}N shielding in comparison with the effect observed for the corresponding Boc analogues. The difference was smaller than 30 ppm and depended on the position of the substitution (Table 4). In the dialkylated compounds, the ^{15}N



Scheme 3 Reagents and conditions: EtI, K_2CO_3, TBABr, DMF, r.t.

Table 4 The ^1H, ^{13}C and ^{15}N NMR chemical shifts measured in the $CDCl_3$ solutions at T = 298 K

	1 [a]	**1i**	**1j**	**1l**	**2** [a]	**2h**	**2j**	**3** [a]	**3h** [b]	**3i** [b]	**3j** [b]	**3m** [b]
N1/H1	−243.5/ 10.48	−246.0/9.83	−237.7	−179.0	−252.2/ 10.16	−244.3	−245.5	−249.7/ 10.40	−241.2	−253.4/9.78	−242.3	−224.4
C2	151.6	152.9	151.6	155.6	151.4	150.8	151.0	151.3	150.9	152.8	151.0	156.5
N3/H3	−225.0/9.58	−213.4	−212.8	−215.8	−223.5/9.98	−225.0/9.17	−212.1	−221.3/ 10.16	−222.9	−209.8	−210.0	−*161*[c]
C4	164.3	163.1	162.1	163.1	164.9	164.3	163.6	164.3	164.1	163.0	163.0	171.3
C5	—	—	—	—	108.3	110.9	109.9	—	—	—	—	—
C5/H5	99.1/5.21	100.4/5.56	101.7/5.55	105.9/5.95	—	—	—	100.4/5.28	102.4/5.72	102.3/5.77[7.7]	101.8/5.70	95.8/5.85[7.2]
C6	152.4	149.4	150.9	162.5	—	—	—	—	—	—	—	—
C6/H6	—	—	—	—	136.5/6.74	140.0/7.00	137.8/6.98	140.6/6.87	144.0/7.19	138.1/7.16 [5.6,7.7]	141.7/7.13	146.1/7.40 [7.2]
CH_3 at C5	—	—	—	—	11.6/1.61	12.4/1.93	13.0/1.93	—	—	—	—	—
CH_3 at C6	18.4/1.94	18.7/2.15	19.5/2.22	23.6/2.16	—	—	—	—	—	—	—	—
$\underline{CH_2}CH_3$	—	—	at N1:40.1/ 3.87	at O2:64.7/ 4.45	—	at N1:43.6/ 3.7814	at N1:44.4/ 3.7814	—	at N1: 43.9/ 3.80	—	at N1:44.7/ 3.79	at N1:45.4/ 3.90
			14.1/1.26	14.2/1.41		4/1.29	3/1.31		14.4/1.32		14.3/1.30	14.5/1.35
$\underline{CH_2}CH_3$	—	at N3: 35.5/ 3.96	at N3:36.3/ 3.96	at N3:36.1/ 4.03	—	—	at N3:36.5/4.03	—	—	at N3: 35.8/3.99 [7.1]	at N3:36.2/ 3.98	at O4:63.0/ 4.43
		12.9/1.22	12.8/1.19	13.3/1.23			12.8/1.22			12.8/1.24[7.1]	12.8/1.20	14.2/1.36

[a] The $^1H/^{13}C$ and ^{15}N chemical shifts measured for **1**, **2** and **3** in the $CDCl_3$/DMSO solution (10 : 1). [b] The values in square brackets denote the ^3J(H–H) spin–spin coupling constants, in Hz. [c] Predicted nitrogen signals (based on the linear regression line) of the compounds presented in Table 4, unrecorded in the ^1H–^{15}N HMBC experiment.

shielding decrease by *ca.* 8 and 11 ppm for N1 and N3 nuclei was observed. Likewise, we noticed a similar effect in the compounds with monosubstituted nitrogen nuclei (*ca.* 8 ppm for N1; *ca.* 12 ppm for N3). However, for the unsubstituted nitrogens an inverse effect for the ^{15}N shielding was observed (*ca.* 2–3 ppm, Table 4) when compared with the respective shieldings in substrates **1–3**. For the carbon atoms, these changes were smaller and less representative. The position of the substitution was determined by analyzing the ^{1}H/^{13}C NMR chemical shifts for methylene of ethyl groups. In the case of the N^1-ethyl derivatives, the ^{1}H/^{13}C chemical shifts were as follows: *ca.* 3.80/45 ppm, whereas for N^3-ethyl uracils: *ca.* 4.00/36 ppm.

The NMR parameters, similarly to the **1**/**1d** pair of the compounds, reflected the changes in the bond location/hybridization in the O^2- and N^3-dialkylation products. The substitution of the protons at O2 and N3 in **1** caused the ^{15}N shielding decrease of the nitrogen nuclei by *ca.* 60 (N1) and 10 ppm (N3) in compound **1l**. The same trend (shielding decrease) was noticeable in the^{13}C NMR spectrum for the C2/C5/C6 and CH_3 carbon atoms, and was 4, 6, 10 and 5 ppm, respectively. Likewise, the replacement of the protons in compound **3** at the N^1- and O^4- positions leading to **3m** resulted in the minor ^{13}C/^{15}N shielding changes. For the nuclei N1/C2/C4 and C6, these effects were as follows: *ca.* 17, 5, 7 and 6 ppm, respectively. A bigger change should have been noticed for the N3 atom but unfortunately the correlation spot for this nucleus was not observed in the ^{1}H–^{15}N HMBC experiment. Only based on the calculated regression line could we determine the ^{15}N chemical shift of the N3 nucleus in the **3m** compound (*ca.* −161 ppm). The positions of the ethyl groups in **1l** and **3m** strongly influenced the ^{1}H/^{13}C NMR chemical shifts of these alkyl moieties. In particular, the atoms of the methylene groups exhibited characteristic ^{1}H/^{13}C NMR chemical shifts (Table 4) depending on whether they adjoined nitrogen or oxygen. The carbon nuclei of the methylene groups at the nitrogen atoms N1/N3 are more shielded than the respective carbons connected to the oxygen atom by *ca.* 20–30 ppm. Moreover, the ^{13}C shielding for the methylene carbon depends strongly on the position of the methylene group (N^1- or N^3-). For the N^1-derivatives, carbon nuclei are less shielded by *ca.* 8–10 ppm than in the N^3-substituted products. Thus, this effect allowed easy determination of the alkylation position in the uracil derivatives, which made the entire identification possible and faster.

The quantum mechanical DFT calculations

Such a diversity of the substituted reaction products turned out to be quite unexpected. Therefore, theoretical calculations were performed to estimate the molecular structure and NMR chemical shifts as well as the thermodynamic conditions of these reactions. The molecular structures and their molecular energies were calculated by using the quantum mechanical DFT B3LYP method with the 6-31G(d) and 6-311G(d) basis sets and simulating the solvent effects (acetonitrile or DMF) with the SMD model of the solvent.[20]

The predicted chemical shifts and a comparison of the empirical and theoretical values are given in Tables S1 and S2 (ESI†). These tables show that the theoretically predicted chemical shifts and the experimental data of the ^{1}H, ^{13}C or ^{15}N nuclei are highly correlated with R^2 in the range 0.87–0.99 (Table S3†). Based on the calculated regression lines, it was possible to estimate the lacking chemical shift of the N3 nitrogen in the **2c** and **3m** compounds. Also, the regression lines supported the signal assignment of the NMR chemical shifts to certain nuclei in the investigated molecules.

The Gibbs free energy outputs for the model reactions of Boc_2O and EtI with uracil derivatives were estimated using the B3LYP/6-311G(d) method (Tables S4 and S5 in the ESI†). In the majority of cases, the theoretical prediction of the Gibbs free energy of the model reactions leading to the mono/di-substitutions of Boc/Et of uracil, thymine and 6-methyluracil corresponded qualitatively to the experimental reaction yield (Tables S6 and S7†). For example, for a synthesized substitution product the corresponding Gibbs free energy of the model reactions was predicted to be negative (for **1b** (−10.49 kcal mol^{-1}), **1c** (−18.86), and **1d** (−6.71)) except for N^1-Boc-6-methyluracil, where Gibbs free energy was slightly positive (**1a** (+2.11 kcal mol^{-1})).

The experimental results also showed that product **1a** is usually obtained with the lowest yield.

An interesting feature of the theoretical calculations was the predicted formation of many more substituted derivatives than those found experimentally. We can only speculate that in the theoretical modeling some unknown details of the reaction mechanism had been omitted or perhaps the kinetic barrier was too high, preventing the detection of the supposed derivatives in the course of the experiment. A newly found compound **1g** and intermediate structures of its (hypothetical) synthesis were studied using the DFT B3LYP method with a smaller 6-31G(d) basis set. The molecular structure of **1g** is given in Table S8,† while free energies of the (hypothetical) reaction components are given in Table S9.† It is predicted that the free energy of the **1g** formation is about 14 kcal mol^{-1}. In the course of the reaction, a transition structure **6** is postulated to have been formed (Table S10†).

Another interesting theoretical result was the predicted susceptibility of the C5 and C6 positions of uracil derivatives to the electrophilic attack. It is worth mentioning that such a property corresponded well with the C5 and C6 gas theoretical phase acidities which were larger than the N1 or N3 acidities for uracil and 6-methyluracil, following calculations by means of the B3LYP/6-31+G(d) method by Kurinovich *et al.*[21] Another argument can be raised based on the reactivity index in the form of the Fukui's cumulative function[22] for the electrophilic attack, f^-, calculated for the C5/C6 positions (Table S11†). Considerably negative f^- values for 6-methyluracil suggested that the C5 position should be one of the possible regions for the electrophilic attack, as it occurred with the reaction of **1b** with Boc-pyridine **5** to give **1g**. However, most of the predicted C^5- or C^6-substituted derivatives were not found in the present experiments. One cannot exclude a possibility of synthesizing them *via* alternative reaction routes, other than those proposed here. We have also observed that the pyridine and DMAP sol-

vents used in the reaction of Boc_2O with thymine and 6-methyluracil affected the reaction yield of the N^1,N^3-derivatives. The theoretical calculations showed a minor effect on the reaction ΔG caused by the hydrogen bonds between pyridine molecules and mono/di-Boc derivatives of 6-methyluracil (Table S12†). A more pronounced effect on ΔG can be seen for DMAP H-bonded to the reagents, though essentially it does not change the order of ΔG predicted with the model where no explicit solvent was used.

Experimental

Chemicals and analytical grade solvents were purchased from commercial suppliers and used without further purification unless stated otherwise. Flash column chromatography was performed on silica gels (200–300 mesh).

The melting points were determined by using a Melting Point System (Mettler Toledo MP70).

The ^{1}H NMR, ^{13}C NMR and ^{15}N NMR (as 2D experiments) spectra were recorded in $CDCl_3$ and $CDCl_3$/DMSO-d_6 solutions with a Varian-NMR-vnmrs600 spectrometer (at 298 K) equipped with a 600 MHz PFG Auto XID (^{1}H/^{15}N–^{31}P 5 mm) indirect probehead. Standard experimental conditions and standard Varian programs (ChemPack 4.1) were used. The ^{1}H and ^{13}C NMR chemical shifts are given relative to the TMS signal at δ = 0.0 ppm, whereas neat nitromethane at δ = 0.0 ppm was used as a standard for the ^{15}N NMR chemical shifts. The concentration of the solutions used for the measurements was about 10–20 mg of the compounds in 0.6 cm^3 of solvents. In the case of uracil and its methyl derivatives (thymine and 6-methyluracil), the appropriate compounds were measured in saturated solution in $CDCl_3$/DMSO (10 : 1).

The mass spectra were recorded on a MaldiSYNAPT G2-S HDMS (Waters) Spectrometer *via* 25 electrospray ionization (ESI-MS).

Alkylation of the pirymidine base (uracil, 6-methyluracil, thymine) with Boc_2O

Procedure A. 4-DMAP (8 mg, 0.065 mmol) was added to the solution of Boc_2O (2.08 g, 9.52 mmol) and 6-methyluracil (400 mg, 3.17 mmol) in MeCN (20 ml). The reaction mixture was stirred for 48 h at room temperature. The solvent was evaporated *in vacuo* and the residue was purified by flash chromatography on silica gel (eluent: hexane–AcOEt, 8 : 2, 7 : 3, 1 : 1).

Procedure B. Thymine (400 mg, 3.17 mmol), Boc_2O (2.08 g, 9.52 mmol), pyridine (4 ml), and MeCN (20 ml) were stirred together for 4 h at 55 °C. The reaction mixture was concentrated *in vacuo* and the residue was partitioned between CH_2Cl_2 (25 ml) and water (25 ml). The organic layer was separated and the aqueous phase was extracted twice with CH_2Cl_2 (15 ml). The combined organic layers were dried over $MgSO_4$ and filtered. The solvent was removed by evaporation *in vacuo* and the residue was purified by flash chromatography on silica gel (eluent: hexane–AcOEt, 7 : 3).

Derivatives of 6-methyluracil. 1a: (22 mg, 3%), white solid; Mp 296.4 °C dec. (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{10}H_{14}N_2O_6Na$: 249.0843, found: 249.0851.

1b: (21 mg, 3%), solid; Mp 249.8 °C dec. (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{10}H_{14}N_2O_6Na$: 249.0843, found: 249.0851.

1c: (111 mg, 11%), white solid; Mp 121.5 °C (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{15}H_{22}N_2O_6Na$: 349.1376, found: 349.1367.

1d: (327 mg, 32%), solid; Mp 68.7 °C (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{15}H_{22}N_2O_6Na$: 349.1373, found: 349.1376.

1g: (257 mg, 20%), foam; HRMS (ESI) *m*/*z*: [M − H] calcd for $C_{20}H_{26}N_3O_6$: 404.1822, found: 404.1818.

Thymines. 2a: (423 mg, 59%), solid; Mp 324.4 °C dec. (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{10}H_{14}N_2O_6Na$: 249.0846, found: 249.0851.

2b: (28 mg, 4%), white solid; (lit.[12]) Mp 309.8 °C dec. (hexane–AcOEt); HRMS (ESI) *m*/*z*: [M − H] calcd for $C_{10}H_{13}N_2O_4$: 225.0875, found: 225.0878.

2c: (403 mg, 39%), solid; (lit.[12]) Mp 145.6 °C (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{15}H_{22}N_2O_6Na$: 349.1376, found: 349.1365.

Uracils. 3a: (84 mg, 12%), solid; (lit.[14]) Mp 290.1 °C dec. (hexane–AcOEt); HRMS (ESI) *m*/*z*: [M − H] calcd for $C_9H_{11}N_2O_4$: 211.0719, found: 211.0717.

3b: (83 mg, 12%), solid; Mp 322.4 °C dec. (hexane–AcOEt); HRMS (ESI) *m*/*z*: [M − H] calcd for $C_9H_{11}N_2O_4$: 211.0719, found, 211.0722.

3c: (464 mg, 42%), white solid; Mp 98.8 °C (hexane–AcOEt); HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_{14}H_{20}N_2O_6Na$: 335.1219, found, 335.1209.

Alkylation of the pirymidine base (uracil, 6-methyluracil and thymine) with EtI

Procedure. EtI (144 μl, 1.79 mmol) was added to the stirred solution of thymine (150 mg, 1.19 mmol), K_2CO_3 (329 mg, 2.38 mmol) and TBABr (48 mg, 0.148 mM) in anhydrous DMF (2.5 ml). The reaction mixture was stirred for 48 h at room temperature. Water (15 ml) and CH_2Cl_2 were added. The organic layer was separated, and the aqueous phase was extracted twice with CH_2Cl_2 (15 ml). The combined organic layers were dried over $MgSO_4$ and filtered. Then, the solvent was removed by evaporation *in vacuo*. The residue was purified by flash chromatography on silica gel (eluent: hexane–AcOEt 7 : 3 than hexane–AcOEt 1 : 1).

Thymines. 2h: (41 mg, 22%), white solid; (lit.[18,25]) Mp 227.6 °C (hexane–AcOEt); HRMS (EI) *m*/*z*: [M] calcd for $C_7H_{10}N_2O_2$: 154.0742, found: 154.0740.

2j: (59 mg, 30%), solid; Mp 63.2 °C (hexane–AcOEt); HRMS (EI) *m*/*z*: [M] calcd for $C_9H_{14}N_2O_2$: 182.1055, found: 182.1055.

6-Methyluracils. 1i: (9 mg, 5%), white solid; Mp 203.6 °C (hexane–AcOEt); HRMS (ESI) *m*/*z*: [M] calcd for $C_7H_9N_2O_2$: 153.0664, found: 153.0659.

1j: (162 mg, 75%), solid; (lit.[26]) Mp 59.7 °C (hexane–AcOEt); HRMS (EI) *m*/*z*: [M] calcd for $C_9H_{14}N_2O_2$: 182.1055, found: 182.1057.

1l: (6 mg, 3%), oil; HRMS (EI) *m*/*z*: [M] calcd for $C_9H_{14}N_2O_2$: 182.1055, found: 182.1051.

Uracils (1.5 eq. of EtI). 3h: (93 mg, 12%), white solid; (lit.[5,27]); Mp 152.7 °C (hexane–AcOEt); HRMS (EI) *m*/*z*: [M] calcd for $C_6H_8N_2O_2$: 140.0586, found: 140.0584.

3i: (5 mg, 1%), oil; HRMS (ESI) *m*/*z*: [M] calcd for $C_6H_7N_2O_2$: 139.0508, found: 139.0509.

3j: (282 mg, 31%), oil (lit.[5]); HRMS (EI) *m*/*z*: [M] calcd for $C_8H_{12}N_2O_2$: 168.0899, found: 168.0895.

3m: (5 mg, 1%), oil; HRMS (ESI) *m*/*z*: $[M + Na]^+$ calcd for $C_8H_{12}N_2O_2Na$: 191.0796, found: 191.0791.

Quantum mechanical DFT calculations

The theoretical calculations have been performed with the Gaussian G09 suite of programs [Gaussian]. The molecular geometries, harmonic frequencies and isotropic nuclear shieldings (GIAO) were calculated following standard settings within the G09 code.

Conclusions

In general, alkylation/acylation of pyrimidine bases afforded mainly N^1-mono-substituted and N^1,N^3-di-substituted derivatives than N^3- or acid labile *O*-alkyl derivatives.[23] The regioselectivity of these reactions was determined by the acidity of the ionisable protons of the heterocyclic ring. As mentioned in the literature,[24] the acidity of the N^1–H proton (pK_a = 9.43 for U and 9.86 for thymine) is higher than that of N^3–H (pK_a > 13 for U and 13.96 for thymine). Despite a big difference in the pK_a of the protons it is difficult to selectively obtain the N^3-mono alkylation product.

In conclusion, the outcome of the alkylation of 6-methyluracil, uracil and thymine with Boc_2O depends strongly on the reaction conditions. The regioselectivity of the Boc/Et substitution of uracil, thymine and 6-methyluracil can be steered by a reasonable choice of the experimental setup. We have studied the following parameters controlling regioselectivity: the molar concentration ratio of the substrate and alkylation agent, the presence/absence of the catalyst (DMAP in this case), and the temperature of the reaction (ambient, elevated). We have isolated different acylation/alkylation products of three pyrimidine bases: uracil, 6-methyluracil and thymine and discovered a new product containing the *N*-Boc-pyridinium moiety at the C^5-position of 6-methyluracil. All the products were fully characterized by using multinuclear NMR data.

The experimental findings were supported by the quantum mechanical DFT calculations of the molecular energies and theoretically predicted chemical shifts.

Author contributions

Olga Michalak: conceived the idea, structure and design of the paper, conceived and designed the synthesis and performed the experiments, analyzed the data, and wrote the majority of the manuscript; Piotr Cmoch: conceived and designed experiments, performed the NMR experiments, analyzed the data and wrote part of the paper; Piotr Krzeczyński: conceived and designed experiments, performed the experiments, analyzed the data, and wrote part of the paper; Marcin Cybulski: conceived and designed experiments, analyzed the data, and wrote part of the paper; Andrzej Leś: conceived and designed theoretical modeling, performed the calculations, analyzed the data, and wrote part of the paper. All authors consulted their results, have read, critically reviewed and agreed to the final version of the manuscript.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgements

This work was supported under the framework of the statutory project funded by the Polish Ministry of Science and Higher Education. The calculations were performed at the Interdisciplinary Centre for Mathematical and Computational Modeling of the University of Warsaw (ICM UW, the computer grant G18-6) which is kindly acknowledged for allocating facilities and computer time.

Notes and references

1 (*a*) I. M. Lagoja, *Chem. Biodiversity*, 2005, **2**, 1–50; (*b*) J. I. Bardagía and R. A. Rossi, *Org. Prep. Proced. Int.*, 2009, **41**, 479–514.
2 A. Pałasz and D. Cież, *Eur. J. Med. Chem.*, 2015, **97**, 582–611.
3 L. Arias, A. Guzman, S. Jaime–Figueroa, F. J. Lopez, D. J. Morgans Jr., F. Padilla, A. Perez–Medrano, C. Quintero, M. Romero and L. Sandoval, *Synlett*, 1997, 1233–1234.
4 A. L. Pogolotti Jr., D. Failla and D. V. Santi, *J. Heterocycl. Chem.*, 1972, **9**, 1423.
5 I. Ishikawa, T. Itoh, H. Takayanagi, J. Oshima, N. Kawahara, Y. Mizuno and H. Ogura, *Chem. Pharm. Bull.*, 1991, **39**, 1922–1930.
6 M. Frieden, M. Giraud, C. B. Reese and Q. Song, *J. Chem. Soc., Perkin Trans. 1*, 1998, 2827–2832.
7 N. G. Kundu and S. G. Khatri, *Synthesis*, 1985, 323–326.
8 S. Manfredine, P. G. Baraldi, R. Bazzanini, M. Guarneri and D. Simoni, *J. Chem. Soc., Chem. Commun.*, 1994, 583–584.
9 (*a*) A. Sarkar, S. R. Roy, N. Parikh and A. K. Chakraborti, *J. Org. Chem.*, 2011, **76**, 7132–7140; (*b*) E. A. Englund, H. N. Gopi and D. H. Appella, *Org. Lett.*, 2004, **6**, 213–215; (*c*) R. Moumne, S. Lavielle and P. Karoyan, *J. Org. Chem.*, 2006, **71**, 3332–3334; (*d*) D. M. Shendage, R. Froehlich and G. Haufe, *Org. Lett.*, 2004, **6**, 3675–3678.

10 J. R. Hwu, M. L. Jain, S.-C. Tsay and G. H. Hakimelahi, *Tetrahedron Lett.*, 1996, **37**, 2035–2038.
11 S. Jaime–Figueroa, A. Zamilpa, A. Guzmán and D. J. Morgans, *Synth. Commun.*, 2001, **31**, 3739–3746.
12 M. F. Jacobsen, M. M. Knudsen and K. V. Gothelf, *J. Org. Chem.*, 2006, **71**, 9183–9190.
13 M. Bessières, V. Roy and L. A. Agrofoglio, *RSC Adv.*, 2014, **4**, 59747–59749.
14 L. Lipani, D. Odadzic, L. Weizel, J.-S. Schwed, B. Sadek and H. Stark, *Eur. J. Med. Chem.*, 2014, **86**, 578–588.
15 J.-In Lee, *Bull. Korean Chem. Soc.*, 2010, **31**, 749–752.
16 W.-B. Qin, J.-Yi Zhu, Y.-B. Kong, Y.-H. Bao, Z.-W. Chen and L.-X. Liu, *Org. Biomol. Chem.*, 2014, **12**, 4252–4259.
17 Z. Paryżek and B. Tabaczka, *Org. Prep. Proced. Int.*, 2001, **33**, 400–405.
18 T. Tanabe, K. Yamauchi and M. Kinoshita, *Bull. Chem. Soc. Jpn.*, 1977, **50**, 3021–3025.
19 A. Gambacorta, D. Tofani, M. A. Loreto, T. Gasperi and R. Bernini, *Tetrahedron*, 2006, **62**, 6848–6854.
20 M. J. Frisch, G. W. Trucks, G. E. Schlegel, H. B. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT, 2013.
21 M. A. Kurinovich and J. K. Lee, *J. Am. Soc. Mass Spectrom.*, 2002, **13**, 985–995.
22 (*a*) K. Fukui, *Science*, 1982, **218**, 747–754; (*b*) W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708–5711.
23 G. Shaw, in *Comprehensive Heterocyclic Chemistry*, ed. A. R. Katritzky and C. W. Rees, Pergamon, Oxford, 1985, p. 499.
24 B. Nawrot, O. Michalak, S. Olejniczak, M. W. Wieczorek, T. Lis and W. J. Stec, *Tetrahedron*, 2001, **57**, 3979–3985.
25 Q.-Q. Wu and Q.-H. Song, *J. Phys. Chem. B*, 2010, **114**, 9827–9832.
26 H. Egg and I. Volgger, *Synthesis*, 1982, 1071–1073.
27 A. Robin, K. Julienne, J.-C. Meslin and D. Deniaud, *Eur. J. Org. Chem.*, 2006, 634–643.