



# A multicomponent tetrazolo indole synthesis†

 Cite this: *Chem. Commun.*, 2021, 57, 6652

 Received 5th May 2021,  
 Accepted 9th June 2021

DOI: 10.1039/d1cc02384e

rsc.li/chemcomm

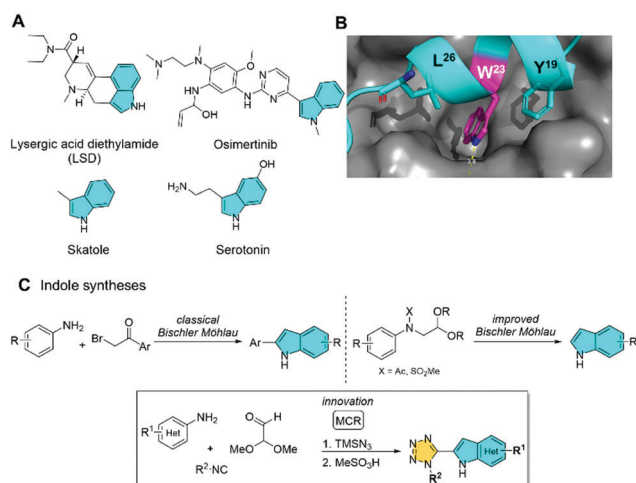
 Xiaofang Lei,<sup>ab</sup> Panagiota Lampiri,<sup>a</sup> Pravin Patil,<sup>b</sup> Giasemi Angeli,<sup>b</sup> Constantinos G. Neochoritis<sup>b</sup>\*<sup>a</sup> and Alexander Dömling<sup>b</sup>\*<sup>b</sup>

**The ubiquitous presence of the indole fragment in natural products and drugs asks for ever novel syntheses. We report an unprecedented mild, two-step synthesis of 2-tetrazolo substituted indoles based on the Ugi-tetrazole reaction combined with an acidic ring closure. A gram-scale synthesis, a bioactive compound and further transformations were performed.**

Indole was discovered in 1866 by Adolf von Baeyer and is the parent compound of an archetypical class of heterocycles which are essential substructures in many drugs and natural products.<sup>1–3</sup> Examples range from the smelly feces component skatole, to the neurotransmitter serotonin (5-hydroxy tryptamine), to the psychedelic drug LSD, the recently introduced oral irreversible EGFR inhibitor Osimertinib targeting the T790M mutation, or the anticancer drug vincristine (Fig. 1A). Indole containing bioactive compounds often have a distinct binding mode to their receptors. For example, in the protein–protein interaction between MDM2 and p53, the tryptophan W23 plays an extraordinary anchoring role in the hot spot triad YWL (Fig. 1B). No doubt indole is a privileged scaffold in nature.<sup>4,5</sup> Accordingly, many different indole syntheses have been invented.<sup>6,7</sup> The most famous and highly versatile indole synthesis was discovered by Emil Fischer in 1883.<sup>8</sup> Examples of a three component indole syntheses include the reaction of monosubstituted alkynes, trifluoro acetylated anilines and bromoarenes.<sup>9</sup> Some special features of the indole moiety include its size and its hydrogen bond donating NH (exemplified in the amino acid tryptophane) together with the electron rich 5-membered pyrrole ring which is prone to undergo electrophilic additions (exemplified in natural product chemistry). Despite

many known indole syntheses, the number of multicomponent reactions indole syntheses is rather limited.<sup>10,11</sup>

Based on our ongoing interest in novel indole syntheses<sup>12,13</sup> we were inspired by the Bischler–Möhlau indole synthesis which involves the alkylation of anilines with bromoacetophenones, followed by the acid induced indole formation (Fig. 1C).<sup>14,15</sup> However, due to the drastic reaction conditions (~200 °C, HBr) the reaction is neither practical nor compatible with many functional groups. Even milder variations involving the cyclization of (*N*-acetylated or sulfonated) aniline acetaldehyde diethyl acetals to indoles suffer from missing variability due to the sequential synthesis of and limited access to the aniline acetaldehyde acetals.<sup>16,17</sup> To solve these issues we redesigned the synthesis of key acetal aniline intermediates by including a mild multicomponent reaction approach which would allow for a greater variability of the different components



**Fig. 1** The nature of indole. (A) Several indole containing natural products and drugs. (B) Indole in structural biology (PDB ID 1YCR). W23 (tryptophan) binding to MDM2 by shape complementarity and a hydrogen bond. (C) Exemplary syntheses of indole.

<sup>a</sup> University of Crete, Department of Chemistry, Heraklion, Greece.  
 E-mail: kneochor@uoc.gr

<sup>b</sup> University of Groningen, Department of Pharmacy, Drug Design group, Groningen, The Netherlands. E-mail: a.s.s.domling@rug.nl

† Electronic supplementary information (ESI) available. CCDC 2077271. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/d1cc02384e



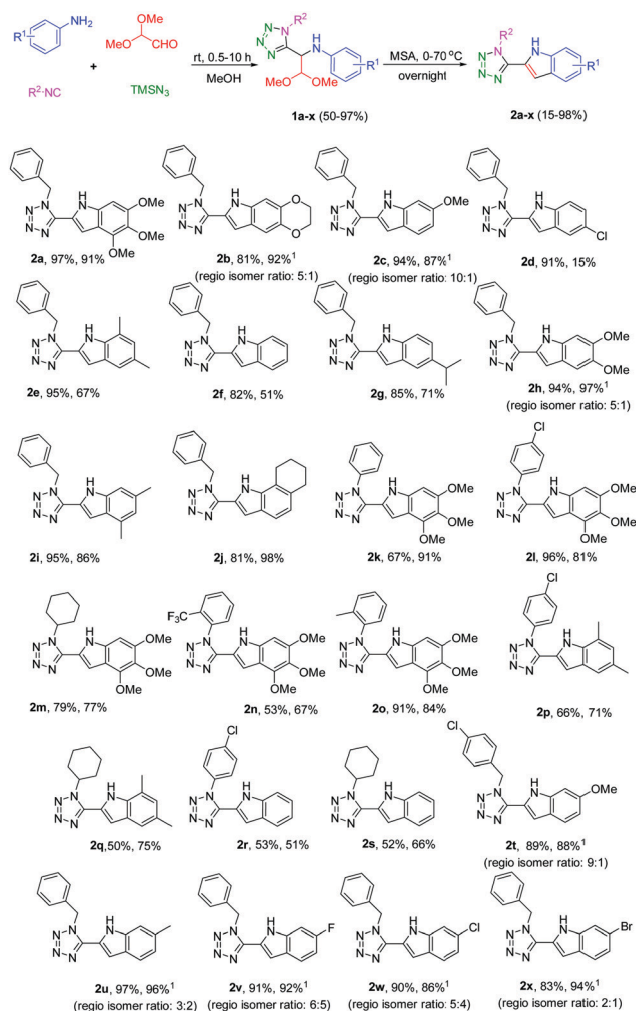
and positions ( $R^1$ ,  $R^2$ ) to access highly substituted indoles (Fig. 1C).

The multicomponent of choice was the Ugi-tetrazole four component reaction (UT-4CR).<sup>18,19</sup> Thus, the reaction of substituted anilines, isocyanides, 2,2-dimethoxyacetaldehyde in aqueous solution and  $TMSN_3$  proceeded smoothly at room temperature affording the UT-4CR derivatives **1a-x** in good to excellent yields (Scheme 1 and ESI†).

The scope of this reaction is very broad, with substituents on both isocyanide and aniline moieties as electron-donating (EDG) and withdrawing groups (EWG) as well. Then, we performed an acidic closure reaction.<sup>20-24</sup> We have screened several acidic conditions with both organic and inorganic acids at different temperatures but the best results were obtained by anhydrous methane sulfonic acid (MSA,  $\geq 99\%$  purity, freshly opened bottle). The water content in MSA proved very important for the cyclization yield as increased levels of water led to side reactions lowering the overall yield. Noteworthy, the

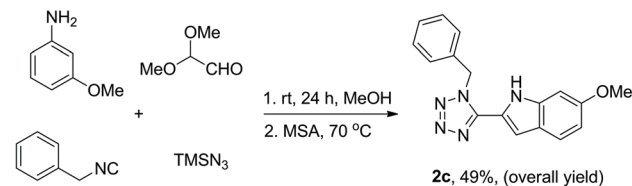
cyclization *via* formic or acetic acid afforded the *N*-formylated or acetylated tetrazole indoles in substantially reduced yields (see ESI†). To our delight, the targeted multi-substituted tetrazole indole derivatives **2a-x** have been successfully and very efficiently synthesized *via* MSA at 70 °C in good to very good yields. In several cases, both the UT and the final adducts precipitated out during the reaction mixture after short reaction times (see ESI†). The scope of the isocyanides is very broad, all the isocyanides that were employed, aryl, benzylic and aliphatic with different substituents reacted efficiently. We preferentially employed anilines with EDGs as substituents due to the presumed electrophilic ring closure mechanism. Substituted with EWGs anilines, *i.e.* **2d**, reacted sluggish in the ring closure reaction. The indoles derived from *meta*-substituted anilines, such as **2b**, **2c**, **2h** and **2u-x**, expectedly were formed as two different isomers. However, after some optimization in the reaction temperature (in case of **2b**) or selectively precipitation (in case of **2c**, **2y** and **2v**), we were able to get only the major isomer that is shown in Scheme 1 in isolated yields of 77%, 79%, 69% and 50%, respectively. Compounds **2u**, **2w**, **2x** were obtained as a mixture of the regioisomers (see ESI†).

We have synthesized 24 multi-substituted tetrazole indoles in our 2-step approach. Next, we investigated the feasibility of a one-pot procedure. Indeed, we were able to access **2c** in a one-pot fashion in 49% overall yield (Scheme 2A). Furthermore, our approach is easily scalable; we have synthesized **1a** in a 10 mmol scale (97% yield, 4.15 g) and subsequently, in a 5 mmol scale (2.15 g) compound **2a** which was isolated as a brown solid by simple precipitation, without the need of column chromatography (Scheme 2B). Noteworthy, the final tetrazole indoles **2** demonstrate a characteristic fluorescence at 366 nm compared to the starting materials **1** (Scheme 2C).

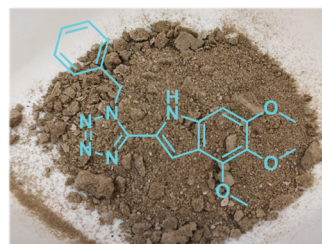


**Scheme 1** Preparation of the substituted tetrazole indoles. The yields of both steps are reported (first and second step, respectively); <sup>1</sup>obtained as regioisomers (ratio shown in parenthesis, yield refers to the mixture), structure of the major regioisomer is given.

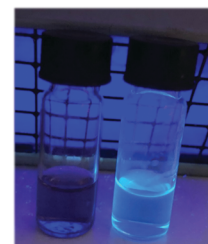
### A One-pot approach



### B gram-scale synthesis

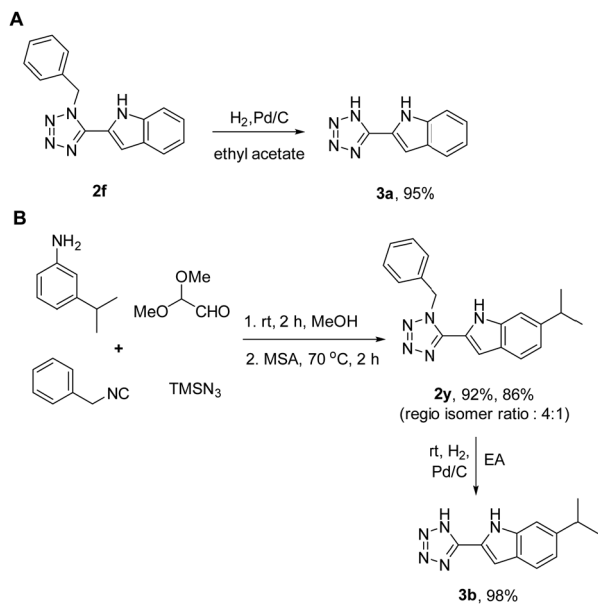


### C Fluorescence at 366 nm



**Scheme 2** (A) The one-pot approach without intermediate purification. (B). Multi gram-scale synthesis of **2a**. (C). Fluorescence of the solution of **2a** (right vial) in acetone at 366 nm compared to the starting material (left vial).



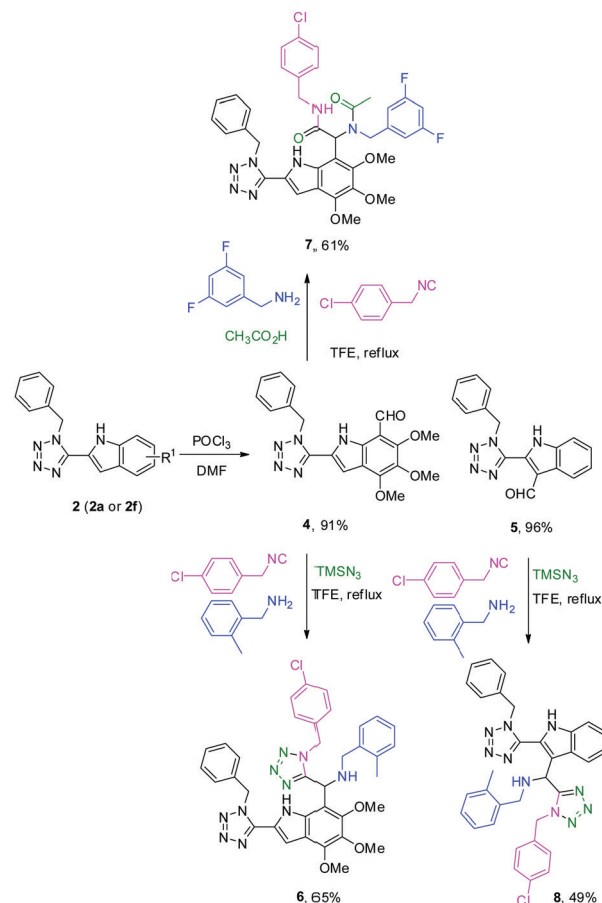


**Scheme 3** (A). Synthesis of the 2-indole carboxylic acid bioisostere **3a**. (B). Synthesis of the eIF4A3 inhibitor **3b**.

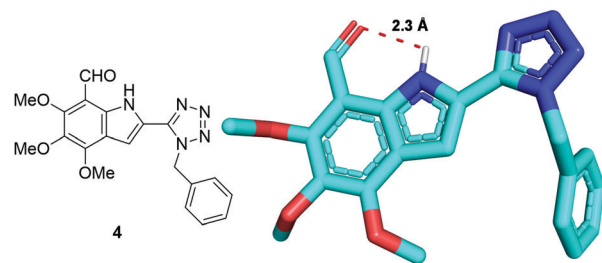
1,5-Disubstituted free NH tetrazoles are known bioisosteres of carboxylic acids. Moreover, 2-indole carboxylic acids have been utilized as effective inhibitors against various targets.<sup>25–28</sup> Therefore, we also established an example of 2-indole carboxylic acid bioisostere. Employing the convertible benzyl isocyanide, we hydrogenated the tetrazole indole **2f** towards the free NH tetrazole indole **3a** (Scheme 3A). Establishing the usefulness of our synthetic pathway, we synthesized the bioactive compound **3b** in only 3 steps with an overall yield of 62% (Scheme 3B). The tetrazolo indole **3b** is a potent eukaryotic initiation factor 4A3 (eIF4A3) inhibitor.<sup>29</sup>

In order to further demonstrate the potential of our synthetic strategy, we functionalized the tetrazole indole derivatives **2**. We installed first a formyl group *via* a Vilsmeier–Haack formylation and then performed a second MCR, establishing a union of MCRs.<sup>30</sup> We noticed that the formylation on the **2a** led to a mixture (1 : 1) of two formylated adducts on the 3- and 7-position of the indole ring due to the probably electron-rich aromatic ring (see ESI<sup>†</sup>). Tuning the formylation by changing the addition ratio of POCl<sub>3</sub> and temperature, we formylated exclusively the 7-position of the indole ring, affording compound **4** in 91% yield. When we switched to the formylation on the less electron-rich tetrazole indole **2f**, then we obtained the indole derivative **5** in 96% (Scheme 4). Next, we functionalized the formylated indoles by performing an additional UT-4CR and the classical variation of the Ugi reaction (U-4CR). Thus, we obtained the UT and U-4CR adducts **6**, **7** and **8** increasing both the complexity and diversity of our initial tetrazole indoles (Scheme 4).

In support of both the proposed scaffold **2** and **4**, we solved the crystal structure of the latter (Fig. 2). Noteworthy, an intramolecular hydrogen bond of 2.3 Å between the –NH and the –CHO can be observed.



**Scheme 4** Functionalization of the formyl indole derivatives **4** and **5** via additional UT-4CR and U-4CR.



**Fig. 2** Crystal structure of the formylated tetrazolo-indole **4** (CCDC 2077271).<sup>†</sup>

The herein disclosed 2-step approach is a useful addition to the indole syntheses toolbox due to the mildness of the reaction conditions. It offers access to 1,5-indolo-tetrazoles with the beneficial physicochemical properties<sup>31</sup> and their bioisosterism to carboxylic acids.<sup>18</sup> Tetrazole-indole derivatives have known important biological activity such as selective ATP-competitive eIF4A3 inhibitors,<sup>32</sup> angiotensin II-1 antagonists,<sup>33</sup> nociceptin/orphanin FQ (N/OFFQ) receptor antagonists<sup>34</sup> and potential anti-allergy agents.<sup>35</sup>

Our short, diverse, and scalable Ugi synthesis outperforms all currently known multistep syntheses towards this scaffold,



where commonly three up to seven sequential steps are needed even starting from a basic indole core. Our method provides easy and rapid access to functionalized indole derivatives that can be used in drug discovery campaigns as bioisosteres of 2-indole carboxylic acids and amides, amongst other applications. Moreover, the great potential of the synthesis was supported by a multi gram synthesis and several post-modifications increasing both the diversity and the complexity of target compounds.

A. D. and C. G. N. conceptualized and directed the project. X. L. and P. L. performed the syntheses and collected the analytical data. P. P. performed the data curation and formal analysis. S. A. determined the single crystal X-ray structure. A. D., C. G. N. and X. L. contributed to the manuscript writing. The research project (to C. G. N) was supported by the Hellenic Foundation for Research and Innovation (H. F. R. I.) under the “2nd Call for H. F. R. I. Research Projects to support Post-Doctoral Researchers” (Project Number: 0911). Xiaofang Lei acknowledges the China Scholarship Council for support. This project has received funding (to AD) from the European Lead Factory (IMI) under Grant Agreement 115489, the Qatar National Research Foundation (NPRP6-065-3-012), the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie (ITN “Accelerated Early stage drug discovery”, Grant Agreement 675555; Cofunds ALERT (665250) and PROMINENT (754425)) and the National Institutes of Health ((NIH 2R01GM097082-05)).

## Conflicts of interest

There are no conflicts to declare.

## Notes and references

- 1 A. Baeyer, *Ann. Chem. Pharm.*, 1866, **140**, 295.
- 2 P. V. Thanikachalam, R. K. Maurya, V. Garg and V. Monga, *Eur. J. Med. Chem.*, 2019, **180**, 562.
- 3 A. Dorababu, *RSC Med. Chem.*, 2020, **11**, 1335.
- 4 M. E. Welsch, S. A. Snyder and B. R. Stockwell, *Curr. Opin. Chem. Biol.*, 2010, **14**, 347.
- 5 H. Zhao and J. Dietrich, *Expert Opin. Drug Discovery*, 2015, **10**, 781.
- 6 G. R. Humphrey and J. T. Kuethe, *Chem. Rev.*, 2006, **106**, 2875.
- 7 M. Inman and C. J. Moody, *Chem. Sci.*, 2013, **4**, 29.
- 8 E. Fischer and F. Jourdan, *Ber. Dtsch. Chem. Ges.*, 1883, **16**, 2241.

- 9 B. Z. Lu, W. Zhao, H.-X. Wei, M. Dufour, V. Farina and C. H. Senanayake, *Org. Lett.*, 2006, **8**, 3271.
- 10 O. Leogane and H. Lebel, *Angew. Chem., Int. Ed.*, 2008, **47**, 350.
- 11 Y.-Y. Liu, X.-Y. Yu, J.-R. Chen, M.-M. Qiao, X. Qi, D.-Q. Shi and W.-J. Xiao, *Angew. Chem., Int. Ed.*, 2017, **56**, 9527.
- 12 K. Wang, E. Herdtweck and A. Dömling, *ACS Comb. Sci.*, 2011, **13**, 140.
- 13 C. G. Neochoritis and A. Dömling, *Org. Biomol. Chem.*, 2014, **12**, 1649.
- 14 A. Bischler, *Ber. Dtsch. Chem. Ges.*, 1892, **25**, 2860.
- 15 R. Möhlau, *Ber. Dtsch. Chem. Ges.*, 1881, **14**, 171.
- 16 B. Pete, G. Simig, L. Poszavác and L. Toke, *Heterocycles*, 2003, **60**, 2441.
- 17 B. G. Richardson, A. D. Jain, H. R. Potteti, P. R. Lazzara, B. P. David, C. R. Tamatam, E. Choma, K. Skowron, K. Dye, Z. Siddiqui, Y.-T. Wang, A. Krunic, S. P. Reddy and T. W. Moore, *J. Med. Chem.*, 2018, **61**, 8029.
- 18 L. G. Neochoritis, T. Zhao and A. Dömling, *Chem. Rev.*, 2019, **119**, 1970.
- 19 C. Kalinski, M. Umkehrer, S. Gonnard, N. Jäger, G. Ross and W. Hiller, *Tetrahedron Lett.*, 2006, **47**, 2041.
- 20 W. Wang, S. Ollio, E. Herdtweck and A. Dömling, *J. Org. Chem.*, 2011, **76**, 637.
- 21 V. Tyagi, S. Khan, V. Bajpai, H. M. Gauniyal, B. Kumar and P. M. S. Chauhan, *J. Org. Chem.*, 2012, **77**, 1414.
- 22 L. El Kaim, M. Gageat, L. Gaultier and L. Grimaud, *Synlett*, 2007, 0500.
- 23 H. Wang and A. Ganesan, *Org. Lett.*, 1999, **1**, 1647.
- 24 L. Cárdenas-Galindo, A. Islas-Jácome, N. Alvarez-Rodríguez, L. El Kaim and R. Gámez-Montaño, *Synthesis*, 2013, **49**, 49.
- 25 T. Sparey, P. Abeywickrema, S. Almond, N. Brandon, N. Byrne, A. Campbell, P. H. Hutson, M. Jacobson, B. Jones, S. Munshi, D. Pascarella, A. Pike, G. S. Prasad, N. Sachs, M. Sakatis, V. Sardana, S. Venkatraman and M. B. Young, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 3386.
- 26 T. Kariya, J. M. Grisar, N. L. Wiech and T. R. Blohm, *J. Med. Chem.*, 1972, **15**, 659.
- 27 J. D. Moore and A. Potter, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 4283.
- 28 D. A. Holt, D. S. Yamashita, A. L. Konialian-Beck, J. I. Luengo, A. D. Abell, D. J. Bergsma, M. Brandt and M. A. Levy, *J. Med. Chem.*, 1995, **38**, 13.
- 29 M. Ito, M. Iwatani, Y. Kamada, S. Sogabe, S. Nakao, T. Tanaka, T. Kawamoto, S. Aparicio, A. Nakanishi and Y. Imaeda, *Bioorg. Med. Chem.*, 2017, **25**, 2200.
- 30 T. Zarganes-Tzitzikas, A. L. Chandgude and A. Dömling, *Chem. Rec.*, 2015, **15**, 981.
- 31 G. M. Schroeder, S. Marshall, H. Wan and A. V. Purandare, *Tetrahedron Lett.*, 2010, **51**, 1404.
- 32 M. Ito, M. Iwatani, Y. Kamada, S. Sogabe, S. Nakao, T. Tanaka, T. Kawamoto, S. Aparicio, A. Nakanishi and Y. Imaeda, *Bioorg. Med. Chem.*, 2017, **25**, 2200.
- 33 S. Russell Stabler and Jahangir, *Synth. Commun.*, 1994, **24**, 123.
- 34 Y. Sugimoto, A. Shimizu, T. Kato, A. Satoh, S. Ozaki, H. Ohta and O. Okamoto, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 3569.
- 35 P. C. Unangst, D. T. Connor, S. R. Stabler, R. J. Weikert, M. E. Carethers, J. A. Kennedy, D. O. Thueson, J. C. Chestnut, R. L. Adolphson and M. C. Conroy, *J. Med. Chem.*, 1989, **32**, 1360.

