BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Medicinal chemistry and drug discovery in an opening era

Edited by Matthew Todd

Generated on 02 November 2026, 02:50

Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

The copyright to this document as a whole, which is published in the *Beilstein Journal of Organic Chemistry*, is held by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften. The copyright to the individual articles in this document is held by the respective authors, subject to a Creative Commons Attribution license.

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

An improved, scalable synthesis of Notum inhibitor LP-922056 using 1-chloro-1,2-benziodoxol-3-one as a superior electrophilic chlorinating agent

Nicky J. Willis[1], Elliott D. Bayle[1,2], George Papageorgiou[2], David Steadman[1], Benjamin N. Atkinson[1], William Mahy[1] and Paul V. Fish[*1,2]

Full Research Paper

Open Access

Address:
[1]Alzheimer's Research UK UCL Drug Discovery Institute, The Cruciform Building, University College London, Gower Street, London WC1E 6BT, UK and [2]The Francis Crick Institute, 1 Midland Road, Kings Cross, London NW1 1AT, UK

Email:
Paul V. Fish[*] - p.fish@ucl.ac.uk

* Corresponding author

Keywords:
brain penetration; 1-chloro-1,2-benziodoxol-3-one; electrophilic chlorination; LP-922056; Notum inhibitor

Beilstein J. Org. Chem. **2019,** *15,* 2790–2797.
doi:10.3762/bjoc.15.271

Received: 12 July 2019
Accepted: 05 November 2019
Published: 19 November 2019

Guest Editor: M. Todd

© 2019 Willis et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Background: The carboxylesterase Notum has been shown to act as a key negative regulator of the Wnt signalling pathway by mediating the depalmitoleoylation of Wnt proteins. LP-922056 (**1**) is an orally active inhibitor of Notum. We are investigating the role of Notum in modulating Wnt signalling in the central nervous system and wished to establish if **1** would serve as a peripherally restricted control. An accessible and improved synthetic route would allow **1** to become more readily available as a chemical tool to explore the fundamental biology of Notum and build target validation to underpin new drug discovery programs.

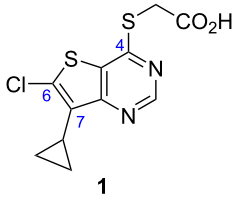
Results: An improved, scalable synthesis of **1** is reported. Key modifications include: (1) the introduction of the C7-cyclopropyl group was most effectively achieved with a Suzuki–Miyaura cross-coupling reaction with MIDA-boronate **11** (**5** → **6**), and (2) C6 chlorination was performed with 1-chloro-1,2-benziodoxol-3-one (**12**) (**6** → **7**) as a mild and selective electrophilic chlorination agent. This 7-step route from **16** has been reliably performed on large scale to produce multigram quantities of **1** in good efficiency and high purity. Pharmacokinetic studies in mouse showed CNS penetration of **1** is very low with a brain/plasma concentration ratio of just 0.01. A small library of amides **17** were prepared from acid **1** to explore if **1** could be modified to deliver a CNS penetrant tool by capping off the acid as an amide. Although significant Notum inhibition activity could be achieved, none of these amides demonstrated the required combination of metabolic stability along with cell permeability without evidence of P-gp mediated efflux.

Conclusion: Mouse pharmacokinetic studies demonstrate that **1** is unsuitable for use in models of disease where brain penetration is an essential requirement of the compound but would be an ideal peripherally restricted control. These data will contribute to the understanding of drug levels of **1** to overlay with appropriate in vivo efficacy endpoints, i.e., the PK-PD relationship. The identification of a suitable analogue of **1** (or **17**) which combines Notum inhibition with CNS penetration would be a valuable chemical probe for investigating the role of Notum in disease models.

Introduction

The Wnt signalling pathway has been shown to regulate crucial aspects of cell fate determination, organogenesis, cell migration and polarity [1]. Importantly, compromised Wnt signalling has been implicated in the perturbation of synaptic integrity and function in Alzheimer's disease (AD) [2]. Palmitoleoylation of Wnt proteins is required for efficient binding to Frizzled receptors and the subsequent signal transduction. The carboxylesterase Notum has been shown to act as a key negative regulator of the Wnt signalling pathway by specifically mediating the depalmitoleoylation of Wnt proteins [3,4].

LP-922056 (**1**, Figure 1) is an orally active inhibitor of Notum recently reported by Lexicon Pharmaceuticals [5,6]. Their research with **1** has shown that Notum is a potential drug target for stimulating bone formation and treating osteoporosis [7]. However, although **1** demonstrates low plasma clearance, the structure contains an essential carboxylic acid and acids tend to have low passive brain penetration [8-12]. We are investigating the role of Notum in modulating Wnt signalling in the central nervous system (CNS) [13] and wished to establish if **1** would serve as a peripherally restricted control compound. Hence, we required a synthetic route to **1** that could be reliably and safely performed on large scale.



1

Figure 1: Chemical structure of Notum inhibitor LP-922056 (**1**).

A synthesis of **1** has been published in the patent literature [6], although many of the experimental procedures are described in terms of 'general procedures' which do not seem to work well when applied to **1** which contains functional groups sensitive to certain reagents employed (vide infra). An improved synthetic route should allow **1** to become more readily available as a chemical tool to explore the fundamental biology of Notum and build target validation to underpin new drug discovery programs for non-CNS disease.

Results and Discussion

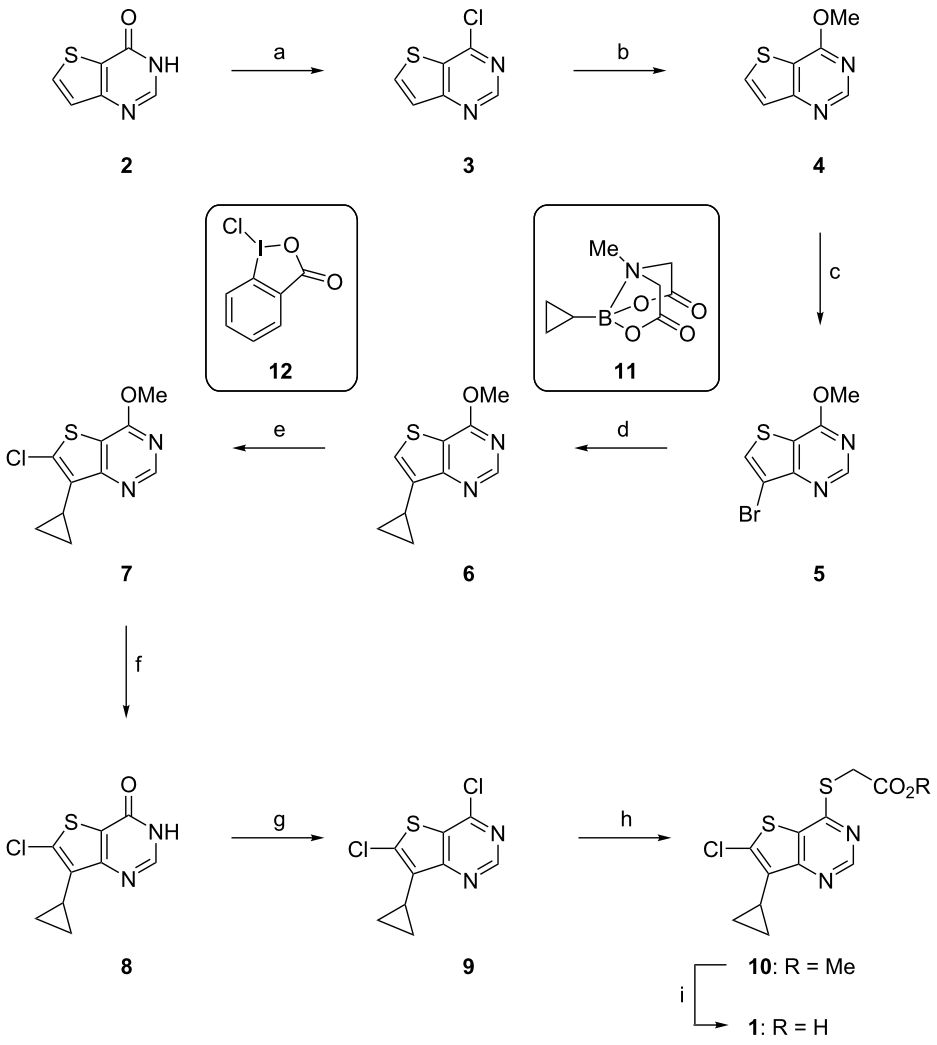
Improved synthesis of 1; first generation

Our first complete synthesis of **1** is presented in Scheme 1 (see Supporting Information File 1 for experimental procedures and characterisation data).

This short sequence starts with 4-chlorothieno[3,2-*d*]pyrimidine (**3**), which is readily available from commercial suppliers, and generally follows published procedures [5,6] but with key modifications to increase yields/selectivities and significantly improve ease of purification of key intermediates. Our modifications include: (1) the introduction of the C7-cyclopropyl group was most effectively achieved with a Suzuki–Miyaura cross-coupling reaction with MIDA-boronate **11** (**5** → **6**); and (2) C6 chlorination was performed with 1-chloro-1,2-benziodoxol-3-one (**12**) (**6** → **7**) as a mild selective electrophilic chlorination agent.

4-Chlorothieno[3,2-*d*]pyrimidine (**3**) was either purchased or prepared from thieno[3,2-*d*]pyrimidin-4(*3H*)-one (**2**) by C4 chlorination with oxalyl chloride/DMF following the method of Mitchell et al. [14]. Treatment of **3** with NaOMe displaced the C4–Cl to give **4** in a good yield as described by Atheral et al. [15]. Thieno[3,2-*d*]pyrimidine **4** is now suitably functionalised for the introduction of the C7-cyclopropyl group, the C6-chlorine atom and elaboration of the thioacetic acid moiety at C4.

Electrophilic bromination at C7 with *N*-bromosuccinimide gave **5** as the major regioisomer reproducibly on 100 mmol scale in modest yield (41–48%). This proved to be the least efficient step in our sequence and justified further optimisation (vide infra). Suzuki–Miyaura cross coupling of bromide **5** with cyclopropylboronic acid (2.5 equiv) produced **6** in good yield (62–89%) but the product required extensive chromatographic purification. We reasoned that switching from the boronic acid (*c*-PrB(OH)$_2$) to the corresponding MIDA-boronate **11** would improve the quality of the reagent and slow release of the active boron species during the course of the reaction would allow us to reduce the number of molar equivalents required to improve conversion [16]. Palladium-mediated cross coupling of **5** with **11** (1.5 equiv) gave **6** in reproducibly high yield (ca. 95%) when performed on gram scale. However, when performed on larger scale (9.3 g), the reaction stalled after ca. 90% conversion and

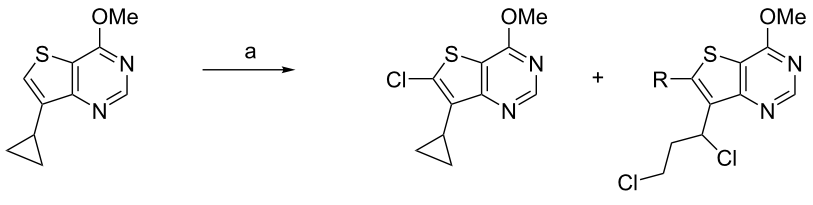


Scheme 1: Synthesis of LP-922056 (**1**). Reagents and conditions[a]: (a) $(COCl)_2$ (3.3 equiv), DMF, CH_2Cl_2, 55 °C , 16 h, 63–78%; (b) NaOMe (5 equiv), 1,4-dioxane, 0 °C then rt, 16 h, 92–93%; (c) NBS (1.1 equiv), AcOH/MeCN (1:100), 85 °C, 16 h, 41–48%; (d) **11** (1.5 equiv), $Pd(PPh_3)_2Cl_2$ (5 mol %), K_3PO_4 (6 equiv), PhMe/H_2O (3:1, 0.25 M), 100 °C, 16 h, 94–95%; (e) 1-chloro-1,2-benziodoxol-3-one (**12**, 1.5 equiv), DMF, 50 °C, 16 h, 77–94%; (f) HCl (12 M, 40 equiv), 70 °C, 16 h; (g) $POCl_3$ (20 equiv), 90 °C, 16 h, 81% over 2 steps; (h) $HSCH_2CO_2Me$ (**13**, 1.2 equiv), NEt_3 (2.1 equiv), MeOH, 0 °C to rt, 16 h, 84%; (i) NaOH (1 M, 2 equiv), THF, 0 °C, 1 h, then HCl (1 M), 0 °C, 30 min, 98%. [a]These reactions have been performed several times but have not been systematically optimised. Yields are the ranges obtained from repeated reactions. DMF, *N,N*-dimethylformamide; NBS, *N*-bromosuccinimide; THF, tetrahydrofuran.

addition of extra catalyst $Pd[(PPh_3)_2Cl_2]$ and/or **11** failed to drive the reaction to completion. We found the most efficient way to complete the reaction conversion was to isolate the crude product (mostly **6**) and subject this material to a repeat reaction; this procedure gave **6** in a good yield 95% and simplified purification.

With multigram quantities of **6** in hand, attention was turned to the C6 chlorination step. Unfortunately, despite this reaction being reported in the literature, there were no experimental details for this specific transformation as only a 'general procedure' was described [6]. Our attempts to use this procedure with *N*-chlorosuccinimide (NCS) as the chlorinating agent gave poor yields of the desired product **7** (ca. 15–32%) due to competing ring opening reactions of the 7-cyclopropyl group (Scheme 2). Clearly, a better procedure was required.

A large number of electrophilic chlorinating reagents for the direct chlorination of aromatic rings have been reported [17]. Recently, Xue et al. described the electrophilic chlorination of arenes and heterocycles by 1-chloro-1,2-benziodoxol-3-one (**12**) [18,19]. The hypervalent iodine(III) reagent **12** is reported to be a mild and effective reagent for the chlorination of nitrogen containing heterocycles which is easy to prepare and is



Scheme 2: Chlorination of **6** with *N*-chlorosuccinimide (NCS). Reagents and conditions: (a) NCS (1.2 equiv), AcOH, 55 °C, 7 h, 15–32%.

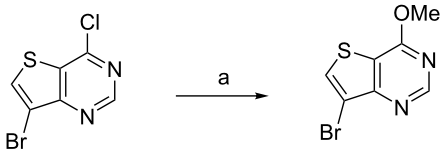
air- and moisture-stable. The scope of published substrates includes chlorination of 7*H*-pyrrolo[2,3-*d*]pyrimidines and we wished to see if we could extend the scope to include sulphur containing heterocycles such as thieno[3,2-*d*]pyrimidines (e.g., **6**). It was also important to explore if **12** would efficiently chlorinate **6** at the less activated C6 position in the presence of the C7 cyclopropyl group.

Treatment of **6** with **12** (1.5 equiv) in DMF at 50 °C for 16–24 h gave the desired chloro product **7** in 77–94% isolated yield. Analysis of the crude reaction mixture showed only trace amounts of cyclopropyl ring opened products such as **14**/**15** as detectible by LC–MS. Hence, **12** proved to be a far superior reagent, when compared to NCS, for the C6 chlorination of thieno[3,2-*d*]pyrimidine **6** (i.e., **6** → **7**).

Completion of our synthesis of **1** followed established procedures although it proved expedient to carry material through several of these later steps without the need for extensive purification beyond a simple work-up procedure (**7** → **8** → **9**). Activation of C4 was accomplished by a two-step procedure of acid hydrolysis of the C4-OMe of **7** to give thieno[3,2-*d*]pyrimidin-4(3*H*)-one **8**, followed by chlorination with $POCl_3$ to give **9**. Finally, nucleophilic displacement of the C4–Cl of **9** by methyl thioglycolate (**13**) gave ester **10** which was hydrolysed with NaOH to afford **1**. This route has been reliably performed on large scale to produce multigram quantities of **1** in good efficiency (total yield over 8-steps from **3**: 18–26%) and high purity (>99%).

Improved synthesis of 1; second generation

A shorter synthesis was then developed by accessing bromide **5** by an alternative route. The low-yielding C7 bromination of **4** with NBS to give **5** as described above (Scheme 1, step c) was avoided by starting with 7-bromo-4-chlorothieno[3,2-*d*]pyrimidine (**16**) which is readily available from commercial suppliers. Treatment of **16** with NaOMe displaced the C4–Cl to give **5** in good yield on 10 g scale (Scheme 3). Even though **16** is somewhat more expensive than **2** or **3** per unit cost (by ca. 5-fold), this updated route shortens the sequence to just 7 steps from **16** and improves the overall yield to 40–50 %.



Scheme 3: Improved synthesis of **5**. Reagents and conditions: (a) NaOMe (5 equiv), 1,4-dioxane, 0 °C then rt, 16 h, 84%.

Mouse pharmacokinetics for 1

Assessment of **1** in mouse liver microsomes (MLM) showed excellent metabolic stability (Cl_i 1.0 μL/min/mg protein) which predicts for low clearance in vivo. Binding to mouse plasma proteins (mPPB) was very high with percent unbound drug (f_u) of just 0.1%; this mPPB value can be used to calculate free drug concentrations from measured drug levels in plasma taken during in vivo experiments. The high mPPB is entirely consistent with the physicochemical properties of **1** as a lipophilic acid (mw 300; *c*Log*P* 3.1; *c*pK_a 3.1).

Pharmacokinetic (PK) data for **1** was generated in vivo in mouse to evaluate brain penetration (Table 1; Figure 2) (see

Table 1: Mouse pharmacokinetic data for **1**; oral (p.o.) dose at 10 mg/kg.[a]

PK Parameter	Plasma	Brain
$T_{1/2}$	8.8 h	7.1 h
T_{max}	2.0 h	2.0 h
C_{max}	35,400 ng/mL	500 ng/g
$AUC_{0\rightarrow t}$	303,000 h·ng/mL	3,700 h·ng/g
$AUC_{0\rightarrow \infty}$	354,000 h·ng/mL	4,080 h·ng/g

[a]Male fed CD1 mouse; suspension formulation in 0.1% Tween80 in water; *n* = 3 per time point; terminal blood and brain levels measured at seven time points: 0.17, 0.50, 1, 2, 5, 7.5 and 24 h.

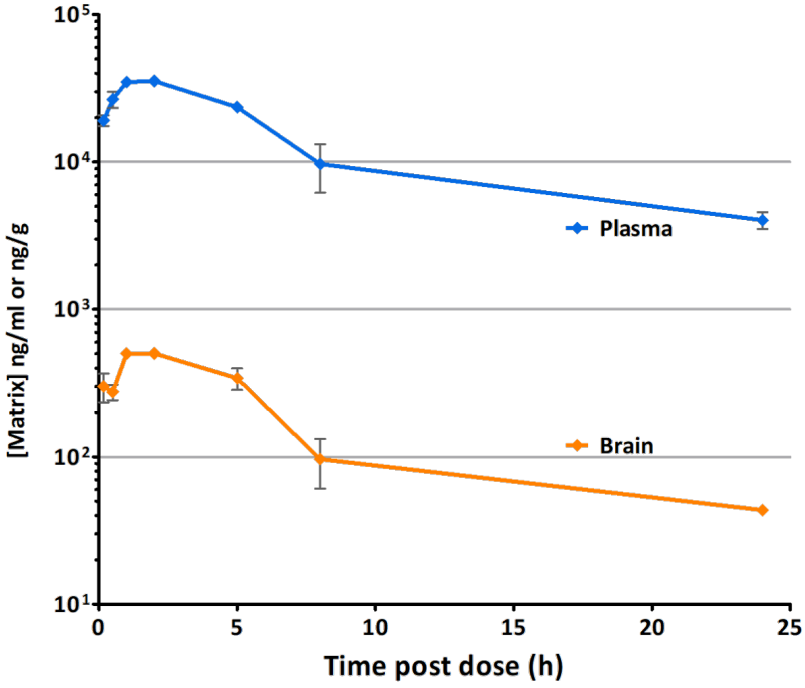


Figure 2: Concentrations of **1** in mouse following oral administration (p.o.) at 10 mg/kg.

Supporting Information File 1, Tables S1–S3). The route of administration and dose were selected to most closely match relevant published mouse disease model studies [5,7]. Following single oral dose (p.o.) of 10 mg/kg, plasma exposure was high and plasma clearance was low relative to liver blood flow resulting in a plasma elimination half-life of 8.8 hours. The plasma parameters from these mouse PK experiments (C_{max} and AUC) are consistent with published preclinical PK data [5].

CNS penetration of **1** was very low with a brain/plasma concentration ratio of ≈0.01 at all time points measured and was also 0.01 based on $AUC_{(0\rightarrow\infty)}$. At this level of exposure, a significant proportion of the compound detected in brain samples is likely to have arisen from residual blood in the brain tissue.

Amide analogues of **1** to explore CNS penetration and future opportunities

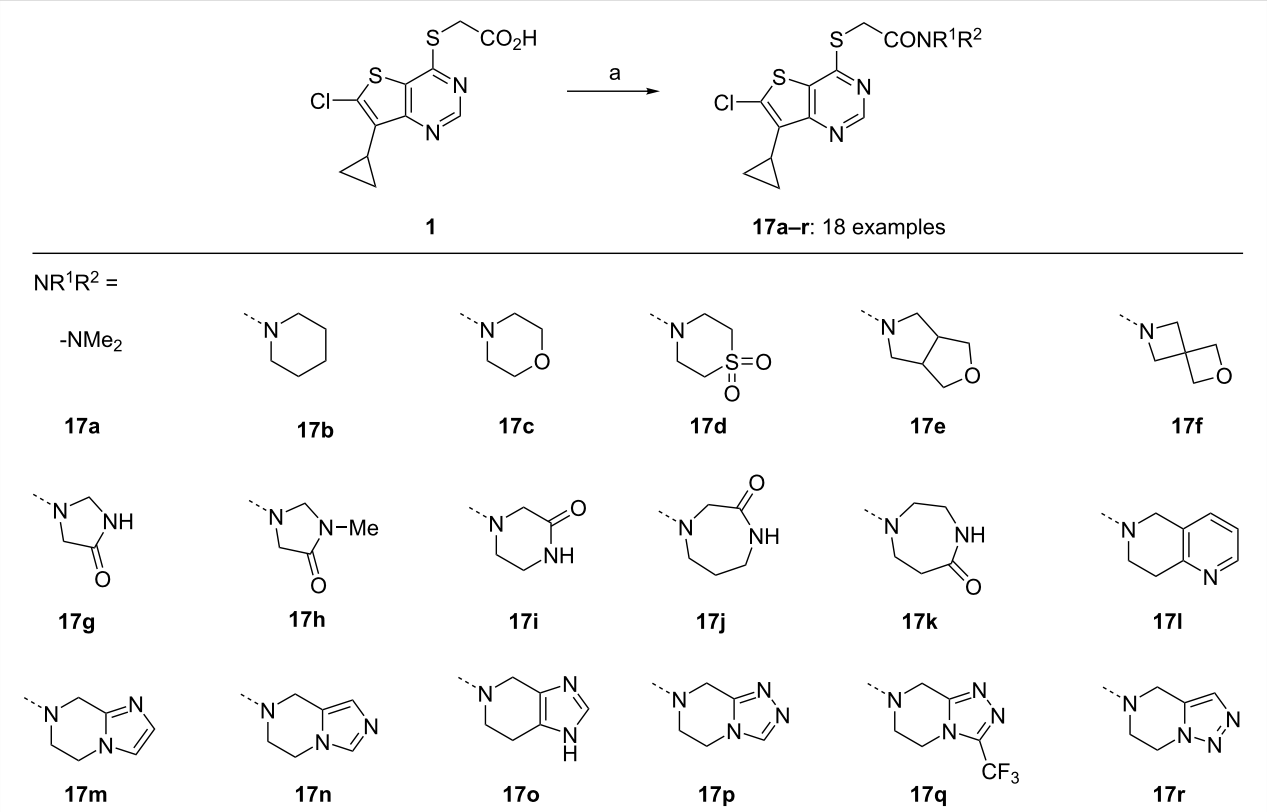
With a peripherally restricted control in hand, we elected to explore if **1** could be modified to deliver a CNS penetrant tool by capping off the acid as an amide. A small library of amides **17** were prepared from acid **1** by activation with HBTU and then subsequent reaction with the amine **18** (Scheme 4). These amides **17** were designed to have molecular properties (mw, *c*Log*P*, tPSA, HBD, pK_a) consistent with CNS drug-like space [20].

Although significant Notum inhibition activity could be achieved (IC_{50} < 100 nM), none of these specific amides demonstrated the required combination of sufficient MLM stability along with cell permeability as measured by transit performance across a MDCK-MDR1 monolayer without evidence of P-gp mediated efflux. This collection of data for amides **17a**–**r** is shared as 'open data' to assist others in evaluating these results with the objective of solving this challenge (see Supporting Information File 1, Table S4 and Supporting Information File 2). Our own efforts took us to an alternative chemotype [21].

Conclusion

An improved, scalable synthesis of Notum inhibitor **1** is reported. Key modifications include: (1) the introduction of the C7-cyclopropyl group was most effectively achieved with a Suzuki–Miyaura cross-coupling reaction with MIDA-boronate **11** (**5** → **6**); and (2) C6 chlorination was performed with 1-chloro-1,2-benziodoxol-3-one (**12**) (**6** → **7**) as a mild selective electrophilic chlorination agent. This 7-step route from **16** has been reliably performed on large scale to produce multigram quantities of **1** in good efficiency.

Pharmacokinetic studies in mouse showed CNS penetration of **1** is very low with brain/plasma concentration ratio of just 0.01 based on $AUC_{(0\rightarrow\infty)}$. Hence, **1** is unsuitable for use in



Scheme 4: Preparation of amides **17**. Representative reagents and conditions[a]: (a) HBTU (1.1 equiv), iPr_2NEt (2.5 equiv), DMF, rt, 15 min; then amine **18** (1.05 equiv). [a]These reactions were performed once for each amide and have not been optimised. DMF, *N,N*-dimethylformamide; HBTU, *O*-(1*H*-benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate.

models of disease where brain penetration is an essential requirement of the compound but would be an ideal peripherally restricted control. These data will contribute to the understanding of drug levels of **1** to overlay with appropriate in vivo efficacy endpoints, i.e., the PK-PD relationship. The full PK data set is presented and shared as 'open data'. This complete data set (along with others) will also assist with the creation of improved predictive pharmacokinetic models.

A small library of amides **17** were prepared from acid **1** to explore if **1** could be modified to deliver a CNS penetrant tool by capping off the acid as an amide. Although significant Notum inhibition activity could be achieved, none of these amides demonstrated the required combination of metabolic stability along with cell permeability without evidence of P-gp mediated efflux. The identification of a suitable analogue of **1** (or **17**) which combines Notum inhibition with CNS penetration would be a valuable chemical probe for investigating the role of Notum in disease models. This collection of data for amides **17a**–**r** is shared as 'open data' to assist others in evaluating these results with the objective of solving this challenge.

Supporting Information

(1) experimental procedures and characterisation data for **3**–**10** and **1**; (2) mouse PK data which includes: study design summary; plasma concentrations; brain concentrations; (3) Notum IC_{50} (nM), MLM Cl_i (μL/min/mg protein) and MDCK-MDR1 AB/BA P_{app} ($\times 10^{-6}$ cm/s) data for **17a**–**r**.

To view the NMR spectra, use the file within the pdata folder of Supporting Information File 3.

Supporting Information File 1

Experimental section, mouse pharmacokinetics and profiles of amides.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-15-271-S1.pdf]

Supporting Information File 2

Profiles of amides.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-15-271-S2.csv]

Supporting Information File 3
Raw NMR data files for compound LP-922056.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-15-271-S3.zip]

Acknowledgements

This work was supported by Alzheimer's Research UK (ARUK) and The Francis Crick Institute. The ARUK UCL Drug Discovery Institute is core funded by Alzheimer's Research UK (520909). The Francis Crick Institute receives its core funding from Cancer Research UK (FC001002), the UK Medical Research Council (FC001002), and the Wellcome Trust (FC001002).

We thank our colleagues Sarah Frew, Amy Monaghan, Fiona Jeganathan and Magda Bictash of the ARUK UCL DDI Screening and Pharmacology team for Notum inhibition data. We thank Abil Aliev and Kersti Karu at the UCL Department of Chemistry for spectroscopic and analytical services. Mouse PK studies were performed by Pharmidex (London, UK) and ADME studies reported in this work were performed by GVK Biosciences (Hyderabad, India) and Cyprotex (Macclesfield, UK).

ORCID® iDs

Nicky J. Willis - https://orcid.org/0000-0003-3245-5280
George Papageorgiou - https://orcid.org/0000-0002-8960-5836
David Steadman - https://orcid.org/0000-0003-4271-5525
Benjamin N. Atkinson - https://orcid.org/0000-0001-5511-9859
Paul V. Fish - https://orcid.org/0000-0002-2117-2173

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint doi:10.3762/bxiv.2019.70.v1

References

1. Komiya, Y.; Habas, R. *Organogenesis* **2008,** *4,* 68–75. doi:10.4161/org.4.2.5851
2. Palomer, E.; Buechler, J.; Salinas, P. C. *Front. Cell. Neurosci.* **2019,** *13,* No. 227. doi:10.3389/fncel.2019.00227
3. Kakugawa, S.; Langton, P. F.; Zebisch, M.; Howell, S. A.; Chang, T.-H.; Liu, Y.; Feizi, T.; Bineva, G.; O'Reilly, N.; Snijders, A. P.; Jones, E. Y.; Vincent, J.-P. *Nature* **2015,** *519,* 187–192. doi:10.1038/nature14259
4. Zhang, X.; Cheong, S.-M.; Amado, N. G.; Reis, A. H.; MacDonald, B. T.; Zebisch, M.; Jones, E. Y.; Abreu, J. G.; He, X. *Dev. Cell* **2015,** *32,* 719–730. doi:10.1016/j.devcel.2015.02.014
5. Tarver, J. E., Jr.; Pabba, P. K.; Barbosa, J.; Han, Q.; Gardyan, M. W.; Brommage, R.; Thompson, A. Y.; Schmidt, J. M.; Wilson, A. G. E.; He, W.; Lombardo, V. K.; Carson, K. G. *Bioorg. Med. Chem. Lett.* **2016,** *26,* 1525–1528. doi:10.1016/j.bmcl.2016.02.021
6. Barbosa, J.; Carson, K. G.; Gardyan, M. W.; He, W.; Lombardo, V.; Pabba, P.; Tarver, J., Jr. Inhibitors of Notum pectinacylesterase and methods of their use. U.S. Pat. Appl. US 2012/0065200 A1, March 15, 2012.
7. Brommage, R.; Liu, J.; Vogel, P.; Mseeh, F.; Thompson, A. Y.; Potter, D. G.; Shadoan, M. K.; Hansen, G. M.; Jeter-Jones, S.; Cui, J.; Bright, D.; Bardenhagen, J. P.; Doree, D. D.; Movérare-Skrtic, S.; Nilsson, K. H.; Henning, P.; Lerner, U. H.; Ohlsson, C.; Sands, A. T.; Tarver, J. E.; Powell, D. R.; Zambrowicz, B.; Liu, Q. *Bone Res.* **2019,** *7,* 2. doi:10.1038/s41413-018-0038-3
8. Manallack, D. T. *Perspect. Med. Chem.* **2007,** *1,* 25–38. doi:10.1177/1177391x0700100003
9. Wager, T. T.; Chandrasekaran, R. Y.; Hou, X.; Troutman, M. D.; Verhoest, P. R.; Villalobos, A.; Will, Y. *ACS Chem. Neurosci.* **2010,** *1,* 420–434. doi:10.1021/cn100007x
10. Manallack, D. T.; Prankerd, R. J.; Yuriev, E.; Oprea, T. I.; Chalmers, D. K. *Chem. Soc. Rev.* **2013,** *42,* 485–496. doi:10.1039/c2cs35348b
11. Charifson, P. S.; Walters, W. P. *J. Med. Chem.* **2014,** *57,* 9701–9717. doi:10.1021/jm501000a
12. Di, L.; Rong, H.; Feng, B. *J. Med. Chem.* **2013,** *56,* 2–12. doi:10.1021/jm301297f
13. Atkinson, B. N.; Steadman, D.; Zhao, Y.; Sipthorp, J.; Vecchia, L.; Ruza, R. R.; Jeganathan, F.; Lines, G.; Frew, S.; Monaghan, A.; Kjær, S.; Bictash, M.; Jones, E. Y.; Fish, P. V. *Med. Chem. Commun.* **2019,** *10,* 1361–1369. doi:10.1039/c9md00096h
14. Mitchell, I. S.; Spencer, K. L.; Stengel, P.; Han, Y.; Kallan, N. C.; Munson, M.; Vigers, G. P. A.; Blake, J.; Piscopio, A.; Josey, J.; Miller, S.; Xiao, D.; Xu, R.; Rao, C.; Wang, B.; Bernacki, A. L. Akt protein kinase inhibitors. WO Patent WO2005/051304, June 9, 2005.
15. Atheral, J. F.; Hough, T. L.; Lindell, S. D.; O'Mahony, M. J.; Parsons, J. H.; Saville-Stones, E. A. Fungicides. U.S. Patent US6432964B1, Aug 13, 2002.
16. Gillis, E. P.; Burke, M. D. *J. Am. Chem. Soc.* **2007,** *129,* 6716–6717. doi:10.1021/ja0716204
17. Rogers, D. A.; Bensalah, A. T.; Espinosa, A. T.; Hoerr, J. L.; Refai, F. H.; Pitzel, A. K.; Alvarado, J. J.; Lamar, A. A. *Org. Lett.* **2019,** *21,* 4229–4233. doi:10.1021/acs.orglett.9b01414
18. Wang, M.; Zhang, Y.; Wang, T.; Wang, C.; Xue, D.; Xiao, J. *Org. Lett.* **2016,** *18,* 1976–1979. doi:10.1021/acs.orglett.6b00547
19. Matoušek, V.; Pietrasiak, E.; Schwenk, R.; Togni, A. *J. Org. Chem.* **2013,** *78,* 6763–6768. doi:10.1021/jo400774u
20. Rankovic, Z. *J. Med. Chem.* **2015,** *58,* 2584–2608. doi:10.1021/jm501535r
21. Fish, P. V. Development of Potent, Selective, CNS Penetrant Small Molecule Inhibitors of Notum to Potentiate Wnt Signalling. Presented at EFMC: XXVth International Symposium on Medicinal Chemistry; 2–6th September 2018; Ljubljana, Slovenia.

License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (https://www.beilstein-journals.org/bjoc)

The definitive version of this article is the electronic one which can be found at: doi:10.3762/bjoc.15.271

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Opening up connectivity between documents, structures and bioactivity

Christopher Southan[1,2]

Review

Open Access

Address:
[1]Deanery of Biomedical Sciences, University of Edinburgh, Edinburgh, EH8 9XD, UK and [2]TW2Informatics Ltd, Västra Frölunda, Gothenburg, 42166, Sweden

Email:
Christopher Southan - cdsouthan@hotmal.com

Keywords:
activity data; databases; drug discovery; chemical structures; protein targets

Beilstein J. Org. Chem. **2020,** *16,* 596–606.
doi:10.3762/bjoc.16.54

Received: 22 November 2019
Accepted: 12 March 2020
Published: 02 April 2020

This article is part of the thematic issue "Medicinal chemistry and drug discovery in an opening era".

Guest Editor: M. Todd

© 2020 Southan; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Bioscientists reading papers or patents strive to discern the key relationships reported within a document “D“ where a bioactivity “A” with a quantitative result “R” (e.g., an IC_{50}) is reported for chemical structure “C” that modulates (e.g., inhibits) a protein target “P”. A useful shorthand for this connectivity thus becomes DARCP. The problem at the core of this article is that the community has spent millions effectively burying these relationships in PDFs over many decades but must now spend millions more trying to get them back out. The key imperative for this is to increase the flow into structured open databases. The positive impacts will include expanded data mining opportunities for drug discovery and chemical biology. Over the last decade commercial sources have manually extracted DARCP from ≈300,000 documents encompassing ≈7 million compounds interacting with ≈10,000 targets. Over a similar time, the Guide to Pharmacology, BindingDB and ChEMBL have carried out analogues DARCP extractions. Although their expert-curated numbers are lower (i.e., ≈2 million compounds against ≈3700 human proteins), these open sources have the great advantage of being merged within PubChem. Parallel efforts have focused on the extraction of document-to-compound (D-C-only) connectivity. In the absence of molecular mechanism of action (mmoa) annotation, this is of less value but can be automatically extracted. This has been significantly accomplished for patents, (e.g., by IBM, SureChEMBL and WIPO) for over 30 million compounds in PubChem. These have recently been joined by 1.4 million D-C submissions from three major chemistry publishers. In addition, both the European and US PubMed Central portals now add chemistry look-ups from abstracts and full-text papers. However, the fully automated extraction of DARCLP has not yet been achieved. This stands in contrast to the ability of biocurators to discern these relationships in minutes. Unfortunately, no journals have yet instigated a flow of author-specified DARCP directly into open databases. Progress may come from trends such as open science, open access (OA), findable, accessible, interoperable and reusable (FAIR), resource description framework (RDF) and WikiData. However, we will need to await the technical applicability in respect to DARCP capture to see if this opens up connectivity.

Introduction

This article assesses a key aspect of data sharing that has the potential to accelerate the progress and impact of medicinal chemistry. To achieve this the community needs to increase the outward flow of experimental results locked-up in millions of published PDFs into structured open databases that explicitly capture the connectivity between structures, documents and bioactivity results. But isn't there enough of this out there already? This can be answered in two parts. The first is that a conservative estimate of the capture backlog would be at least two-fold more data still entombed in PDFs that is not currently indexed in database records. The second part is the imperative to enable open science data mining at all scales. This applies not only to individual documents (i.e., small data) but scaling up to all papers and patents (i.e., big data). The potential of the latter is huge, especially since artificial intelligence (AI) is being increasingly applied to knowledge distillation. This report will outline the principles of connectivity capture, selected sources, progress, impediments and prospects for their amelioration.

Review

Defining terms

It is necessary to outline the topics covered:

Medicinal chemistry: As directed towards drug discovery this needs no introduction. However, in the broader context of bioactive chemistry, it becomes indivisible from the related domains of chemical biology (directed towards mechanistic insight rather that direct drug discovery), enzymology, pharmacology, and toxicology in addition to the development of insecticides or herbicides.

Connectivity: This term is used for an explicit link (e.g., a URL) between a published document and the chemical structures specified therein. Implicit is not only manual navigation (e.g., link-clicking) but also that such connectivity can be made machine-readable and thus computationally interrogated at large scale via an application programming interface (API) or a resource description framework (RDF).

Papers as documents: This typically refers to research papers from journals but increasingly needs to encompass their associated supplementary data. Note also that by far the majority of medicinal chemistry, biological chemistry and pharmacology papers are still behind subscription paywalls. However, the full-text for some of them is not only open but also available to be mined in both PubMed Central (PMC) [1] and European PubMed Central (EPMC) [2]. Connectivity can extend to other document types such as review articles and vendor catalogues. In this article the main document type referred to will be the PubMed identifier (PMID). These have open abstracts and are also indexed in the digital object identifier system (DOI). However, significant numbers of papers in the bioactive chemistry domain (including preprints) may be DOI-only.

Patents as documents: Academics tend to overlook that patents a) include several fold more medicinal chemistry than papers b) appear years earlier c) most academic drug discovery operations apply for them d) they include a proportion of high quality data that never appears in journals e) they can be text-mined and d) consequently, over 30 million structures have entered PubChem via automated extraction [3].

Non-document sources: While this article has to be restricted to documents an increasing amount of drug discovery data is beginning to surface on the web that may never be instantiated in document form. Although this started with PubChem Bioassay as far back as 2004, the more recent proliferation is via open-notebook science. More projects are using open electronic laboratory notebooks (ELNs) that are not only accessible to anyone by web browsing, but also, crucially, crawled by Google and indexed for chemistry searching [4,5].

Structures: A necessary focus of this article will be traditional small-molecule chemistry that is not too far outside the rule-of-five lead-like property space. In terms of connectivity antibodies, other protein biotherapeutics, as well as large peptides or polynucleotides, are also important to encompass. However, capture into structured records is more challenging for these larger therapeutic modalities than for small-molecules that can be merged on the basis of chemistry rules. Notwithstanding, space limitations mean that non-small molecule connectivity is out of scope for this article.

Bioactivity: This covers a wide spectrum of assay read-outs but with a focus on in vitro, in cellulo, in vivo and in clinico. Ideally this should also include low or inactive analogues which are crucial for SAR elucidation but documents are (understandably) biased towards positive results.

Open: As the theme of this special issue this term will doubtless be expanded on in other articles. However, brief qualification in the context of this work is necessary. Regardless of licensing complications, open is taken here to mean public data sources accessible via a web browser (signing in may be an impediment but not a stopper). These are thus distinct from commercial offerings where access has to be purchased.

Relationship representation

As outlined in the abstract the connectivity between documents, structures and bioactivity can be expressed in shorthand as "D-A-R-C-P" (DARCP). This is shown schematically in Figure 1.

The entity specifications can be adapted to different use-cases. For example, the substitution of "P" with target "T" can be used where "T" is a cell or a microorganism. Another example would be where an SAR series can be represented as a multiplexed set of one-A-to-many-R-C. It can also be extended to "D-A-R-C-L-P" where L refers to the explicit location references for C in the document (e.g., "compound **10b**" in a paper or "example 503" in a patent). However, as a formalism for bioactivity there are exceptions and mechanistic nuances that do not fit a DARCP simplification. An example would be heparin (GtoPdb ligand 4214). This could be a commercial partially purified extract of 1200–1500 Mw which consequently does not have a defined chemical structure as "C". However, as a curatorial expedient, the chemically defined form (as PubChem CID 22833565 with 1040 Mw) has been annotated (even though the sodium salt is the active form in vivo). Note also that while formally "P" in the heparin case is SERPINC1 (ATIII) the mechanism is an indirect one involving the activation of binding to F2 (activated thrombin) for inhibition. Another problematic example is mechanism-based covalent inhibition where the time dependence of IC50 for "A" is not captured.

The structured capture of DARCLP by curation (or at least DARCP) has very high value from the additional relationships that can be explored via the entities and attributes as outlined below:

- Documents: clustering by content relatedness, position within citation networks, connections via authors or institutional affiliations.
- Assays: classified by various assay ontologies.
- Results: log transformations (e.g., pIC_{50} or pK_i) for potency ranking and implicit molecular mechanism of action (mmoa), (e.g., where A-R indicates C to be a potent inhibitor of P).
- Compounds: a full range of cheminformatic analysis including 2D or 3D clustering, property prediction and chemical ontology assignments.
- Proteins: a full range of bioinformatic analysis including gene ontology (GO) assignments, pathway annotation, structural homology, disease associations and genetic variation (e.g., for target validation).

Those cases where the link is only compound-to-document can be referred to as D-C (or c2d). These have become available in a large excess over full DARCLP since they are technically easier to obtain and can be automated to a usable level of specificity. This needs the introduction of the intuitive concept of "aboutness" (ABNS). The title of a document from which D-C could be extracted usually includes an explicit ABNS statement. For example, the code name of a lead compound would be what a medicinal chemistry journal article would be "about". In the same way, the ABNS of a clinical pharmacology journal article some years later could be describing the clinical trial results for the identical structure which, by then, could have an international non-proprietary name (INN). However, the extraction of multiple compounds (i.e., one-D-to-many-Cs) immediately becomes problematic in the absence of full relationship chains. For example, the medicinal chemistry article may describe the testing of a useful set of analogues for SAR but (as is usually the case) the A-R-P data was not extracted.

At this point we need to introduce the additional concept of name-to-structure (n2s). This is an important determinant of both ABNS and D-C utility. Using the example again from a paper this would mean that both the code name and the INN

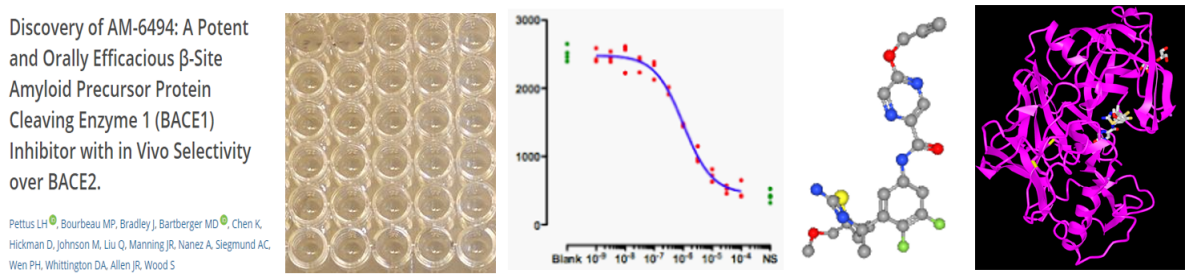


Figure 1: A schematic of the document > assay > result > compound > protein target relationships, D-A-R-C-P [microplate photo adapted from https://commons.wikimedia.org/wiki/File:Microtiter_plate.JPG by Jeffrey M. Vinocur, source licensed under the Creative Commons Attribution Generic 2.5 license (CC-BY 2.5), https://creativecommons.org/licenses/by/2.5/].

would be included in the D-C capture record (i.e., n2s) even if 50 analogues were also tested. Other examples that present particular ABNS problems are review articles, synthetic chemistry papers and patents. A review could exemplify 20 lead compounds all with different company code numbers and/or INNs, an extended synthesis report could give rise to 200 D-C records and a patent could have over 500. Discerning the ABNS for patents can be especially problematic since frankly obfuscatory titles and abstracts are common (e.g., "Novel Compounds").

The "hamburger" problem

This can be summarised by the following (unattributed) quote "We have spent millions putting data into the literature but now have to spend millions more getting it back out". This alludes to entombing the DARCP "meat" within a PDF "hamburger". The paradox is that electronic text formats typically used for drafting papers are machine-readable (certainly with modern parsing techniques). However, this is systematically obviated by the PDF conversion. For example, a chemist may have SMILEs and/or InChIs in their ELN and/or molfiles in an institutional data repository. However, they have to convert this to a ChemDraw proprietary file format in order to render the structural image that eventually appears in the PDF. This means getting the structures "back out" for database capture needs either manual re-sketching or use of an image-to-structure (i2s) tool such as optical structure recognition (OSRA) [6], both of which are error-prone processes.

The common practice of including tables of Markush representations, while they improve SAR readability, makes the extraction problem worse. While most medicinal chemistry journals will include IUPAC names in the synthesis descriptions, these also have to be pulled "back out" of the PDF. This can be done via PDF-to-text optical character recognition (OCR) or curated by pasting across to the Open Parser for Systematic IUPAC Nomenclature (OPSIN) tool [7]. Here again, both the automatic and manual procedures are error prone. Locally-stored SAR data from an Excel sheet or an ELN can be used to populate draft manuscripts (and with lower error rates) but the irony is conversion to PDF (i.e., entombment) makes the ARC in result tables more difficult to extract.

A specific example of the problem can be given for a 2017 article on new antimalarial compounds entering development [8]. Because the chemistry representations were restricted only to images in the PDF a blog post was necessary to manually map the structures to PubChem identifiers [9]. The MyNCBI link to the 16 CID entries given at the top of the blog post is still live (indicating reassuring persistence for this system after four years). While this initial connectivity was only D-C (and where D was a review article rather than a primary activity report) this example had an important sequel. During the curation of the new Guide to Malaria Pharmacology 14 of the 16 compounds now have full DARCLP annotation where D is the primary activity report, P is the *Plasmodium* target and activity values against the parasite are included in the records [10].

Commercial capture

Since this report is about open connectivity it might not seem pertinent to review commercial resources. However, a brief assessment of these is relevant in several contexts. The first is that, despite occasional use of the adjective "proprietary" in their descriptions, the primary content of commercial databases is almost entirely derived from open sources. Notwithstanding, they capture, curate, annotate, collate, integrate and index this in value-added ways (including user-friendly query front-ends and customer-specific APIs) to justify subscription costs. The second aspect is that by virtue of being able to apply more internal resources than open databases, their statistics give some indication of where the practical upper limits might lie. The third aspect is that they can give insights into the challenges of extraction, although technical details of how this is done are sparingly presented externally.

The largest relevant commercial source is CAS-SciFinder [11] While the mmoa may be indexed (i.e., providing C-P mappings) it does not include a complete DARCP capture. Consequently, this has to be classified as primarily D-C-only source. By November of 2019 SciFinder reached 157 million unique organic plus inorganic substances, having passed 100 million in June 2015. While some of these are virtual structures (i.e., never been synthesised) this large enterprise (with over 4,500 employees according to LinkedIN) has the de facto largest searchable collection of small-molecule structures extracted from papers, patents and other sources. A presentation from 2016 declared that in the first 7 months of that year ≈10.5 million substances were extracted from ≈0.5 million patents and ≈1.0 million documents. In addition, ≈75% of current novel structures are from patents. However, the 157 million is exceeded by the latest public UniChem release of just under 160 million [12]. In addition, a 2019 scaffold diversity analysis stringently filtered the CAS collection down to only ≈30 million compounds with direct links to literature and patents [13]. Since its first release in 2009 Elsevier Reaxys has emerged as another large-scale D-C capture endeavour, the statistics and search characteristics of which have recently been compared with SciFinder [14]. It has reached 31 million structures but also subsumes PubChem which brings it up to 105 million.

The two leading commercial sources that capture DARCLP at scale are the Global Online Structure Activity Relationship Database (GOSTAR) [15] from Excelra (formerly GV000Bio)

and Elsevier Reaxys Medicinal Chemistry [16]. The current statistics for these are shown in Table 1.

The GOSTAR numbers have a more detailed breakdown in a paper from 2013 (see Table 1 in that reference) which includes the calculated averages of 12 compounds per paper and 43 per-patent [17]. GOSTAR's compound total has doubled in the intervening six years but the extraction averages and ratio of compounds from papers: patents of ≈1:2.7 recorded in 2013 are likely to be similar. Comparable metrics for RMC curation have not been disclosed so it remains unclear what procedural differences that might explain their considerably larger activity, target and document counts compared to GOSTAR but connected to a million less compounds. Notwithstanding, using nominally the same medicinal chemistry corpus the extracted chemical structure ratios between SciFinder, Reaxys, GOSTAR and RMC are very approximately 30:30:8:7. Several technical differences may explain these ratios but the most important is the primary focus of the latter two on full DARCLP and SAR capture rather than just D-C. This selectivity in the choice of which journals and patents are curated, maintains the quality of target and activity mappings.

Public DARCP resources

The first relevant web-instantiated curated resource, BindingDB, was published in 2001 [18]. This was followed by the IUPHAR Ion Channels Compendium of papers in 2003. This had developed into the IUPHAR-DB website by 2009 and was updated to the current IUPHAR/BPS Guide to Pharmacology (GtoPdb) by 2012 [10]. That same year also saw the first ChEMBL publication for which the website was live by 2010 [19]. All three of these resources focus on expert-curated DARCP extractions from the literature. In addition, PubChem, first appearing in 2004 has now become the de facto global hub for DARCP because all the three databases above submit their structures that are integrated with ≈700 other sources [20]. Comparative statistics of the four are shown below in Table 2.

As for the commercial sources, comparing content statistics between databases is not straightforward because the numbers in Figure 2 were generated in slightly different ways. Not all the nuances can be addressed here but some salient ones can be pointed out. Moving across the columns there is an element of circularity in the compounds. The first reason is that ChEMBL subsumes 0.53 million compounds from confirmed PubChem BioAssays and 1.3 million curated from papers. The second reason is that BindingDB and ChEMBL have a reciprocal mirroring collaboration where BindingDB subsumes the protein target assay results from ChEMBL and the latter subsumes BindingDB patent extraction data (e.g., the 137,000 compounds in release 25). This is separated from their total data counts in rows three and four. It also means that the overlap of compound structures, target and document identifiers between the two sources have extensive circularity (but some are independently curated). The PubChem figures for bioactivities seem large because these are factored by substances not compounds, whereas ChEMBL (as the dominant contributor to PubChem BioAssay) collapses their assay counts to compounds. For GtoPdb the lower count reflects the curation of mainly lead compounds with curated binding constants from papers.

The 18000 targets in PubChem include automated assignments that result in an element of over-counting. Those in the other three sources are classified manually and have species-specific cross-references in UniProt [26]. These give the following human Swiss-Prot counts of 3644, 2585 and 1457 for ChEMBL, BindingDB and GtoPdb, respectively. We can see the document counts for the curated sources in column five of

Table 1: Statistics of GOSTAR (top row, from their website [15]) and RMC (from the information sheet [16]).

compounds (millions)	bioactivities (millions)	binding assays (millions)	targets (1000s)	papers (1000s)	patents (1000s)
7.8	9.7	8.7	9	191	76
6.8	35.2	–	27	370	133

Table 2: Content statistics for three DARCP sources and PubChem.

database	reference	compounds	nioactivities	targets	papers	patents
PubChem (9/19)	[21]	96 million	268 million	18000	14.2 million	3.2 mil
ChEMBL(release 25)	[22]	1.9 million	15.5 million	124000	72000	–
BindingDB (9/19)	[23]	772000	1.7 million	52000	29000	–
BindingDB patents	[24]	225000	406000	1000	–	3000
GtoPdb (2019.4)	[25]	75000	17000	20000	11000	600

Table 2. From the ChEMBL release notes their literature extraction average out at ≈15 compounds-per-document (n.b., the majority will have ARCP connectivity but some have only non-bioactivity A-R data such as plasma clearance).

Content overlaps and differences

Despite differences in the way their internal statistics are computed, standardised content comparison between open databases can use outputted lists for D,C and P (comparing A-R is not so straightforward). The intersects and differences between these are shown in the series of three Venn diagrams (Figures 2–4). See also Supporting Information File 1 for technical details on how these were prepared.

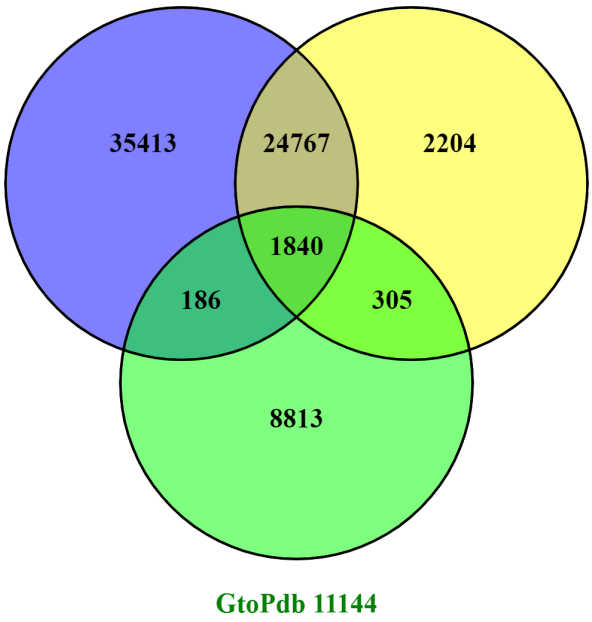


Figure 2: PMID content with totals appended to each segment. Those for ChEMBL were downloaded from European PubMed Central (EPMC) via the query (HAS_CHEMBL:y). For BindingDB the list was provided with courtesy of Michael Gilson. For GtoPdb the list was downloaded via the PubMed connectivity for the SIDs. The OR union is 73500 PMIDs (but all three sources also curate a proportion of DOI-only documents).

While there are technical caveats, we can briefly consider the implications of Figures 2–4. The PMID capture in Figure 2 shows a pattern of intersects and differences that is to some extent reflected for the other entities also. Each indicates some unique capture but ChEMBL and BindingDB overlap for ≈25000 papers. Despite being the smallest of the three, GtoPdb shows proportionally more unique PMIDs. This is substantially due to the curators adding additional references into the SID records beyond those from which the binding data were extracted (e.g., in vivo and clinical reports published after the

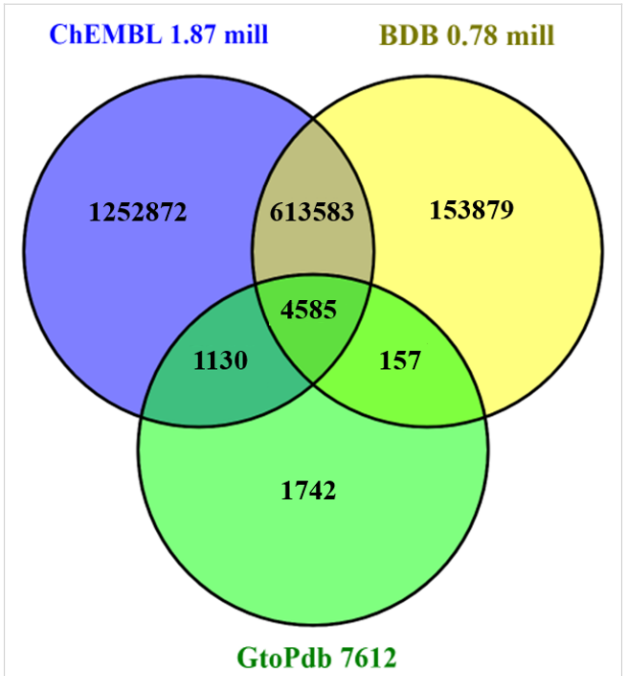


Figure 3: Chemistry content comparison for the three sources with totals appended to each segment, selected and downloaded as PubChem CID lists. The OR total is 2,033,127. Note that the uniqueness in the Venn sectors is just between these three. Overall uniqueness in PubChem as whole is 266000, 49000 and 143 for ChEMBL, BindingDB and GtoPdb, respectively.

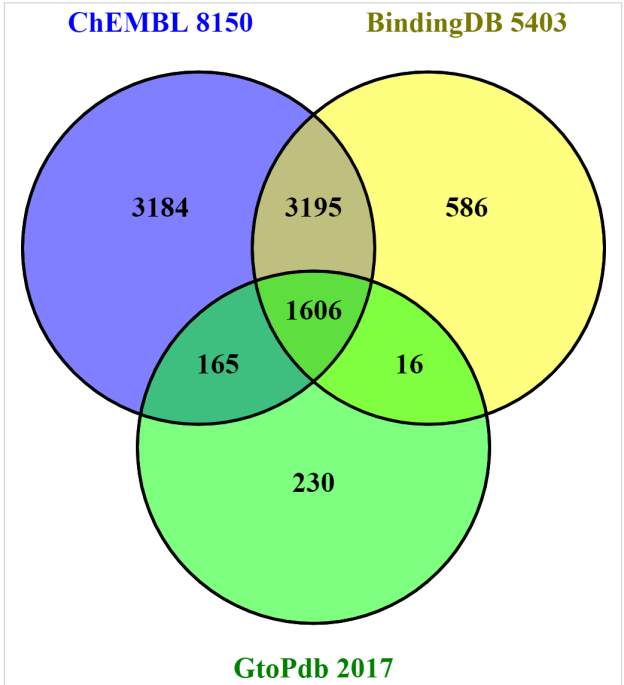


Figure 4: Target comparisons with totals in each segment. These were downloaded as UniProt ID lists selected via the chemistry cross-references in UniProt. These include all species but, as expected, human and rodent predominate in each case. The total of human proteins covered by the three is 3745 representing 18% of the UniProt 20,365 proteome.

initial in vitro results). Notably, the public total from all four of ≈75000 is less than 50% of the journal document counts declared by the two commercial sources (Table 1). While the limited resources of the public sector are clearly a factor, it would be informative to know explicitly what was behind the differences. Journal selectivity is likely to be dominant but other factors may come into play.

The chemistry content in Figure 3 shows similar disproportionation with ChEMBL, as expected, dominating unique content at over 1.2 million. While this is skewed by the BioAssay subsumation of ≈0.5 million, most will be a consequence of extracting ≈35000 unique PMIDs. For BindingDB most of their 153000 unique structures are from the ≈200000 protein-ligand binding data points that were curated from 1,100 US Patents during 2019 (n.b., these will eventually be subsumed into ChEMBL release 26). We can further rationalise the proportionality between compounds and PMIDs by noting that GtoPdb extract on average ≈1 lead compound per paper, ChEMBL ≈14 per paper with BindingDB extracting similar numbers from papers but ≈40 per patent.

The differences in target coverage (i.e., as "P" in DARCP) shown in Figure 4 are noteworthy and persist despite the ChEMBL/BindingDB selective mirroring. As for PMID coverage it would be useful to know which types of selectivity were responsible for this divergence in connectivity. For BindingDB some unique proteins are likely to be patent-only but exploring further causes of complementary target coverage are outside the scope of this work.

Journals connecting to PubChem

As anomalous as it may seem, no individual journals have put in place a direct feed of author-specified DARCP into PubChem BioAssay (or any other database for that matter). Historically, four journals have initiated D-C feeds in PubChem but two of these, *Prous Science Drugs of the Future* and *Nature Communications*, ceased in 2012 and 2014 respectively. This has left only *Nature Chemical Biology* and *Nature Chemistry* as still active with 12481 and 15276 author-specified CID structures respectively (plus some on-hold submissions) but the latter journal does not typically include bioactivity reports. Some Elsevier journals do list CIDs in their abstracts but without submitted links.

One journal that has pioneered a first approximation to DARCP flow into PubChem is the *British Journal of Pharmacology* (although the links are technically indirect) [27]. The annotation task was initially done by editors but since 2016 authors have been incorporating GtoPdb ligand and target identifiers in their text that became clickable out-links in the published HTML and PDF versions. This has the additional advantage of setting up a virtuous circle of reciprocal connectivity with PubMed where DARCP curated by GtoPdb has been submitted to PubChem. This is outlined in Figure 5.

Anomalies in the system

The wider informatics ecosystem exhibits a range of quirks related to DARCP and DC capture. These can complicate connectivity, confound standardisation and make navigation difficult, especially where they are non-obvious. The technical decisions that have caused such anomalies were generally been made to accommodate different submitter requirements (i.e., no one is trying to make the system more complicated, it just seems that way). The following is a selection:

1.) PubChem presents users with the complexity of parallel systems of D-C connectivity [28]. For medical subject headings (MeSH) the publication links are biased towards common name matches in many papers (e.g., the MeSH category for chemicals and drugs links 127000 PubChem CIDs to over 14 million PMIDs). Somewhat surprisingly, the largest D-C source by far is the IBM automated patent extraction system. This has operated on not only patents but also PubMed (plus MeSH terms in those abstracts) as well as full text from PMC articles. By 2016 this was responsible for 56% of all PMID-CID mappings (although IBM made what may have been their final submission in 2017). PubChem has a third substantial category of D-C connectivity from the publishers Springer, Thieme and most recently Wiley. These three sources have added document links for 660, 740 and 118 thousand CIDs respectively (with an overlap of only 74000). However, those having DOI-only document links are not connected into Entrez. They are made accessible via cross-references in the CIDs for Thieme and Wiley but only via SIDs for Springer. Since these publishers have provided a proof of concept for these D-C efforts it is to be hoped that they will be followed by equivalent undertakings from other chemistry publishers, for example ACS, Elsevier and ChemRxiv.

2.) The three DARCP sources with conceptually equivalent curated links (Table 2) are indexed within the NCBI systems in different ways. GtoPdb links into Entrez via reciprocal connectivity between 11000 PMIDs and 9800 SIDs. It also has just under 2000 BioAssay links by target class. ChEMBL has just over 1 million BioAssay entries largely as per publication indexing and also has its own target hierarchy in PubChem. While the SIDs have no PMID links 34000 of the ChEMBL-extracted PMIDs are indexed in Entrez but only via BioAssay. BindingDB has neither extensive BioAssay links nor Entrez SID connectivity. However, the 28000 curated PMIDs in this case are connected into the NCBI system as a LinkOut source.

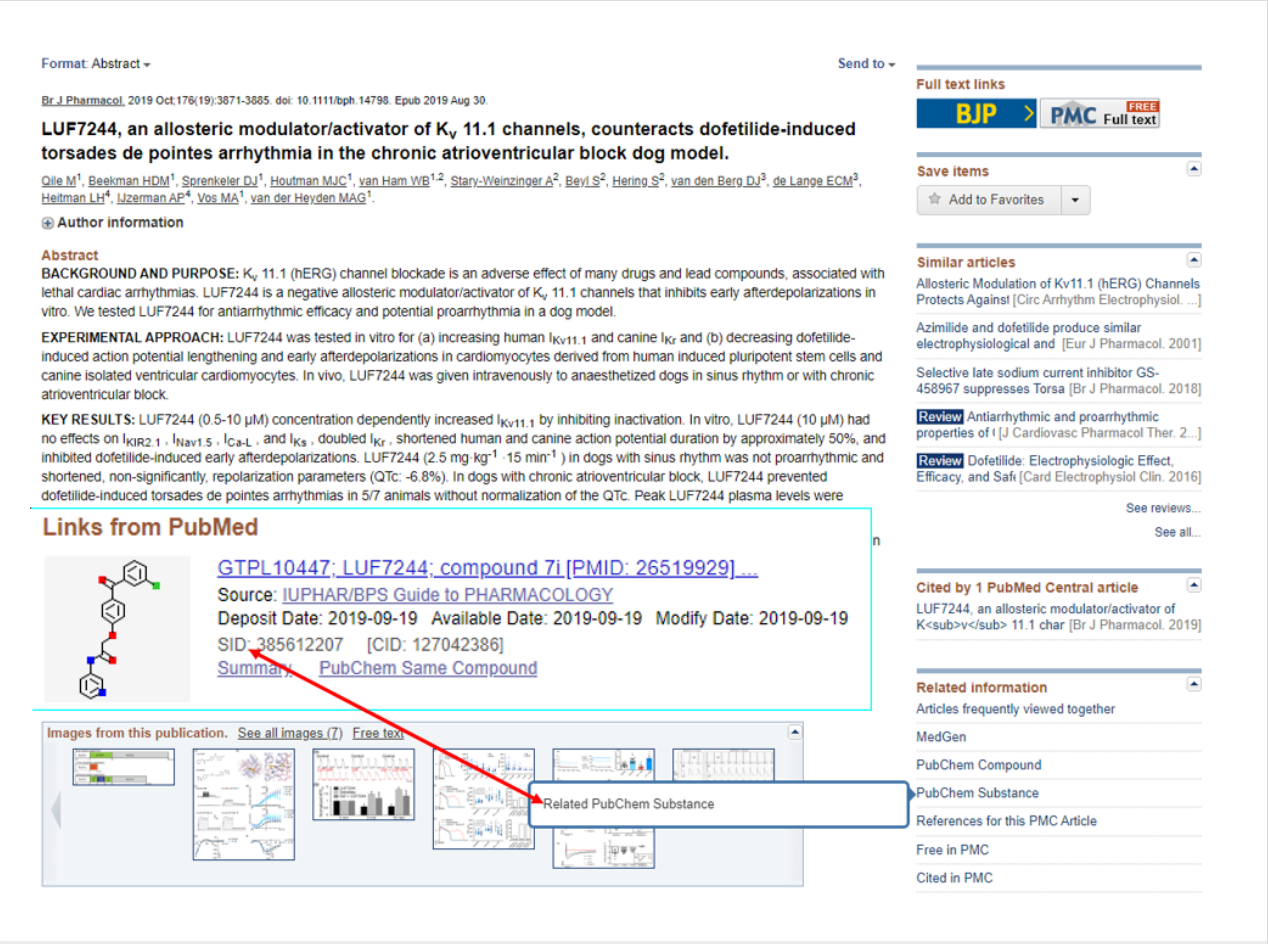


Figure 5: Connectivity for PMID 31339551. The lead structures from the paper, LUF7224 was curated by GtoPdb as ligand ID 10447 for which the DARCP is thus annotated (with P being Kv11.1 in this case). The submission of PubChem SID 385612207 by GtoPdb is indexed in Entrez with a live PubChem link from the PubMed right hand facet. The arrow indicates the reciprocity in that users can navigate "in" to the system via GtoPdb or "out" either via the CID or the PMID.

3.) As a commendable initiative the *Journal of Medicinal Chemistry* requires authors to provide SMILES [29] and in some cases, they may also add activity values, as supplementary data. These are made available as comma-separated (.csv) files. However, while this was envisaged to facilitate automated extraction, no one actually does this (or at least has not openly surfaced the results). These files thus useful contain C-R but A and P remain in the paper (although DARCP from this journal is extensively curated by GtoPdb, ChEMBL and BindingDB).

4.) Despite the pioneering efforts of *Nature Chemical Biology* there are caveats associated with their D-C mappings. The first is that in their SID records they index DOIs in the Depositor Comments but not PMIDs. This means there is no connection into the Entrez system (although this may have been an expedient choice to avoid the lag time associated with post-publication PMID assignments). As another quirk, there are 2,447 structures submitted by the journal that do not have an exact match to those extracted by the Springer automated pipeline for the same documents. It would be advantageous (including increasing traffic to the journal) if they could extend the author data submissions to enable full DARCP representation in PubChem BioAssay for suitable data sets.

5.) The transfer of data from the literature into on-line open resources (by an individual or a curation enterprise) could conceivably come up against copyright issues. The complexities of permissions to mine scholarly content were reviewed in 2016 [30]. It was reported therein that small amounts of data (e.g., presumably encompassing DARCP) would not generally be considered "creative expression" and thus should be exempt from copyright. We can also note that, after the fact, many public databases have been adding both curated and automatically mined "data facts" for well over a decade.

6.) The extraction of DARCP and D-C connectivity from patents presents a number of anomalies specific to this docu-

ment type. The first is that, compared to journal articles in which a proportions of the same data are later republished, in terms of compound structures in-common the appropriate PubChem query records a CID intersect of 29% between ChEMBL and all the major automated patent chemistry extraction sources adding up to 29.7 million), the document corpus has no paywalls. This means it is not only free-to-mine but also the HTML (before hamburgerisation) available from the USPTO greatly facilitates automated extraction. The second is that, in contrast to the commercial DARCP curation efforts by Excelra and Elsevier (implying the perceived high value), public extraction of patent ARCP is limited almost exclusively to BindingDB. However, as a consequence of being free-to-mine a number of operations have carried out public large-scale automated D-C extractions. These include, SureChEMBL [31], IBM, World Intellectual Property Organization (WIPO) and most recently Google Patents. However, the problem arises in PubChem and other sources) of what could be called "swamping" from the continual re-indexing (i.e., D-C linking) of common chemistry and structures without any ARCP data. As an example, in PubChem CID2244 for aspirin there are 143,180 connections to patent documents.

Conclusion

Comparing the historical connectivity between bioinformatics and cheminformatics points towards the root of the problems we currently face. Over more than three decades the links between sequence data and the literature have become a blazing success, first for molecular biology followed by genomics. This was driven mainly by the combination of journal mandates for author inclusion of sequence accession numbers and somewhat later, pointers to genomic and expression data sets. This has needed herculean technical integration efforts not only from the NCBI Entrez system and the equivalent EBI resources but also global coordination by the International Nucleotide Sequence Database Collaboration (INSDC) [32]. While compliance is not 100%, extensive literature and data set connections are now captured by both PubMed, PMC and EPMC.

The paradox is that no open equivalent ever emerged in the chemistry domain, in part due to the dominance of SciFinder. Thus, despite PubChem CIDs appearing in 2004 and the InChI identifier being implemented in 2013, publishers (with a few exceptions) have neither mandated nor encouraged the inclusion of open, machine readable chemical representations and/or open chemical database identifiers in their journals (external to the PDFs). The consequent shortfall for chemistry capture in general and DARCP in particular, shows no signs of diminishing. A rough estimate, derived from comparing commercial numbers from Figure 3 with public ones in Table 1, would be a chemical structure ratio of roughly 2 million to 7 million (i.e., a public shortfall in the order of ≈5 million although the major part of the later comes from patents).

An important corollary is that despite progress in automated chemical and biomedical entity recognition from text (e.g., via Natural Language Processing, NLP) [33] the fully automated extraction of explicit DARCLP relationships from documents has not yet been achieved (although AI efforts are doubtless pushing towards this). This stands in contrast to the ability of biocurators to discern such relationships from a paper in minutes (but needing extra minutes for a patent) [34]. The expansion of automated D-C capture in PubChem (e.g., by Springer, Thieme and Wiley) as well as automated chemical look-up in PMC and EPMC are certainly welcome developments. Notwithstanding, the associated ABNE problems severely limit knowledge distillation from D-C connections alone.

So, where do we go from here? The good news is that GtoPdb, ChEMBL and BindingDB should continue their expert capture role. The bad news is that it looks like, even by 2020, no journal will yet have instigated a formal process to extract DARCP and pipe it directly into open databases. One can only surmise that there is neither sufficient publisher "pull" nor author-incentivised "push" to make this happen. An alternative solution would be for authors to independently facilitate the transfer of their own annotated DARCP data into, for example, PubChem BioAssay. While the key connectivity to a PMID (via Entrez) could be added later, the necessary database submissions (possibly directly from an ELN) could, in principle, be de-coupled from the publishing process and thereby bypass PDF-entombment. Here again, we come up against the impasse of which stakeholders would value this high enough to instigate it.

Notwithstanding, we can take a more optimistic look at recent developments in the wider context of knowledge sharing including semantics and linked data that have the potential to improve the situation for DARCP capture. These include open data [35], open access (OA) [36], FAIR (findable, accessible, interoperable and reusable) [37,38], resource description framework (RDF) [39] and [40] WikiData [40]. While there is certainly momentum behind these trends, the persistence of publisher paywalls still remains a serious obstacle (e.g., of the 62000 papers curated by ChEMBL in EPMC only 85000 are full-text and only 600 OA). Strong community adoption (including from publishers) is also being seen for FAIR, which, in principle, should encompass accessibility to D, A, R, C and P (even if not their explicit connectivity) Planning is underway for the capture of FAIR data in various repositories (e.g., Figshare) but quite how this would practically expedite the flow

of connected DARCP into major databases (including core resources of the ELIXIR - distributed infrastructure for biological data [41]) is not yet clear. Another new development in the list, WikiData [40], is a community-maintained knowledge base that builds on the principles of FAIR. Here again, we will have to see how the practicalities of crowdsourcing DARCP curation and feeds into open databases can be accomplished.

Supporting Information

Supporting Information File 1

Technical details.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-16-54-S1.pdf]

Acknowledgments

Paul A. Thiessen from PubChem is acknowledged for details on document-to-chemistry connectivity. The referees' comments were much appreciated for improving the final version.

ORCID® iDs

Christopher Southan - https://orcid.org/0000-0001-9580-0446

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint doi:10.26434/chemrxiv.10295546

References

1. Sayers, E. W.; Beck, J.; Brister, J. R.; Bolton, E. E.; Canese, K.; Comeau, D. C.; Funk, K.; Ketter, A.; Kim, S.; Kimchi, A.; Kitts, P. A.; Kuznetsov, A.; Lathrop, S.; Lu, Z.; McGarvey, K.; Madden, T. L.; Murphy, T. D.; O'Leary, N.; Phan, L.; Schneider, V. A.; Thibaud-Nissen, F.; Trawick, B. W.; Pruitt, K. D.; Ostell, J. *Nucleic Acids Res.* **2020,** *48,* D9–D16. doi:10.1093/nar/gkz899
2. Levchenko, M.; Gou, Y.; Graef, F.; Hamelers, A.; Huang, Z.; Ide-Smith, M.; Iyer, A.; Kilian, O.; Katuri, J.; Kim, J.-H.; Marinos, N.; Nambiar, R.; Parkin, M.; Pi, X.; Rogers, F.; Talo, F.; Vartak, V.; Venkatesan, A.; McEntyre, J. *Nucleic Acids Res.* **2018,** *46,* D1254–D1260. doi:10.1093/nar/gkx1005
3. Southan, C. *Drug Discovery Today: Technol.* **2015,** *14,* 3–9. doi:10.1016/j.ddtec.2014.12.001
4. Harding, R. J. *PLoS Biol.* **2019,** *17,* e3000120. doi:10.1371/journal.pbio.3000120
5. Southan, C. *J. Cheminf.* **2013,** *5,* 10. doi:10.1186/1758-2946-5-10
6. Filippov, I. V.; Nicklaus, M. C. *J. Chem. Inf. Model.* **2009,** *49,* 740–743. doi:10.1021/ci800067r
7. Lowe, D. M.; Corbett, P. T.; Murray-Rust, P.; Glen, R. C. *J. Chem. Inf. Model.* **2011,** *51,* 739–753. doi:10.1021/ci100384d
8. Phillips, M. A.; Burrows, J. N.; Manyando, C.; van Huijsduijnen, R. H.; Van Voorhis, W. C.; Wells, T. N. C. *Nat. Rev. Dis. Primers* **2017,** *3,* 17050. doi:10.1038/nrdp.2017.50
9. Southan, C. Name-to-struc-to-target for the latest antimalarial review. 2017; https://cdsouthan.blogspot.com/2017/08/name-to-struc-resolution-for-yet.html.
10. Armstrong, J. F.; Faccenda, E.; Harding, S. D.; Pawson, A. J.; Southan, C.; Sharman, J. L.; Campo, B.; Cavanagh, D. R.; Alexander, S. P. H.; Davenport, A. P.; Spedding, M.; Davies, J. A.; NC-IUPHAR. *Nucleic Acids Res.* **2019,** *48,* D1006–D1021. doi:10.1093/nar/gkz951
11. Gabrielson, S. W. *J. Med. Libr. Assoc.* **2018,** *106,* 588–590. doi:10.5195/jmla.2018.515
12. Chambers, J.; Davies, M.; Gaulton, A.; Papadatos, G.; Hersey, A.; Overington, J. P. *J. Cheminf.* **2014,** *6,* 43–50. doi:10.1186/s13321-014-0043-5
13. Lipkus, A. H.; Watkins, S. P.; Gengras, K.; McBride, M. J.; Wills, T. J. *J. Org. Chem.* **2019,** *84,* 13948–13956. doi:10.1021/acs.joc.9b02111
14. Mutton, T.; Ridley, D. D. *J. Chem. Educ.* **2019,** *96,* 2167–2179. doi:10.1021/acs.jchemed.9b00268
15. Exelra. GOSTAR: Applications. 2019; https://www.gostardb.com/gostar/newui/applications.jsp (accessed Nov 13, 2019).
16. Elsevier. Improve content discoverability across the entire drug discovery workflow - Reaxys Medicinal Chemistry. 2019; https://www.elsevier.com/solutions/reaxys/who-we-serve/pharma-rd/reaxys-medicinal-chemistry (accessed Nov 13, 2019).
17. Southan, C.; Varkonyi, P.; Boppana, K.; Jagarlapudi, S. A. R. P.; Muresan, S. *PLoS One* **2013,** *8,* e77142. doi:10.1371/journal.pone.0077142
18. Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. *Nucleic Acids Res.* **2015,** *44,* D1045–D1053. doi:10.1093/nar/gkv1072
19. Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. *Nucleic Acids Res.* **2019,** *47,* D930–D940. doi:10.1093/nar/gky1075
20. Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. *Nucleic Acids Res.* **2016,** *44,* D1202–D1213. doi:10.1093/nar/gkv951
21. PubChem Statistics. 2020; https://pubchemdocs.ncbi.nlm.nih.gov/statistics.
22. ChEMBL Database Statistics. 2020; https://www.ebi.ac.uk/chembl/.
23. The Binding Database. 2020; https://www.bindingdb.org/bind/index.jsp.
24. Patents In BindingDB. 2020; https://www.bindingdb.org/bind/ByPatent.jsp.
25. Database Content. 2020; https://www.guidetopharmacology.org/about.jsp.
26. The UniProt Consortium. *Nucleic Acids Res.* **2018,** *47,* D506–D515. doi:10.1093/nar/gky1049
27. McGrath, J. C.; Pawson, A. J.; Sharman, J. L.; Alexander, S. P. H. *Br. J. Pharmacol.* **2015,** *172,* 2929–2932. doi:10.1111/bph.13112
28. Kim, S.; Thiessen, P. A.; Cheng, T.; Yu, B.; Shoemaker, B. A.; Wang, J.; Bolton, E. E.; Wang, Y.; Bryant, S. H. *J. Cheminf.* **2016,** *8,* 32. doi:10.1186/s13321-016-0142-6
29. Gilson, M. K.; Georg, G.; Wang, S. *J. Med. Chem.* **2014,** *57,* 1137. doi:10.1021/jm5002056

30. Molloy, J.; Haeussler, M.; Murray-Rust, P.; Oppenheim, C. Responsible Content Mining. In *Working with Text: Tools, Techniques and Approaches for Text Mining;* Tonkin, E. L.; Tourte, G. J. L., Eds.; Chandos Information Professional Series; 2016; pp 89–109. doi:10.1016/b978-1-84334-749-1.00004-4
31. Papadatos, G.; Davies, M.; Dedman, N.; Chambers, J.; Gaulton, A.; Siddle, J.; Koks, R.; Irvine, S. A.; Pettersson, J.; Goncharoff, N.; Hersey, A.; Overington, J. P. *Nucleic Acids Res.* **2015,** *44,* D1220–D1228. doi:10.1093/nar/gkv1253
32. Karsch-Mizrachi, I.; Takagi, T.; Cochrane, G.; on behalf of the International Nucleotide Sequence Database Collaboration. *Nucleic Acids Res.* **2018,** *46,* 48–51. doi:10.1093/nar/gkx1097
33. Krallinger, M.; Rabal, O.; Lourenço, A.; Oyarzabal, J.; Valencia, A. *Chem. Rev.* **2017,** *117,* 7673–7761. doi:10.1021/acs.chemrev.6b00851
34. International Society for Biocuration. *PLoS Biol.* **2018,** *16,* e2002846. doi:10.1371/journal.pbio.2002846
35. Farrell, E. The State of Open Data Report 2019. 2019; https://digitalscience.figshare.com/articles/The_State_of_Open_Data_Report_2019/9980783.
36. Tennant, J. P.; Crane, H.; Crick, T.; Davila, J.; Enkhbayar, A.; Havemann, J.; Kramer, B.; Martin, R.; Masuzzo, P.; Nobes, A.; Rice, C.; Rivera-López, B.; Ross-Hellauer, T.; Sattler, S.; Thacker, P. D.; Vanholsbeeck, M. *Publications* **2019,** *7,* 34. doi:10.3390/publications7020034
37. Sansone, S.-A.; McQuilton, P.; Rocca-Serra, P.; Gonzalez-Beltran, A.; Izzo, M.; Lister, A.; Thurston, M.; Batista, D.; Granell, R.; Adekale, M.; Dauga, D.; Ganley, E.; Hodson, S.; Lawrence, R.; Khodiyar, V.; Tenenbaum, J.; Axton, J. M.; Ball, M.; Besson, S.; Bloom, T.; Bonazzi, V.; Jimenez, R.; Carr, D.; Chan, W. M.; Chung, C.; Clement-Stoneham, G.; Cousijn, H.; Dayalan, S.; Dumontier, M.; Dzale Yeumo, E.; Edmunds, S.; Everitt, N.; Fripp, D.; Goble, C.; Golebiewski, M.; Hall, N.; Hanisch, R.; Hucka, M.; Huerta, M.; Kenall, A.; Kiley, R.; Klenk, J.; Koureas, D.; Larkin, J.; Lemberger, T.; Lynch, N.; Schriml, L.; Ma'ayan, A.; MacCallum, C.; Mons, B.; Moore, J.; Muller, W.; Murray, H.; Nobusada, T.; Noesgaard, D.; Paxton-Boyd, J.; Orchard, S.; Rustici, G.; Schurer, S.; Sharples, K.; Soares, N.; Stanford, J.; Subirats-Coll, I.; Swedlow, J.; Tong, W.; Wilkinson, M.; Wise, J.; Yilmaz, P. *bioRxiv* **2018,** 245183. doi:10.1101/245183
38. Scalfani, V. F.; MacEwen, L. Workshop: FAIR Publishing Guidelines for Spectral Data and Chemical Structures. 2019; https://osf.io/psq7k/.
39. Chen, B.; Dong, X.; Jiao, D.; Wang, H.; Zhu, Q.; Ding, Y.; Wild, D. J. *BMC Bioinf.* **2010,** *11,* 255. doi:10.1186/1471-2105-11-255
40. Waagmeester, A.; Stupp, G.; Burgstaller-Muehlbacher, S.; Good, B. M.; Griffith, M.; Griffith, O.; Hanspers, K.; Hermjakob, H.; Hudson, T. S.; Hybiske, K.; Keating, S. M.; Manske, M.; Mayers, M.; Mietchen, D.; Mitraka, E.; Pico, A. R.; Putman, T.; Riutta, A.; Queralt-Rosinach, N.; Schriml, L. M.; Shafee, T.; Slenter, D.; Stephan, R.; Thornton, K.; Tsueng, G.; Tu, R.; Ul-Hasan, S.; Willighagen, E.; Wu, C.; Su, A. I. *bioRxiv* **2019,** 799684. doi:10.1101/799684
41. Drysdale, R.; Cook, C. E.; Petryszak, R.; Baillie-Gerritsen, V.; Barlow, M.; Gasteiger, E.; Gruhl, F.; Haas, J.; Lanfear, J.; Lopez, R.; Redaschi, N.; Stockinger, H.; Teixeira, D.; Venkatesan, A.; Bateman, A.; Cochrane, G.; Finn, R.; Keane, T.; Leach, A.; McEntyre, J.; Orchard, S.; Parkinson, H.; Porras, P.; Sarkans, U.; Spalding, D.; Velankar, S.; Vizcaíno, J. A.; Yates, A.; Bridge, A.; Glöckner, F. O.; Hanauer, M.; Rath, A.; Rodwell, C.; Licata, L.; Oksvold, P.; Uhlén, M.; von Feilitzen, K.; Orengo, C.; Persson, B.; Rambla, J.; Schomburg, D.; von Mering, C.; Blomberg, N.; Durinx, C.; McEntyre, J. *Bioinformatics* **2020,** btz959. doi:10.1093/bioinformatics/btz959

License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (https://www.beilstein-journals.org/bjoc)

The definitive version of this article is the electronic one which can be found at: doi:10.3762/bjoc.16.54

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Anthelmintic drug discovery: target identification, screening methods and the role of open science

Frederick A. Partridge‡[1], Ruth Forman‡[2], Carole J. R. Bataille‡[3], Graham M. Wynne[3], Marina Nick[1], Angela J. Russell*[3,4], Kathryn J. Else*[2] and David B. Sattelle*[1]

Review

Open Access

Address:
[1]Centre for Respiratory Biology, UCL Respiratory, Division of Medicine, University College London, Gower Street, London, WC1E 6BT, United Kingdom, [2]The Lydia Becker Institute for Immunology and Inflammation, Faculty of Biology, Medicine and Health, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom, [3]Department of Chemistry, Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA United Kingdom and [4]Department of Pharmacology, University of Oxford, Mansfield Road, Oxford, OX1 3QT, United Kingdom

Email:
Angela J. Russell* - angela.russell@chem.ox.ac.uk; Kathryn J. Else* - kathryn.Else@manchester.ac.uk; David B. Sattelle* - d.sattelle@ucl.ac.uk

* Corresponding author ‡ Equal contributors

Keywords:
anthelmintic; antiparasitic; cestode; nematode; trematode

Beilstein J. Org. Chem. **2020,** *16,* 1203–1224.
doi:10.3762/bjoc.16.105

Received: 30 January 2020
Accepted: 12 May 2020
Published: 02 June 2020

This article is part of the thematic issue "Medicinal chemistry and drug discovery in an opening era".

Guest Editor: M. Todd

© 2020 Partridge et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Helminths, including cestodes, nematodes and trematodes, are a huge global health burden, infecting hundreds of millions of people. In many cases, existing drugs such as benzimidazoles, diethylcarbamazine, ivermectin and praziquantel are insufficiently efficacious, contraindicated in some populations, or at risk of the development of resistance, thereby impeding progress towards World Health Organization goals to control or eliminate these neglected tropical diseases. However, there has been limited recent progress in developing new drugs for these diseases due to lack of commercial attractiveness, leading to the introduction of novel, more efficient models for drug innovation that attempt to reduce the cost of research and development. Open science aims to achieve this by encouraging collaboration and the sharing of data and resources between organisations. In this review we discuss how open science has been applied to anthelmintic drug discovery. Open resources, including genomic information from many parasites, are enabling the identification of targets for new antiparasitic agents. Phenotypic screening remains important, and there has been much progress in open-source systems for compound screening with parasites, including motility assays but also high content assays with more detailed investigation of helminth physiology. Distributed open science compound screening programs, such as the Medicines for Malaria Venture Pathogen Box, have been successful at facilitating screening in diverse assays against many different parasite pathogens and models. Of the compounds identified so far in these screens, tolfenpyrad, a repurposed insecticide, shows significant promise and there has been much progress in creating more potent and selective derivatives. This work exemplifies how open science approaches can catalyse drug discovery against neglected diseases.

Review

The need for anthelmintic drug discovery

Anthelmintic drugs is the collective term for the group of drugs which treat infections of animals or humans infected with parasitic worms (helminths). Parasitic worms infect a wide range of species and as such present a major burden on not only human health, but also livestock production and crop production. There are two major phyla of helminths – nematodes (also known as roundworms) and platyhelminths (also known as flatworms). The nematodes include the soil-transmitted human helminths, e.g., *Ascaris lumbricoides*, and the vector transmitted tissue dwelling filarial worms, e.g., *Wuchereria bancrofti* while the platyhelminths include trematodes (also known as flukes) e.g. *Schistosoma mansoni* and cestodes (also known as tapeworms), e.g., *Taenia solium*.

Although the current review focusses on the unmet need in anthelmintic drug discovery to tackle the burden of human helminth infections, infection of livestock with parasitic worms has important animal welfare implications and can result in considerable economic losses to the livestock industry. In industrialised countries most livestock are routinely given anthelmintics to control or prevent infections and it is estimated the number of sheep, goats and cattle treated annually is hundreds of millions [1]. Treatment of horses, other equids, and companion animals is also a major use of anthelmintics.

Anthelmintic drug discovery has been a continued emphasis in the animal health industry, driven by the spread of resistance to the macrocyclic lactones [2]. In the past 25 years, three new classes of anthelmintic drugs have reached the market: derquantel, emodepside and monepantel. However, the continuing emergence of anthelmintic resistance combined with less predictable infection patterns due to changes in climate have resulted in a breakdown of control of these parasites [3].

Human helminth infections are neglected tropical diseases (NTDs). The most common infections, which we will highlight in this review, are caused by soil transmitted helminths (STHs), schistosomes and lymphatic filarial worms (Table 1). These

Table 1: Prevalence of and morbidity caused by major human helminth infections. DALYs are disability-adjusted life years.

disease	main etiologic helminth	number infected (million)	DALYs (million)	morbidity
soil-transmitted helminths				
ascariasis	*Ascaris lumbricoides*	819 [8,9]	1.3 [5]	infections (due in part to size and number of worms) and intestinal blockages (potentially requiring surgery), growth stunting and effects on cognition [10-12].
hookworm	*Necator americanus; Ancylostoma duodenale*	439 [8,9]	1.7 [5]	anaemia which can cause complications during pregnancy and post-birth; growth stunting and effects on cognition [13,14].
trichuriasis	*Trichuris trichiura*	465 [8,9]	0.3 [5]	inflammatory foci and haemorrhaging, growth stunting and effects on cognition [15-17].
filarial nematodes				
lymphatic filariasis	*Wuchereria bancroft; Brugia malayi*	120 [18]	1.2 [5]	lymphedema (elephantiasis), hydrocele, renal pathology manifesting as chyluria, and acute dermatolymphangioadenitis causing regular fevers.
onchocerciasis	*Onchocerca volvulus*	20 [19]	1.0 [5]	itching, skin inflammation and visual impairment or blindness
platyhelminth trematodes				
schistosomiasis	*Schistosoma haematobium, Schistosoma mansoni, Schistosoma japonicum*	Over 250 [20]	1.9 [5]	acute infection: myalgia, abdominal pain in the right upper quadrant, diarrhoea, fatigue, malaise, fever, chronic infection: reactions against eggs trapped in host tissues lead to inflammatory and obstructive symptoms; the tissues and organs affected depend on the *Schistosoma* spp. schistosomiasis is also associated with undernutrition, exercise intolerance, diarrhoea (sometimes bloody), chronic pain and anaemia [20].

infections are largely confined to rural, impoverished areas in tropical and subtropical regions of the developing world and co-infections with several different helminths are common [4]. In general, helminth infections are associated with morbidity (see Table 1) rather than mortality and high-intensity infections are associated with increased morbidity. Combined, they represent a massive global burden estimated at 6.4 million disability-adjusted life years (DALYs) [5] with life-long implications as they limit the educational prospects of children and reduce worker productivity [6,7]. Thus, they effectively trap whole countries in poverty.

Current anthelmintics – STHs

Helminth infections are predominantly located in unseen, rural areas of low income countries; thus despite their prevalence they have been coined the "forgotten diseases of forgotten people" [21]. It is perhaps not a surprise, therefore, that almost all the drugs available for human treatment were initially developed as veterinary medicines.

Effective and safe anthelmintic drugs can reduce prevalence, intensity and morbidity associated with STHs. Two main strategies for delivery exist: the management of diagnosed patients and preventative chemotherapy, which relies on mass scale administration of a single dose treatment to undiagnosed individuals; otherwise known as mass drug administration (MDA) [22-24]. MDA programs are the cornerstone control strategy for these infections and usually target treatment to pre-school or school-attending children.

Currently, the World Health Organization (WHO) recommends annual or biannual (where baseline prevalence is over 50%) intervention with the anthelmintics albendazole (ALB, **1**) or mebendazole (MEB, **2**), to treat STHs [25,26]. While achieving cure rates approaching 100% for Ascaris, these drugs, when used as single dose monotherapies, are less effective against hookworm [27,28] and have shockingly poor cure and egg reduction rates against *T. trichiura* [29]. (It is important to note that these drugs do have much better efficacy when administered as a course of treatment. However, given the practicalities and huge scale of mass drug administration programs, single dose efficacy is the benchmark for MDA.) As a result, single dose combination therapies, for example, with the tetrahydropyrimidines pyrantel pamoate (PYP, **3**) and oxantel pamoate (OXP, **4**) have been advocated over recent years with some success (Table 2, Figure 1) [30].

Unfortunately, drug resistance against benzimidazoles **1** and **2** and other anthelmintics have been detected in veterinary parasites, with mutations in the beta-tubulin gene [31,32]. To date there is limited evidence of resistance to benzimidazoles in human STHs [33], although benzimidazole resistance alleles have been found with increased frequency following anthelmintic treatment [34,35]. However, as the use of these anthelmintics has increased in the past decade, thanks to the donation of millions of doses of these drugs to enable mass drug administration, the selective drug pressure on the STHs has increased. This could trigger the emergence of drug resistance [36]. Additionally, the terrestrial stage of STHs can survive as eggs in the soil for several months (*A. lumbricoides* and *T. trichiura*) and larvae several weeks (hookworms), dependent on prevailing environmental conditions [37]. This environmental pool of the infectious lifecycle stage makes even those individuals successfully treated by MDA schemes at risk of reinfection [38]. It has therefore been proposed that, as well as targeting improvements in sanitation to tackle this issue, an environmentally-acting, egg-targeting agent could play a complementary role to help break transmission [39]. The WHO has set a global target to eliminate morbidity due to soil-transmitted helminth disease in children by 2020 through the regular treatment of at least 75% of the children in endemic areas (an estimated total number of 873 million) [40]. It looks unlikely that this target will be met, thus the development of new potent anthelmintic drugs that act via novel mechanisms of action are urgently needed.

Current anthelmintics – schistosomiasis

Praziquantel (PZQ, **5**), an *N*-acylated tetrahydroisoquinoline-piperazinone derivative has been the basis of schistosomiasis treatment for over 30 years [20]. PZQ is currently administered as a racemic mixture (Figure 1), despite the (*R*)-enantiomer being the biologically active form [41]. As will be discussed later in this review, there has been success in producing the active species in enantiomerically pure form. A clinical trial examining the bioavailability of orally dispersible tablets of levo-praziquantel has been completed [42], and a Phase III trial evaluating safety and efficacy is underway (NCT03845140).

Oral PZQ is safe and efficacious against adult worms of all *Schistosoma* spp., although only very recently have indications of its mechanism of action been advanced [43]. PZQ is the foundation of the global community-based schistosomiasis control programmes which use MDA, to reduce morbidity. A meta-analysis showed that the WHO-recommended dose of PZQ (40 mg/kg) achieved cure rates (CR) of between 76–95% for different *Schistosoma* species and 63.5% for mixed *S. haematobium* and *S. mansoni* infections and mean egg reduction rates (ERR) between 86–95% [44].

However, although PZQ is safe and efficacious against adult worms, it is only given once, and as the drug does not act against the migrating schistosomula stage of the parasite, if

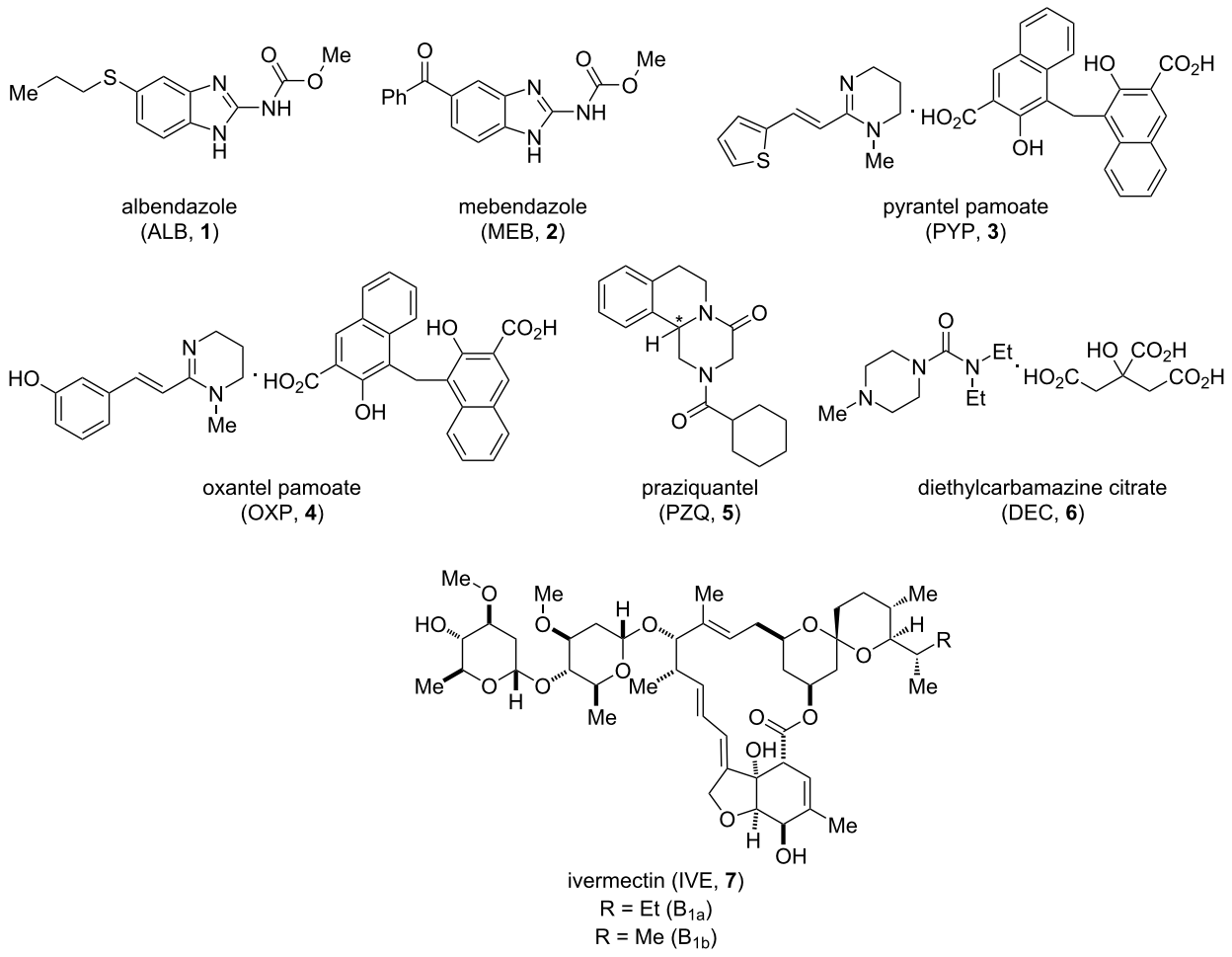


Figure 1: Structures of some current front-line anthelmintics discussed in this review. *Denotes the stereogenic centre: praziquantel is administered as a racemic mixture.

Table 2: Drug combinations show increased single-dose cure rates for soil-transmitted helminths. CR: cure rate – the proportion of infected individuals who become negative when tested after treatment with the indicated drug(s). ERR: egg reduction rate – the reduction in egg count after treatment expressed as a proportion of the egg count before treatment. Data from [30].

	MEB (**1**)		ALB (**2**)		ALB (**2**) + PYP (**3**) + OXP (**4**)		MEB (**1**) + PYP (**3**) + OXP (**4**)	
	CR	ERR	CR	ERR	CR	ERR	CR	ERR
Ascaris lumbricoides	96.8	99.5	96.5	99.7	90.4	98.3	100.0	100.0
Necator americanus and *Ancylostoma duodenale*	41.6	65.1	78.5	92.1	92.8	96.7	84.7	93.4
Trichuris trichiura	44.4	80.7	32.1	64.3	84.2	92.7	75.2	90.7

these larvae are present a new infection will arise. Moreover, PZQ does not prevent re-infection and therefore transmission can rapidly re-establish from just a few infected patients who can contaminate the aquatic environment.

Additionally, the emergence of drug resistance is a concern with some reports of infections that respond poorly to PZQ in areas where there has historically been heavy use of the drug [45]. However, as is the case for benzimidazoles, there is no clear confirmation that clinically relevant praziquantel resistance has developed [46]. Nevertheless, there is a widespread concern of the risk associated with relying on a single drug, particularly as MDA is being further scaled up, with hundreds of millions of doses of praziquantel being donated every year [20]. Another

unknown influence in the future of schistosomiasis treatment is the effect which climate change may have on aquatic environments (and the intermediate freshwater snail host *Biomphalaria glabrata)* and therefore on the distribution of water-borne diseases like schistosomiasis [47].

Similarly to the STHs, the WHO 2020 target for schistosomiasis is morbidity control by reducing the prevalence of heavy-intensity infections to less than 5% amongst school age children. However, whilst these guidelines look likely to be met in lower prevalence regions based on the current WHO MDA guidelines there is a low likelihood that these goals will be achievable in high-prevalence regions where the transmission potential is greater [48].

Current anthelmintics – lymphatic filariasis

Lymphatic filariasis (LF) treatment has been overseen through the WHO Global Program to Eliminate Lymphatic Filariasis (GPELF) which was launched in 2000. To achieve the WHO goal of the elimination of LF as a public health problem, the GPELF has a two-pronged approach involving not only preventative chemotherapy through MDA to treat the at-risk population, thereby interrupting transmission, but also management of disease morbidity.

By 2018, the GPELF had delivered over seven billion treatments to more than 910 million people [49]. Fourteen countries have eliminated LF as a public health problem, and a further ten countries have been able to stop MDA due to progress in controlling the infection.

MDA for LF involves combinations of three anthelmintics – albendazole (ALB, **1**), diethylcarbamazine citrate (DEC, **6**), ivermectin (IVE, **7**, administered as a mixture of B_{1a} and B_{1b}). A recent trial found that a single dose of the triple therapy (IVE + DEC + ALB) was able to clear *W. bancrofti* microfilariae from the blood for three years in almost all treated individuals, and this was superior to a single dose of two drug therapy (DEC + ALB) and non-inferior to three annual doses of the two drug therapy [50]. This trial also provided evidence that, in addition to clearance of microfilariae, both double and triple drug therapies have partial macrofilaricidal effects, as measured by reduction in circulating filarial antigen levels. A second trial also found that a single dose of the three drug (IVE + DEC + ALB) therapy had a greater ability to reduce *W. bancrofti* microfilariae for 24 months after treatment, compared to annual dosing of two drug (IVE + ALB) therapy [51], although microfilarial clearance was not sustained in the treatment population, likely due to reinfection. Importantly, this study also found greater inactivation of adult worm nests (clusters of active adult worms in the lymphatic tissue) in the IVE + DEC + ALB group, demonstrating the macrofilaricidal effect of this treatment combination.

These studies and others have led to WHO recommending triple therapy (IVE + DEC + ALB) for MDA in countries without endemic onchocerciasis or loiasis [52]. Implementation of triple therapy MDA has significant promise for elimination of LF in many countries.

Unfortunately, in countries where onchocerciasis is endemic, DEC (**5**) is contraindicated and achieving a cure is particularly problematic in areas where additionally *Loa loa* is co-endemic and the use of IVE (**7**) is therefore contraindicated. In these circumstances, which apply to some African countries, annual dual therapy (IVE + ALB) or biannual ALB monotherapy are used for MDA as appropriate. Thus, alternative anthelmintics are needed, ideally compounds which can achieve macrofilaricidal (i.e. curative) efficacy but which are safe in regions with onchocerciasis and loiasis.

To this end there has been significant investment into anti-*Wolbachia* treatments. These treatments target the bacterial symbiont *Wolbachia* which is essential for development, growth and survival of many filarial parasites. Targeting of *Wolbachia* with antibiotics has been shown to have curative efficacy against lymphatic filariasis [53] and importantly is safe to administer in *L. loa* co-endemic regions [54]. However, currently available effective antibiotics are unsuitable for public health MDA strategies due to contraindications and treatment duration and therefore novel compounds are required.

Current anthelmintics – onchocerciasis

Onchocerciasis (river blindness) is caused by infection by *Onchocerca volvulus.* There have been great efforts, beginning in the 1970s, to reduce the burden of this disease, first by control of the insect vector, and later by MDA of donated ivermectin. In 2018, over 150 million people in affected areas received ivermectin [55]. This macrocyclic lactone is an effective microfilaricide but does not kill the adult nematodes. Ivermectin must therefore be administered annually or twice-annually for many years to eliminate the parasite in the population. Progress has been impressive: onchocerciasis has been largely controlled as a public health problem in most of Africa [56], and four countries in the Americas – Columbia, Ecuador, Guatemala and Mexico – have achieved elimination of the parasite [57,58].

There has also been success in attempts to improve MDA for onchocerciasis by repurposing drugs from veterinary medicine. It has recently been shown in a Phase III trial that moxidectin is superior to ivermectin at reducing microfilarial density

12 months after a single dose [59]. This greater duration of action would be expected to reduce parasite transmission between annual rounds of MDA, accelerating progress towards elimination. Emodepside is another veterinary drug that is very promising for repurposing. It has shown activity in pre-clinical models of a variety of human helminth pathogens and is being pursued for treatment of onchocerciasis under an agreement between Bayer and the Drugs for Neglected Diseases Initiative [60]. Phase I safety trials have been completed (NCT03383614).

The management and control of the STHs, schistosomes and filarial parasites relies primarily on chemotherapy and education. Whilst vaccines are being developed for roundworms and whipworms, the development is still at the pre-clinical stage [61-65]. A hookworm vaccine is at a more advanced stage of development [66] and a schistosomiasis vaccine is in a clinical Phase III trial [67], however, no vaccines are currently in use in the field.

The current anthelmintic pipeline and the drug discovery landscape for parasitic helminth infections is sparse. This contrasts with the situation for malaria and kinetoplastid infections, where efforts, particularly by the Drug for Neglected Diseases Initiative and Medicines for Malaria Venture, in partnership with various pharmaceutical companies, are now paying off with an improved pipeline of drug development and the approval of tafenoquine (**11**) and fexinadole (**10**) [68,69]. With increasing concerns over the potential emergence of resistance to currently deployed anthelmintics, the possibility of climate change altering the distribution of these parasites, coupled with the inability of the currently available chemotherapies to impact parasite transmission and the poor efficacies of some of these drugs, e.g. against *Trichuriasis*, the need for new approaches to anthelmintic development is pressing. Additionally, the majority of anthelmintics are limited by their poor cross-phyla activity, e.g., praziquantel (PZQ, **5**) has efficacy against trematodes and cestodes but not nematodes. Only the benzimidazoles **1** and **2** show some broader effects but are much more active against nematodes than against cestodes or trematodes [70]. Ideally an anthelmintic with broad activity against different helminth infections would be desirable, although this may be too much to hope for given the evolutionary distance between the different target phyla.

As it stands, the WHO roadmap on NTDs, which set out a comprehensive plan for the control, elimination and eradication of NTDs, looks unlikely to deliver the desired outcomes by 2020. As the NTD 2030 roadmap is being rolled out there is an urgent need for novel anthelmintics to enable eradication of these diseases of poverty.

Application of open science to anthelmintic development

Commercial incentives for anthelmintic development

Despite the important need for new drugs and other control solutions for human helminth infection, these indications have been largely ignored by the pharmaceutical industry, presumably for commercial reasons. Indeed, no new chemical entities were approved between 2000 and 2011 [71,72]. Since then only moxidectin (**8**, 2018) and triclabendazole (**9**, 2019) have been approved. The major drugs used for control of human helminth infections have been in clinical use for many years: ivermectin (**7**, FDA approval in 1996), mebendazole (**2**, 1974), albendazole (**1**, 1996), praziquantel (**5**, 1982), diethylcarbamazine (**6**, 1950). The scarcity of new drugs reflects the limited economic incentive to spur commercial investment in neglected tropical diseases such as human helminth infection.

Efforts have been made to promote such investments. The FDA Tropical Disease Priority Review Voucher Program aims to create a commercial incentive to develop new drugs for otherwise neglected diseases [73,74]. Organisations that have an eligible drug successfully approved receive a transferrable voucher for a further priority review that has substantial value. In recent years two anthelmintic drugs have received the support of this program: moxidectin (**8**) showing superiority to ivermectin (**7**) for onchocerciasis [59], and triclabendazole (**9**) approved for fascioliasis, although both drugs were originally developed for veterinary indications and triclabendazole (**9**) was used for the treatment of fascioliasis for many years before FDA approval associated with the voucher program (Figure 2).

Public-private partnerships (PPPs), which typically bring together diverse organisations such as pharmaceutical companies, governments and charitable organisations are now proving successful at bringing drugs through to approval for neglected tropical diseases. For example, fexinidazole (**10**) was developed by the non-profit Drugs for Neglected Disease Initiative, in partnership with Sanofi, the Swiss Tropical and Public Health Institute, and other organisations. It has now been approved as the first all-oral treatment for all stages of human African trypanosomiasis [69,75]. A partnership between Medicines for Malaria Venture and GSK developed tafenoquine (**11**), which is effective as a single-dose treatment for the radical cure of *Plasmodium vivax* malaria [68]. While neither of these examples is for helminth infection, they show that these organisational models can be successful at bringing new molecules through to clinical practice. At an earlier stage of development, the anti-*Wolbachia* (A·WOL) consortium, a partnership including the Liverpool School of Tropical Medicine and AbbVie discovered ABBV-4083 (**12**), an antibiotic effective in preclinical models as a macrofilaricide by acting on the *Wolbachia*

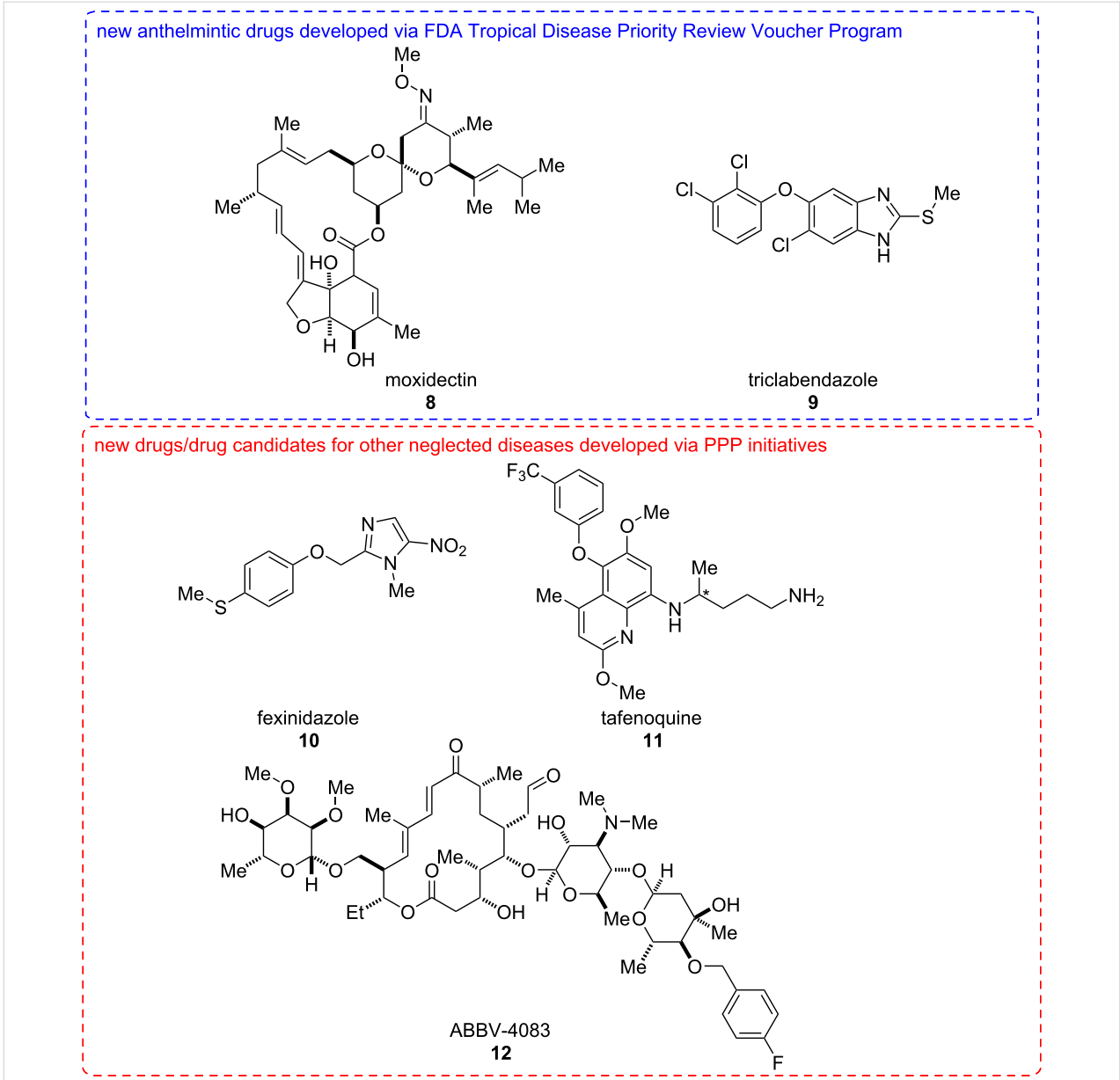


Figure 2: Structures of new anthelmintics drugs developed through repurposing, and new drugs or drug candidates for other neglected diseases developed through public–private-partnership (PPP) initiatives. *Denotes the stereogenic centre: tafenoquine (**11**) is administered as a racemic mixture.

bacterial endosymbiont [76] (Figure 2). WIPO Re:Search is another public–private partnership that facilitates work in neglected tropical diseases by bringing together intellectual property, expertise, facilities and funding from pharmaceutical companies, universities, and non-profit organisations. The number and breadth of projects, including many in the area of helminthiasis, that have been facilitated by this organisation since it was founded in 2011 is impressive [77].

Open science: an efficient model for drug innovation

An alternative way to promote anthelmintic drug discovery is to reduce the cost, by introducing research strategies that make drug innovation more efficient [72]. Open source drug discovery is a model that seeks to completely open up the research process [78]. This has several radical advantages that challenge traditional drug discovery. Secrecy and the hoarding of data in silos, such as individual research groups waiting for publication of their data, stifle our ability to access the best ideas. Openness can create communities that collaborate and attract new expertise when needed or serendipitously create new directions as different people from around the world and different fields bring fresh insights. Timely sharing of data speeds up research and avoids inadvertent repetition of effort. The wider drug discovery community is gradually adapting to these types

of challenges, with notable examples being ChEMBL (European Molecular Biology Laboratory) and PubChem (NIH). Both focus on characterising drug-like molecules and providing information to the public domain. The former is a manually curated database of bioactive molecules, which aims to bring together chemical, bioactivity and genomic data to aid the translation of genomic information into effective new drugs [79]. The latter is an open database for researchers to upload scientific data, including biological results, so that others may use it [80].

Perhaps the biggest example of open source drug discovery in its pure form is the Open Source Malaria project [81,82]. This is a platform for malaria-related research, with an emphasis on drug development. At the heart of this project are the completely open online electronic lab notebooks, that immediately share all work being done on the various strands of research of the project. Results are shared and publicly discussed, and priorities set on the Github issues page of the project. Importantly, anyone is free to jump in with suggestions, and indeed the Open Source Malaria Project has been successful at receiving high-quality contributions, from a wide range of sources. A highlight of this work has been the detailed exploration of an arylpyrrole antimalarial series [82]. We are not aware of a similar real-time, fully open source effort being applied to anthelmintic discovery, but this would be an exciting prospect for the field. However, researchers have been freely releasing open tools useful for drug discovery, openly describing their compound screening efforts, and participating in distributed open library screening projects such as the Medicines for Malaria Venture Pathogen Box. The point has recently been made by Tim Geary and colleagues that millions of compounds have been screened in industrial and academic labs on isolated helminths, but that the rate of drug discovery has been very low, so efforts must be made to enhance cooperation among the various groups pursuing this strategy [83]. They go on to suggest that sharing of both positive and negative screening data via online databases is a priority to focus attention on the most promising compounds and to reduce the redundancy of effort. Such a collaborative data-sharing structure must be a priority for the field.

C. elegans: a model organism for parasitology and an exemplar of an open community

C. elegans as a model nematode

Caenorhabditis elegans (*C. elegans*) is a non-parasitic nematode worm that is found worldwide and was selected by Sydney Brenner as a genetic model organism for biological research with strong potential to contribute to our understanding of developmental biology and neurobiology [84]. In 1998 it became the first complex eukaryote to have its genome sequenced [85]. *C. elegans* occurs as hermaphrodites and males and its capacity for hermaphroditic reproduction (selfing) facilitates the long-term maintenance of genetic strains. The capacity to freeze and store strains in glycerol adds to its utility. The transparency of the worm facilitates studies on the development, and *C. elegans* remains the only complex organism for which the entire cell lineage has been described [86]. For this pioneering work, Brenner, Horvitz and Sulston were awarded the 2002 Nobel Prize in Physiology/Medicine. The nervous system, which makes up 358 of the hermaphrodite's 959 somatic cells, is the only one for which a complete wiring diagram is known [87], facilitating studies on neural signalling and nervous and neuromuscular disorders [88] as well as research in understanding the anthelmintic drug action [89,90].

To grow and maintain *C. elegans* in the laboratory is relatively straightforward. Their small size (1 mm in length as adults) means ease of storage. Their rapid life cycle (approximately 3 days from egg to adult), and short lifespan (approximately 2–3 weeks) when fed on a diet of *E. coli* facilitates genetic studies. Forward and reverse genetics are pursued conveniently in *C. elegans*. A rich diversity of mutants is available via the Caenorhabditis Genetics Centre [91]. The discovery of RNA interference delivered via feeding worms double-stranded DNA [92] has opened the door to genome-scale gene knockdown in the search for new drug targets. These approaches can expedite the validation of drug targets and the identification of new candidate molecular targets.

So how can a free-living worm contribute to our understanding of parasitic nematodes and the development of anthelmintic drugs? A key advantage is the ease of culture of *C. elegans*. Large numbers can be generated rapidly and at low cost which enables high-throughput chemical and genetic screening studies. It is often difficult or impossible to undertake comparable studies on parasitic worms due to the challenges of maintaining parasitic worms outside their host, although rodent models are available for many classes of helminth [93].

Although *C. elegans* is clearly not a target organism, it can be deployed in the search for new anthelmintics for animal health and human health applications. Screens can be pursued for new chemical leads which may then be applied to other parasitic species. *C. elegans* chemistry-to-gene screens, facilitating deconvolution of the molecular target and mechanism of action, are also useful. For example, Burns and colleagues screened 67,012 compounds to identify those that kill *C. elegans* and followed this by rescreening hits in two parasitic nematode species and two vertebrate models (HEK293 cells and zebrafish). By this means, they identified 30 structurally distinct anthelmintic lead molecules [94]. They also determined the target (complex

II of the electron transport chain) of one lead compound, that showed nematode specificity and nanomolar potency. This work shows that *C. elegans* can be effective, cost-efficient, and has a role to play in the anthelmintic drug discovery process.

Another potential attribute in the context of investigating parasites is the ease with which the *C. elegans* genome can be manipulated, enabling the generation of transgenic *C. elegans* expressing anthelmintic drug targets from a parasitic worm [95,96]. These approaches are still in their infancy, but such genetic modifications can give rise to scorable phenotypes reflecting the properties of the parasite drug target which may in future lend themselves to high-throughput chemical and genetic (RNAi) screening approaches.

There are, however, limitations to using *C. elegans* as a research tool, notably its innate physical and enzymatic defences to xenobiotics, factors important for the survival in its natural environment. As a result, *C. elegans* is somewhat inaccessible to some chemicals, meaning that high concentrations of certain compounds may be required to observe changes in the phenotype [90,97].

Aroian and colleagues [98], in line with the work of Burns et al. [94], counsel caution on relying on data from *C. elegans* alone, which makes perfect sense as it is never the primary target organism. They screened a compound library against both adult and free-living larval stages (egg to L3i larval development assay, E2L) of the human hookworm parasite *Ancylostoma ceylanicum* and against *C. elegans*. They found that the *A. ceylanicum* E2L assay was more successful at identifying compounds active against *A. ceylanicum* adults than *C. elegans* assays (lower false negative rate). This works lead to the testing of four compounds with in vitro activity in an in vivo *A. ceylanicum* hamster infection model – sulconazole (**13**), econazole (**14**), pararosaniline (**15**) and cetylpyridinium chloride (**16**) (Figure 3). Of these pararosaniline (**15**) showed a significant reduction in parasite egg production in this model, despite no activity in *C. elegans* assays.

Perhaps one of the best reasons for integrating *C. elegans* into the discovery process is that if interesting new molecules are identified which are active on both the target parasite and the genetic model organism, but the precise target remains unclear, then *C. elegans* genetics can offer a route to target identification that would be difficult by any other route.

The *C. elegans* community: historically an open science model

Research into *C. elegans* was pioneered by researchers including Victor Nigon, Ellsworth Dougherty and Jean-Louis Brun [99]. However, the use of *C. elegans* as a model organism in fields such as genetics, developmental biology and neuroscience was established by Sydney Brenner in the 1960s at the MRC Laboratory of Molecular Biology in Cambridge, UK [100]. An important feature of *C. elegans* research has been the

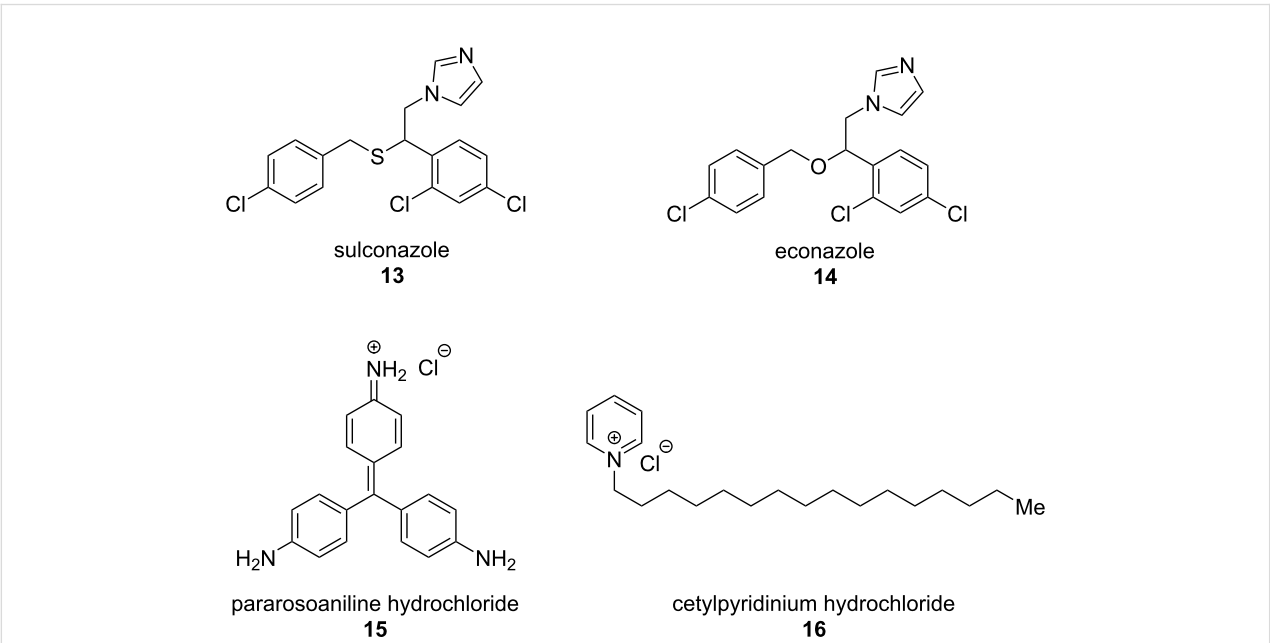


Figure 3: Compounds with anthelmintic activity identified by a combination of screening against *Ancylostoma ceylanicum, C. elegans, and T. muris* [98].

community of researchers, with a long tradition of openness and the sharing of ideas and reagents, which we today can recognise as an example of open science [101].

This openness is exemplified by a newsletter for *C. elegans* researchers, *The Worm Breeder's Gazette,* which combines advice and suggestions on methods with informal communication of new findings in advance of publication. For example, the Nobel prize-winning use of a green fluorescence protein as a marker was reported in the *Gazette* by Martin Chalfie and colleagues five months before the peer-reviewed publication [101-103].

A second example of the history of *C. elegans* open science is the genetic map and genome sequencing project, led by John Sulston and Bob Waterston, which made *C. elegans* in 1998 the first multicellular organism to have its genome sequenced. The project involved teams at the Genome Sequencing Center at the Washington University School of Medicine (St Louis, Missouri, USA) and the Sanger Centre (Hinxton, Cambridge, UK) sharing a belief that 'together, we can do more', rather than 'one against the other' [104]. An important feature was that genomic clones were made freely available, enabling researchers to investigate genes of interest [101]. Furthermore, genetic and genomic information was rapidly distributed, in the form of the ACeDB database, initially via gopher, an early internet service, and later via WormBase [105,106].

These historical examples illustrate the strength of open science in the *C. elegans* research tradition. In the following sections of this review we discuss how open science approaches continue to be important in the field of anthelmintic and antiparasitic drug discovery.

Open approaches to target identification

Genomic resources are important for target identification, particularly in the case of parasites, as the life stages found in the host are often difficult to obtain or culture, and few molecular tools are available. WormBase ParaSite [107] is an important central resource for helminth genomic data [108]. At the time of writing, this database contains information on 142 species of parasites and other helminths, including genomes, comparative genomics data and RNAseq studies, along with a number of tools facilitating access to this data including a genome browser, BioMart (a tool for exporting tables of selected information) and a REST API, an interface for programmatic access to the database.

But how can we get from genomic information to new drug targets? A recent comparative genomics study from a large international consortium of researchers has really helped with this question, by comparing the genomes of 81 species of parasitic and non-parasitic worms (both nematodes and platyhelminths) [109]. This work produced large open datasets, such as expanded gene families relevant to parasitism, and analyses of the metabolism in different parasites across the phyla, important for exploring metabolism in the search for new drugs. Furthermore, this study predicted promising anthelmintic targets and compounds likely to interact with these targets, thereby identifying drugs for potential repurposing as anthelmintics.

Once targets have been identified, it is desirable to obtain genetic/pharmacological proof-of-concept for target validation. Whilst RNA interference and CRISPR methodologies are now being applied to parasites themselves [110-112], inevitably, large-scale functional genomic resources are mainly found in *C. elegans.*

The *C. elegans* Gene Knockout Consortium has obtained putative knockout mutations in around 15,000 genes, and has now adopted CRISPR/Cas9 to extend the resource to every gene in the genome [113]. A complementary collection of knockout mutants from the National Bioresource Project in Japan is also available [114]. Both projects make the mutants openly available for low cost.

Another open source resource with immediate applicability to target identification and validation is the Open Worm Movement Database [115]. This is an open platform for analysing and sharing worm behavioural data, such as that obtained from worm tracking software. For example, the researcher can search for worm strains with a particular movement phenotype, such as low movement speed. These paralysed or poorly moving worm mutants may be a source of novel neuromuscular anthelmintic targets. The researcher can immediately view videos of the identified worm strains on YouTube to confirm their hypothesis.

Open tools for phenotypic screening

Recently, many laboratories throughout the world have recognised the need for the development of new anthelmintic compounds, so have initiated screening programmes against various pathogens and models. Despite our growing knowledge about potential targets for new anthelmintics, phenotypic screens involving assays of parasites or models in vitro remain important [83]. As a result, several methods and screening platforms have been developed to improve the screening speed and reliability.

In this section we discuss recent phenotyping methods and systems, and how they have been applied to anthelmintic

discovery. We concentrate on those where the software source code and/or hardware design is made clearly and openly available. Methods, applications, and location of source code/design are summarised in Table 3.

Open tools for phenotypic screening of motility and viability

Several groups have developed image acquisition and analysis systems for high-throughput phenotypic screening of parasites (Table 3), typically using the approach of thresholding difference images/movies to quantify motility, sometimes segmenting the image by recognising the organism of interest [116].

WormScan is a method that uses a flatbed scanner to capture sequential images, where the scanner high-intensity light usefully stimulates the worm movement [117]. This system has been utilised to screen a 26000 compound library in a *C. elegans* growth assay [119]. An updated version of this software (Automated WormScan) has recently been published [118]. The Lifespan Machine also uses a scanner to acquire images, and has the ability to monitor thousands of worms simultaneously and determine mortality time for individual worms on plates [120].

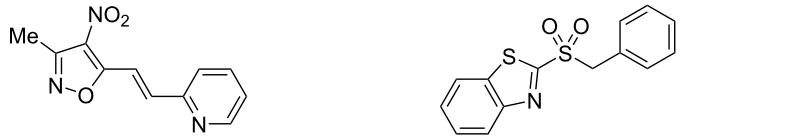
WormAssay is a combination of a video camera and an open source software package. It uses two algorithms (Lucas–Kanade optical flow estimation, and a pixel change method) to determine the motility of macroparasites in microtitre plates [122]. The Worminator builds on WormAssay for the use with microscopic parasites [123]. This system has been validated for determining the anthelmintic activity against a variety of nematodes and schistosomes. A screening using the Worminator identified auranofin as a promising candidate for repurposing as a treatment for of lymphatic filariasis and onchocerciasis [144]. This drug was originally approved as a treatment for rheumatoid arthritis but is currently in trials for amoebiasis and giardiasis (NCT02736968).

An ImageJ [145] macro that determines the motility of *Echinococcus multilocularis* protoscoleces within microtitre plates by pixel difference thresholding has been used to identify an anthelmintic hit compound MMV665807 [125]. Wiggle Index, another ImageJ macro for difference thresholding and motility quantification has been extensively used for library-scale screening of exsheathed *Haemonchus contortus* L3s [127-130]. INVAPP Paragon, an imaging setup and MATLAB analysis script that again uses difference thresholding for motility quantification has been used for library screening with *Trichuris muris* and *C. elegans* [39,135,136].

CellProfiler is a major open source package for quantitatively measuring phenotypes from imaging data, particularly from high-throughput screens [139]. Some groups have developed helminth analysis methods using CellProfiler. A virtual screening approach was used to identify inhibitors of *S. mansoni* thioredoxin glutathione reductase [140]. These virtual hits were then tested in a CellProfiler-based high content screen using *S. mansoni* schistosomula, which determines both the motility and a range of other phenotype scores, leading to the identification of 2 new small molecules with distinct chemical scaffolds **17** and **18** with activity against schistosomula and adult worms at low micromolar concentrations (Figure 4). Another CellProfiler toolbox enables the quantitation of *C. elegans* viability and fluorescence [141].

Table 3: Open tools for high-throughput phenotypic screening of motility and viability and their use for anthelmintic discovery.

tool	validated with	source code/description (license)
WormScan [117] automated WormScan [118]	*C. elegans* high-throughput screen [119]	paper supporting information [117,118]
Lifespan Machine [120]	*C. elegans* lifespan analysis	Github [121] (GPLv3)
WormAssay [122] Worminator [123]	*Brugia malayi* (adults and microfilariae), *Cooperia* spp. L3, *Dirofilaria immitis* microfilariae, *Schistosoma mansoni* [122,123]	Github [124] (GPLv2 or later)
Cestode motility ImageJ macro [125]	*Echinococcus multilocularis* protoscoleces	paper supporting information [125]
Wiggle index ImageJ macro [126-128]	multiple high-throughput library screens with *H. contortus* [129-134]	paper supporting information [126]
INVAPP paragon [135]	library screens with *T. muris* and *C. elegans* [39,135-137]	Github [138] (MIT license)
CellProfiler [139] CellProfiler schistosome pipeline [140] CellProfiler WormToolbox [141]	*S. mansoni* thioredoxin glutathione reductase inhibitor screening [140], *C. elegans* live/dead high-throughput screening [141]	Github [142] (BSD license) CellProfiler published pipelines website [143]



17

schistosomula EC_{50} 3.23 μM
adult EC_{50} 6.43 μM (male)
5.68 μM (female)

18

schistosomula EC_{50} 2.62 μM
adult EC_{50} 21.1 μM (male)
4.91 μM (female)

Figure 4: Inhibitors of *S. mansoni* thioredoxin glutathione reductase with anthelmintic activity [140].

Open tools for more detailed analysis of helminth physiology

Sophisticated software and hardware methods have been developed to phenotype more subtle aspects of worm biology than paralysis/motility/viability assays. These methods could form a fruitful basis for finding compounds that are anthelmintic in vivo but do not cause paralysis, perhaps involving aspects of the interaction with the host, interfering with the secretion of proteins, or other ways of damaging the worm [146]. They are also useful for understanding in more detail the mechanism of action of anthelmintic compounds, since it has been recognised that we do not fully understand how many anthelmintics work – for example the concentrations of macrocyclic lactones that paralyse worms in vitro are much greater than the concentration achieved by effective doses in vivo [147].

Several different open source systems have been established for tracking the worm movement of *C. elegans* [148-151]. These systems typically measure a number of parameters in addition to speed such as bending, reversals, and other aspects of behaviour. CeleST is a similar open source quantitative locomotion analysis system that measures aspects of nematode swim behaviour [152]. Such systems have, to our knowledge, not been utilised with helminth parasites, but such studies would be fruitful to dissect anthelmintic actions in detail.

Microfluidic systems have great potential to aid anthelmintic discovery by enabling finely detailed individual worm longitudinal microscopy. They have the potential to greatly reduce the amount of compound required for a screen hence enabling larger libraries to be economically screened. Encouragingly, some authors have made their microfluidic chip designs openly available, enabling utilisation and modification by other groups. Stress-Chip is a chip that allows the isolation of a hundred worms in single-worm arenas and monitoring as chemicals flow over the worms [153]. The CAD file for producing the microfluidic device has been made available on Figshare under the CC BY 4.0 license [154]. Another 10-chamber worm isolation microfluidic device has been reported, originally for the imaging of worms to quantify the sleep behaviour during development, and the CAD file is made available in the supporting information [155].

The cost of the equipment is of course often a concern, especially for groups working on neglected tropical diseases and/or in developing countries. Recently, an open hardware project has reported Incu-Stream, a long-term imaging system capable of automatically scanning wells across microplates and recording videos of worm movement for further analysis [156]. The authors provided a parts list with a total materials cost of $184. Schematics, CAD files and the associated software are provided on Github [157].

Open approaches to developing therapeutics

The Pathogen Box project

The Pathogen Box is a 400 compound collection that was made freely available by the Medicines for Malaria Venture (MMV), a not-for-profit product development partnership organisation [158]. The compounds have demonstrated activity against a variety of neglected tropical disease pathogens [159]. This is an open access science project, with the only condition that researchers agree to share their results. This model follows on from the successful MMV Malaria Box, where 55 groups compiled results from over 290 different assays in a diversity of screens related not only to malaria, but also to other neglected tropical diseases [160].

The Pathogen Box project is currently active, but several groups have already reported anthelmintic screens using this library [129,135,161-165]. We have compiled the results from these published screens in Figure 5. These results already highlight how the open approach enables the library to be tested against a variety of different organisms, enabling researchers to identify and prioritise compounds active against multiple pathogens. For example MMV690102, which was originally developed as an

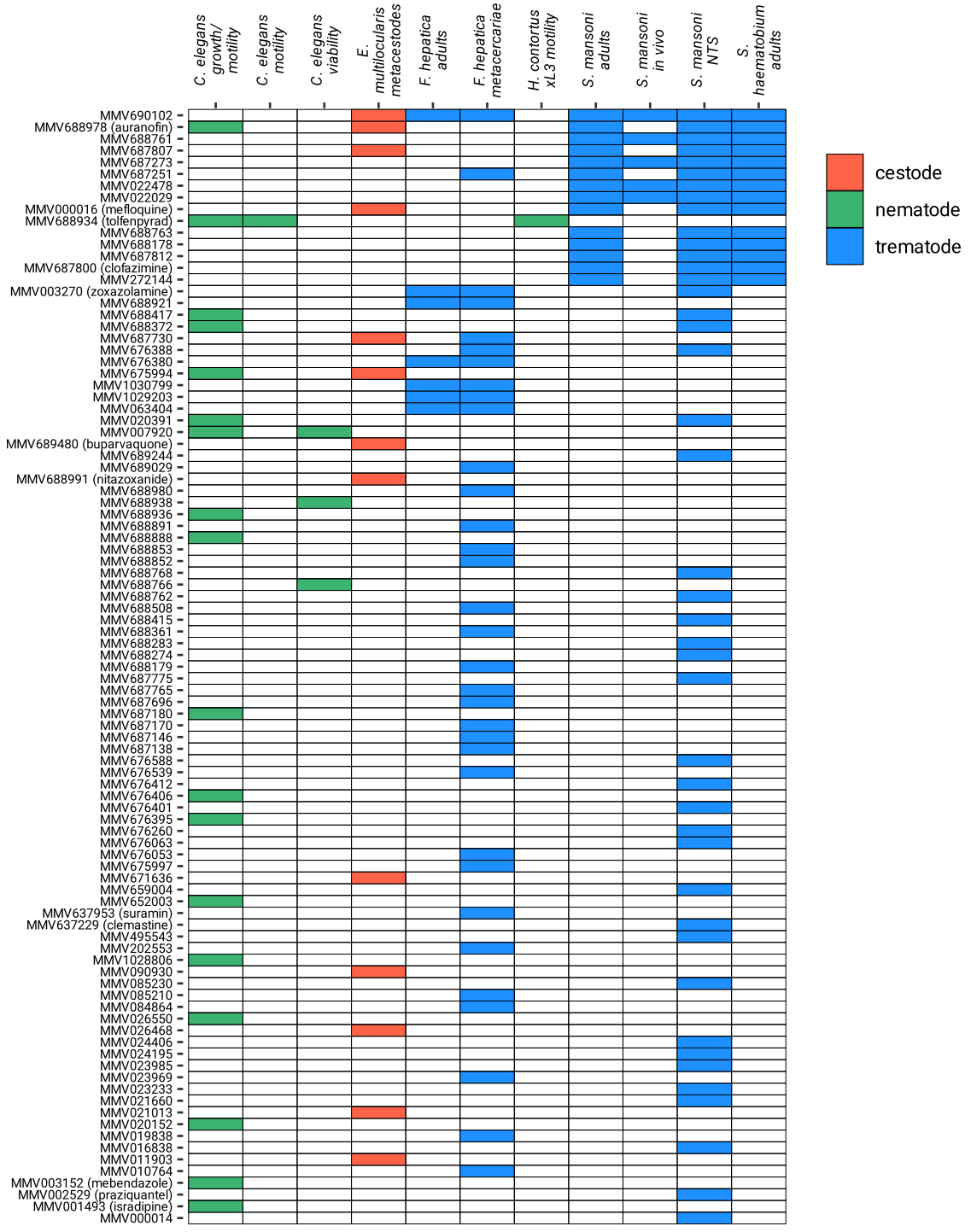


Figure 5: Active compounds from anthelmintic screens using the MMV Pathogen Box. NTS: newly transformed schistosomula, xL3: exsheathed L3. Assay outline and citations for data sources: *C. elegans* growth/motility screen – automated quantification using the INVAPP system [135]. *C. elegans* motility screen – automated quantification using the WMicrotracker ONE system [163]. *C. elegans* viability screen – automated quantification of viability by differential absorption of the dyes DB-1 and propidium iodide [164]. *E. multilocularis* metacestode screen – damage was assessed by quantifying the activity of phosphoglucose isomerase released into the culture media [162]. *F. hepatica* screens – viability of metacercariae was assessed using a scoring system taking into account membrane damage and fluke translucency, and viability of adults was assessed using a scoring system taking into account worm motility, colour and rigidity [165]. *H. contortus* screen – automated quantification of motility of exsheathed L3 worms [129], *S. haematobium* and *S. mansoni* screens – activity was assessed using a scoring system taking into account motility and morphological/tegumental changes [161].

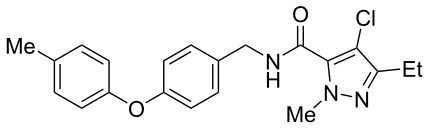
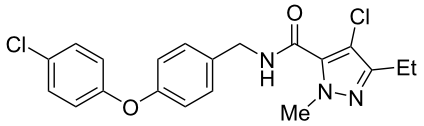
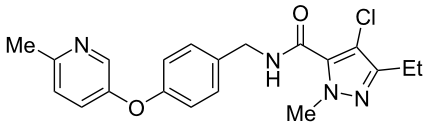
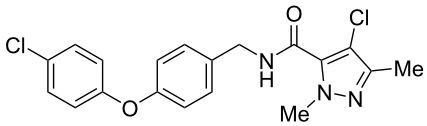
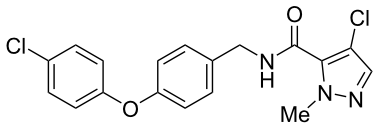
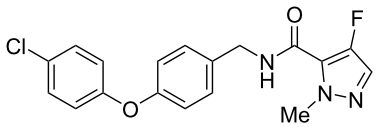
inhibitor of kinetoplastid dihydrofolate reductase [159], is active against three trematode species (*F. hepatica, S. haematobium, S. mansoni* [161]) and the cestode *E. multilocularis* [162].

Perhaps the most promising lead from the Pathogen Box so far is tolfenpyrad, a pyrazole-5-carboxamide insecticide, which was first identified as an anthelmintic with activity against exsheathed L3 and L4 parasitic life stages of *Haemonchus contortus,* a major parasite of ruminants [129]. Subsequent studies have demonstrated activity against the model nematode *C. elegans* [135,163]. Tolfenpyrad (**19**) was found to be highly potent, with an IC_{50} value between 0.02 and 3 μM in various *H. contortus* assays and 0.2 μM in a *C. elegans* assay [129,135]. A follow-up study identified two additional pyrazole-5-carboxamide compounds with activity against *H. contortus,* although not improving on the potency of tolfenpyrad [166]. Tolfenpyrad acts in arthropods as an inhibitor of mitochondrial complex I [167]. It will be interesting if a tolfenpyrad derivative can progress to trials as it would be a new mechanism of action for an anthelmintic, although some mitochondrial uncouplers, such as the veterinary medicine closantel, are active against *Fasciola hepatica* [168].

Recently, a medicinal chemistry effort was undertaken to determine the anthelmintic structure–activity relationships for tolfenpyrad (**19**) [169]. The main objective of this work was to reduce the lipophilicity of tolfenpyrad **19**, which was considered undesirable for an orally administered agent, as typical of anthelmintics, compared to a surface-applied pesticide. This was accomplished through systematic alteration of the pyrazole-5-carboxamide and phenoxybenzyloxy moieties within tolfenpyrad **19** (Table 4).

The systematic variation of the *p*-methylphenyl ring within **19** gave rise to a number of aromatic and heteroaromatic analogues with similar levels of potency to tolfenpyrad (Table 4). For instance, replacement of a methyl group with a chlorine

Table 4: Potent anthelmintic activity of tolfenpyrad (**19**) derivatives against *H. contortus.* The activity is shown for two in vitro assays: one for motility of xL3 (exsheathed L3 stage worms) and a second for development of xL3 into the L4 stage [169].

ID	structure	IC_{50} (μM) in xL3 motility assay	IC_{50} (μM) in L4 development assay
19 (tolfenpyrad)		2.9	0.03
20		2.03	0.0008
21		4.0	0.03
22		3.33	0.01
23		0.38	0.0007
24		0.7	0.0008

atom in **20** maintained similar levels of potency in both the xL3 motility assay and the L4 development assay. Similarly, replacement of *p*-methylphenyl group with a 2-methylpyrid-5-yl group in **21** largely maintained potency whilst lowering lipophilicity. Conversely, modifications to the pyrazole group gave rise to more dramatic changes in the potency. For instance, while changing the ethyl substituent within **20** to a methyl substituent in **22** gave a slight increase in potency, removal of the ethyl group in **23** showed a substantial improvement in the activity, with IC_{50} value for the xL3 motility improved to 0.38 μM and for the L4 development to 0.7 nM, with a similar activity for the corresponding fluoropyrazole derivative **24**. These latter two compounds **23** and **24** showed high selectivity for the parasite, with low or no cytotoxicity. The authors went on to demonstrate that **24** showed a broad activity against other nematode parasite models: *H. polygyrus, A. ceylanicum* and *T. muris*. A broadly-related 1-methyl-1*H*-pyrazole-5-carboxamide series has also been investigated in detail, with compounds identified that show substantially improved potency and selectivity compared to tolfenpyrad [131,170].

Praziquantel (**5**)

Schistosomiasis is a major tropical disease resulting from the infection by a trematode parasite, the blood fluke *Schistosoma mansoni* [171]. After malaria, it is the next most devastating parasitic disease with millions affected worldwide. No vaccine is available but the drug praziquantel (**5**) is an effective treatment. It is administered to children or whole communities often in mass drug administration (MDA) programmes [172]. An unfortunate drawback is that the drug is currently generated and administered as a racemic mixture. The pure active enantiomer would be preferable for several reasons, for example, the inactive enantiomer has been linked to unwanted side-effects and also contributes a very bitter taste.

With a view to finding a synthetic route to the active enantiomer, an open website was established, and several groups became involved, both academic and commercial laboratories. As a result, two different approaches to the problem emerged where hitherto there had been none (Figure 6). The hydrolysis to an intermediate amine **25** which was then resolved with a derivative of tartaric acid was a solution that emerged from this open source approach. Another solution was identified by a sponsored contract research team. This involved a different intermediate **26** which was then, in turn, resolved using tartaric acid itself. A detailed account of the successful resolution process has been published by Matthew Todd and colleagues [173]. This has not yet led to the pure enantiomer being widely available, but the setting up of an open science project was the stimulus to the solution of a challenging problem. A Phase III clinical trial testing safety and efficacy of L-praziquantel is currently recruiting (NCT03845140).

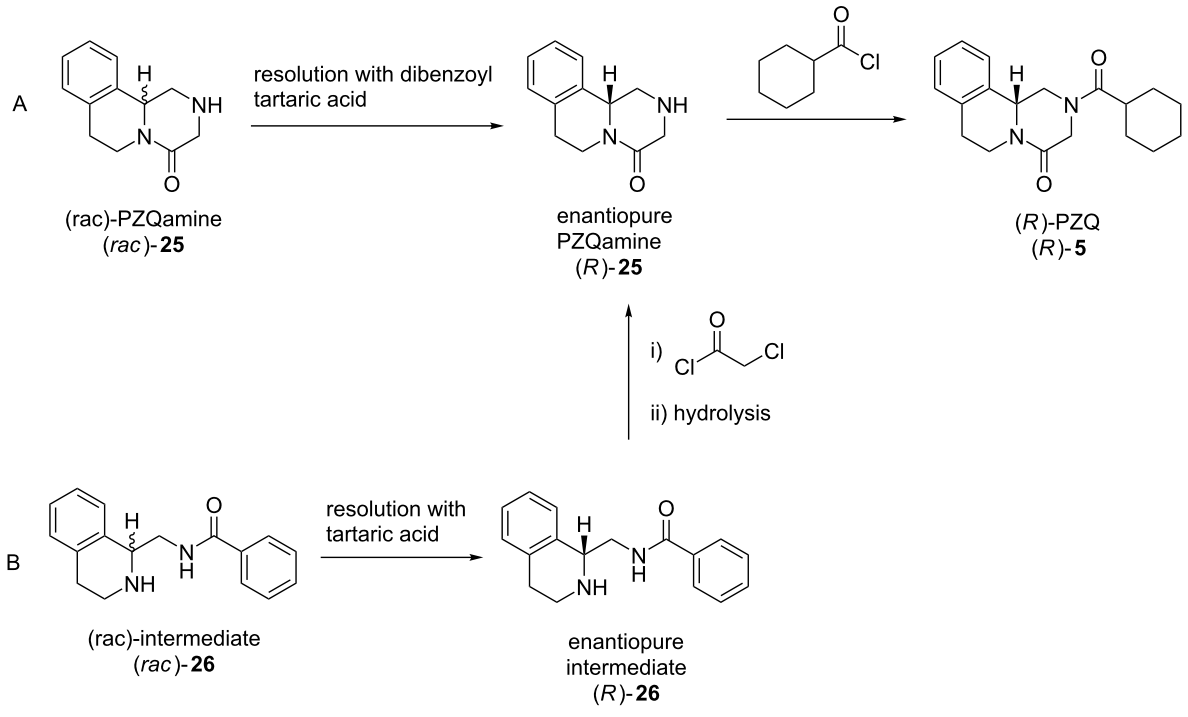


Figure 6: Two resolution approaches to enantiopure PZQ (*R*)-**5** discovered through A) open science and B) contract research [173].

Conclusion

The recent years have seen unprecedented investment and effort to distribute anthelmintic drugs to millions of people in mass drug administration programmes. Onchocerciasis has been eliminated from four countries in the Americas [57,58], with the prospect of elimination from African countries in the coming decade. These advances have been made with our existing major anthelmintic drugs. In addition, there has been great progress in finding more effective combinations of drugs, as well as bringing forward trials of veterinary anthelmintics to combat human disease. However, the risk of mass drug administration leading to resistance, as well as the growing understanding that existing drugs are not ideal for all human helminth infections, has led to new focus on the need to develop new anthelmintics. Unfortunately, there is limited economic incentive to fund drug development adequately.

Open source science seeks to radically open up the drug development process, with the goal of increasing the efficiency, reducing the cost of research duplication, reducing the hoarding of data and creating collaborative communities [174]. This approach has been applied to the discovery of compounds with antimalarial activity [82]. In the field of anthelmintic drug discovery there is much open science. In this review we particularly highlight open-source assay systems that have been developed and openly released by several groups that enabled compound screening against different helminth parasites. The genomic information about helminths is rapidly expanding and most is released freely and can be queried by scientists around the world, helping to find new anthelmintic drug targets. Probably the weakest area for open science is compound screening and subsequent drug development. The authors of this review are probably as guilty as other members of the community. Despite having good intentions and releasing screening data with publications, we could all do more to release data soon after collection, rather than being constrained by academic and publishing timescales. The MMV Pathogen box has demonstrated how many different groups around the world can be recruited to screen compound libraries in their specialist assays, and release data in a relatively timely manner – researchers are asked, as a condition of receiving the compounds, to share any data generated in the public domain within two years.

What are the barriers to making anthelmintic drug discovery more open? Perhaps more could be done to facilitate data sharing. While existing databases such as PubChem and ChEMBL can be used to share screening data, a specialised anthelmintic screening and target database could promote more widespread use and release of data in a consistent and accessible format, enabling more collaborative working and reuse of data. Wider awareness and sharing of open data, will also help spread the knowledge that data can be shared before publication, without diluting academic credit and still allowing later publication. We encourage the community, particularly journal editors and reviewers, to support open science. Ultimately, we are all working in this field to find new medicines to help the millions of people infected with helminths, and data sharing and open science can only expedite this aim.

Funding

FAP and DBS acknowledge the support of Medical Research Council grant MR/N024842/1. RF and KJE acknowledge the support of Medical Research Council grant MR/N022661/1.

ORCID® iDs

Frederick A. Partridge - https://orcid.org/0000-0001-6236-4297
Carole J. R. Bataille - https://orcid.org/0000-0002-3963-0021
Marina Nick - https://orcid.org/0000-0002-9972-963X
Angela J. Russell - https://orcid.org/0000-0003-3610-9369
Kathryn J. Else - https://orcid.org/0000-0001-6660-055X
David B. Sattelle - https://orcid.org/0000-0003-0705-5242

References

1. Doyle, S. R.; Cotton, J. A. *Trends Parasitol.* **2019,** *35,* 289–301. doi:10.1016/j.pt.2019.01.004
2. Epe, C.; Kaminsky, R. *Trends Parasitol.* **2013,** *29,* 129–134. doi:10.1016/j.pt.2013.01.001
3. Morgan, E. R.; Aziz, N.-A. A.; Blanchard, A.; Charlier, J.; Charvet, C.; Claerebout, E.; Geldhof, P.; Greer, A. W.; Hertzberg, H.; Hodgkinson, J.; Höglund, J.; Hoste, H.; Kaplan, R. M.; Martínez-Valladares, M.; Mitchell, S.; Ploeger, H. W.; Rinaldi, L.; von Samson-Himmelstjerna, G.; Sotiraki, S.; Schnyder, M.; Skuce, P.; Bartley, D.; Kenyon, F.; Thamsborg, S. M.; Vineer, H. R.; de Waal, T.; Williams, A. R.; van Wyk, J. A.; Vercruysse, J. *Trends Parasitol.* **2019,** *35,* 52–71. doi:10.1016/j.pt.2018.10.006
4. Hotez, P. J.; Molyneux, D. H.; Fenwick, A.; Kumaresan, J.; Sachs, S. E.; Sachs, J. D.; Savioli, L. *N. Engl. J. Med.* **2007,** *357,* 1018–1027. doi:10.1056/nejmra064142
5. GBD 2016 DALYs and HALE Collaborators. *Lancet* **2017,** *390,* 1260–1344. doi:10.1016/s0140-6736(17)32130-x
6. Hotez, P. *PLoS Neglected Trop. Dis.* **2007,** *1,* e118. doi:10.1371/journal.pntd.0000118
7. Molyneux, D. H.; Dean, L.; Adekeye, O.; Stothard, J. R.; Theobald, S. *Parasitology* **2018,** *145,* 1647–1654. doi:10.1017/s0031182018000069
8. Pullan, R. L.; Smith, J. L.; Jasrasaria, R.; Brooker, S. J. *Parasites Vectors* **2014,** *7,* 37. doi:10.1186/1756-3305-7-37
9. Global Burden of Disease Study 2013 Collaborators. *Lancet* **2015,** *386,* 743–800. doi:10.1016/s0140-6736(15)60692-4
10. de Silva, N. R.; Chan, M. S.; Bundy, D. A. *Trop. Med. Int. Health* **1997,** *2,* 519–528. doi:10.1046/j.1365-3156.1997.d01-320.x
11. de Silva, N. R.; Guyatt, H. L.; Bundy, D. A. *Trans. R. Soc. Trop. Med. Hyg.* **1997,** *91,* 31–36. doi:10.1016/s0035-9203(97)90384-9

12. Murray, C. J. L.; Vos, T.; Lozano, R.; Naghavi, M.; Flaxman, A. D.; Michaud, C.; Ezzati, M.; Shibuya, K.; Salomon, J. A.; Abdalla, S.; Aboyans, V.; Abraham, J.; Ackerman, I.; Aggarwal, R.; Ahn, S. Y.; Ali, M. K.; Alvarado, M.; Anderson, H. R.; Anderson, L. M.; Andrews, K. G.; Atkinson, C.; Baddour, L. M.; Bahalim, A. N.; Barker-Collo, S.; Barrero, L. H.; Bartels, D. H.; Basáñez, M.-G.; Baxter, A.; Bell, M. L.; Benjamin, E. J.; Bennett, D.; Bernabé, E.; Bhalla, K.; Bhandari, B.; Bikbov, B.; Bin Abdulhak, A.; Birbeck, G.; Black, J. A.; Blencowe, H.; Blore, J. D.; Blyth, F.; Bolliger, I.; Bonaventure, A.; Boufous, S.; Bourne, R.; Boussinesq, M.; Braithwaite, T.; Brayne, C.; Bridgett, L.; Brooker, S.; Brooks, P.; Brugha, T. S.; Bryan-Hancock, C.; Bucello, C.; Buchbinder, R.; Buckle, G.; Budke, C. M.; Burch, M.; Burney, P.; Burstein, R.; Calabria, B.; Campbell, B.; Canter, C. E.; Carabin, H.; Carapetis, J.; Carmona, L.; Cella, C.; Charlson, F.; Chen, H.; Cheng, A. T.-A.; Chou, D.; Chugh, S. S.; Coffeng, L. E.; Colan, S. D.; Colquhoun, S.; Colson, K. E.; Condon, J.; Connor, M. D.; Cooper, L. T.; Corriere, M.; Cortinovis, M.; de Vaccaro, K. C.; Couser, W.; Cowie, B. C.; Criqui, M. H.; Cross, M.; Dabhadkar, K. C.; Dahiya, M.; Dahodwala, N.; Damsere-Derry, J.; Danaei, G.; Davis, A.; De Leo, D.; Degenhardt, L.; Dellavalle, R.; Delossantos, A.; Denenberg, J.; Derrett, S.; Des Jarlais, D. C.; Dharmaratne, S. D.; Dherani, M.; Diaz-Torne, C.; Dolk, H.; Dorsey, E. R.; Driscoll, T.; Duber, H.; Ebel, B.; Edmond, K.; Elbaz, A.; Ali, S. E.; Erskine, H.; Erwin, P. J.; Espindola, P.; Ewoigbokhan, S. E.; Farzadfar, F.; Feigin, V.; Felson, D. T.; Ferrari, A.; Ferri, C. P.; Fèvre, E. M.; Finucane, M. M.; Flaxman, S.; Flood, L.; Foreman, K.; Forouzanfar, M. H.; Fowkes, F. G. R.; Fransen, M.; Freeman, M. K.; Gabbe, B. J.; Gabriel, S. E.; Gakidou, E.; Ganatra, H. A.; Garcia, B.; Gaspari, F.; Gillum, R. F.; Gmel, G.; Gonzalez-Medina, D.; Gosselin, R.; Grainger, R.; Grant, B.; Groeger, J.; Guillemin, F.; Gunnell, D.; Gupta, R.; Haagsma, J.; Hagan, H.; Halasa, Y. A.; Hall, W.; Haring, D.; Haro, J. M.; Harrison, J. E.; Havmoeller, R.; Hay, R. J.; Higashi, H.; Hill, C.; Hoen, B.; Hoffman, H.; Hotez, P. J.; Hoy, D.; Huang, J. J.; Ibeanusi, S. E.; Jacobsen, K. H.; James, S. L.; Jarvis, D.; Jasrasaria, R.; Jayaraman, S.; Johns, N.; Jonas, J. B.; Karthikeyan, G.; Kassebaum, N.; Kawakami, N.; Keren, A.; Khoo, J.-P.; King, C. H.; Knowlton, L. M.; Kobusingye, O.; Koranteng, A.; Krishnamurthi, R.; Laden, F.; Lalloo, R.; Laslett, L. L.; Lathlean, T.; Leasher, J. L.; Lee, Y. Y.; Leigh, J.; Levinson, D.; Lim, S. S.; Limb, E.; Lin, J. K.; Lipnick, M.; Lipshultz, S. E.; Liu, W.; Loane, M.; Ohno, S. L.; Lyons, R.; Mabweijano, J.; MacIntyre, M. F.; Malekzadeh, R.; Mallinger, L.; Manivannan, S.; Marcenes, W.; March, L.; Margolis, D. J.; Marks, G. B.; Marks, R.; Matsumori, A.; Matzopoulos, R.; Mayosi, B. M.; McAnulty, J. H.; McDermott, M. M.; McGill, N.; McGrath, J.; Medina-Mora, M. E.; Meltzer, M.; Mensah, G. A.; Merriman, T. R.; Meyer, A.-C.; Miglioli, V.; Miller, M.; Miller, T. R.; Mitchell, P. B.; Mock, C.; Mocumbi, A. O.; Moffitt, T. E.; Mokdad, A. A.; Monasta, L.; Montico, M.; Moradi-Lakeh, M.; Moran, A.; Morawska, L.; Mori, R.; Murdoch, M. E.; Mwaniki, M. K.; Naidoo, K.; Nair, M. N.; Naldi, L.; Narayan, K. M. V.; Nelson, P. K.; Nelson, R. G.; Nevitt, M. C.; Newton, C. R.; Nolte, S.; Norman, P.; Norman, R.; O'Donnell, M.; O'Hanlon, S.; Olives, C.; Omer, S. B.; Ortblad, K.; Osborne, R.; Ozgediz, D.; Page, A.; Pahari, B.; Pandian, J. D.; Rivero, A. P.; Patten, S. B.; Pearce, N.; Padilla, R. P.; Perez-Ruiz, F.; Perico, N.; Pesudovs, K.; Phillips, D.; Phillips, M. R.; Pierce, K.; Pion, S.; Polanczyk, G. V.; Polinder, S.; Pope, C. A.; Popova, S.; Porrini, E.; Pourmalek, F.; Prince, M.; Pullan, R. L.; Ramaiah, K. D.; Ranganathan, D.; Razavi, H.; Regan, M.; Rehm, J. T.; Rein, D. B.; Remuzzi, G.; Richardson, K.; Rivara, F. P.; Roberts, T.; Robinson, C.; De Leòn, F. R.; Ronfani, L.; Room, R.; Rosenfeld, L. C.; Rushton, L.; Sacco, R. L.; Saha, S.; Sampson, U.; Sanchez-Riera, L.; Sanman, E.; Schwebel, D. C.; Scott, J. G.; Segui-Gomez, M.; Shahraz, S.; Shepard, D. S.; Shin, H.; Shivakoti, R.; Singh, D.; Singh, G. M.; Singh, J. A.; Singleton, J.; Sleet, D. A.; Sliwa, K.; Smith, E.; Smith, J. L.; Stapelberg, N. J. C.; Steer, A.; Steiner, T.; Stolk, W. A.; Stovner, L. J.; Sudfeld, C.; Syed, S.; Tamburlini, G.; Tavakkoli, M.; Taylor, H. R.; Taylor, J. A.; Taylor, W. J.; Thomas, B.; Thomson, W. M.; Thurston, G. D.; Tleyjeh, I. M.; Tonelli, M.; Towbin, J. A.; Truelsen, T.; Tsilimbaris, M. K.; Ubeda, C.; Undurraga, E. A.; van der Werf, M. J.; van Os, J.; Vavilala, M. S.; Venketasubramanian, N.; Wang, M.; Wang, W.; Watt, K.; Weatherall, D. J.; Weinstock, M. A.; Weintraub, R.; Weisskopf, M. G.; Weissman, M. M.; White, R. A.; Whiteford, H.; Wiebe, N.; Wiersma, S. T.; Wilkinson, J. D.; Williams, H. C.; Williams, S. R. M.; Witt, E.; Wolfe, F.; Woolf, A. D.; Wulf, S.; Yeh, P.-H.; Zaidi, A. K. M.; Zheng, Z.-J.; Zonies, D.; Lopez, A. D.; AlMazroa, M. A.; Memish, Z. A. *Lancet* **2012,** *380,* 2197–2223. doi:10.1016/s0140-6736(12)61689-4
13. Keiser, J.; Tritten, L.; Silbereisen, A.; Speich, B.; Adelfio, R.; Vargas, M. *PLoS Neglected Trop. Dis.* **2013,** *7,* e2119. doi:10.1371/journal.pntd.0002119
14. Hotez, P.; Whitham, M. *Obstet. Gynecol.* **2014,** *123,* 155–160. doi:10.1097/aog.0000000000000025
15. Gardner, J. M.; Grantham-Mcgregor, S.; Baddeley, A. *Ann. Trop. Med. Parasitol.* **1996,** *90,* 55–63. doi:10.1080/00034983.1996.11813026
16. Simeon, D. T.; Grantham-McGregor, S. M.; Callender, J. E.; Wong, M. S. *J. Nutr.* **1995,** *125,* 1875–1883. doi:10.1093/jn/125.7.1875
17. Stephenson, L. S.; Latham, M. C.; Adams, E. J.; Kinoti, S. N.; Pertet, A. *J. Nutr.* **1993,** *123,* 1036–1046.
18. *Progress Report 2000–2009 and Strategic Plan 2010–2020 of the Global Programme to Eliminate Lymphatic Filariasis: Halfway towards Eliminating Lymphatic Filariasis;* World Health Organization, 2010.
19. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. *Lancet* **2018,** *392,* 1789–1858. doi:10.1016/s0140-6736(18)32279-7
20. McManus, D. P.; Dunne, D. W.; Sacko, M.; Utzinger, J.; Vennervald, B. J.; Zhou, X.-N. *Nat. Rev. Dis. Primers* **2018,** *4,* 13. doi:10.1038/s41572-018-0013-8
21. Hotez, P. *PLoS Neglected Trop. Dis.* **2007,** *1,* e77. doi:10.1371/journal.pntd.0000077
22. Marocco, C.; Bangert, M.; Joseph, S. A.; Fitzpatrick, C.; Montresor, A. *Trans. R. Soc. Trop. Med. Hyg.* **2017,** *111,* 12–17. doi:10.1093/trstmh/trx011
23. Bah, Y. M.; Bah, M. S.; Paye, J.; Conteh, A.; Saffa, S.; Tia, A.; Sonnie, M.; Veinoglou, A.; Amon, J. J.; Hodges, M. H.; Zhang, Y. *Infect. Dis. Poverty* **2019,** *8,* 41. doi:10.1186/s40249-019-0553-5
24. Shumbej, T.; Menu, S.; Girum, T.; Bekele, F.; Gebru, T.; Worku, M.; Dendir, A.; Solomon, A.; Kahase, D.; Alemayehu, M. *Res. Rep. Trop. Med.* **2019,** *10,* 109–118. doi:10.2147/rrtm.s208473
25. Freeman, M. C.; Akogun, O.; Belizario, V.; Brooker, S. J.; Gyorkos, T. W.; Imtiaz, R.; Krolewiecki, A.; Lee, S.; Matendechero, S. H.; Pullan, R. L.; Utzinger, J. *PLoS Neglected Trop. Dis.* **2019,** *13,* e0007201. doi:10.1371/journal.pntd.0007201
26. *WHO Model List of Essential Medicines, 20th List.;* World Health Organization, 2017.

27. De Clercq, D.; Sacko, M.; Behnke, J.; Gilbert, F.; Dorny, P.; Vercruysse, J. *Am. J. Trop. Med. Hyg.* **1997,** *57,* 25–30. doi:10.4269/ajtmh.1997.57.25
28. Soukhathammavong, P. A.; Sayasone, S.; Phongluxa, K.; Xayaseng, V.; Utzinger, J.; Vounatsou, P.; Hatz, C.; Akkhavong, K.; Keiser, J.; Odermatt, P. *PLoS Neglected Trop. Dis.* **2012,** *6,* e1417. doi:10.1371/journal.pntd.0001417
29. Moser, W.; Schindler, C.; Keiser, J. *BMJ [Br. Med. J.]* **2017,** *358,* j4307. doi:10.1136/bmj.j4307
30. Moser, W.; Schindler, C.; Keiser, J. *Adv. Parasitol.* **2019,** *103,* 91–115. doi:10.1016/bs.apar.2018.08.002
31. Demeler, J.; Krüger, N.; Krücken, J.; von der Heyden, V. C.; Ramünke, S.; Küttler, U.; Miltsch, S.; López Cepeda, M.; Knox, M.; Vercruysse, J.; Geldhof, P.; Harder, A.; von Samson-Himmelstjerna, G. *PLoS One* **2013,** *8,* e70212. doi:10.1371/journal.pone.0070212
32. Geurden, T.; Hoste, H.; Jacquiet, P.; Traversa, D.; Sotiraki, S.; Frangipane di Regalbono, A.; Tzanidakis, N.; Kostopoulou, D.; Gaillac, C.; Privat, S.; Giangaspero, A.; Zanardello, C.; Noé, L.; Vanimisetti, B.; Bartram, D. *Vet. Parasitol.* **2014,** *201,* 59–66. doi:10.1016/j.vetpar.2014.01.016
33. Matamoros, G.; Rueda, M. M.; Rodríguez, C.; Gabrie, J. A.; Canales, M.; Fontecha, G.; Sanchez, A. *Trop. Med. Infect. Dis.* **2019,** *4,* 73. doi:10.3390/tropicalmed4020073
34. Zuccherato, L. W.; Furtado, L. F.; da Silva Medeiros, C.; da Silva Pinheiro, C.; Rabelo, É. M. *PLoS Neglected Trop. Dis.* **2018,** *12,* e0006766. doi:10.1371/journal.pntd.0006766
35. Orr, A. R.; Quagraine, J. E.; Suwondo, P.; George, S.; Harrison, L. M.; Dornas, F. P.; Evans, B.; Caccone, A.; Humphries, D.; Wilson, M. D.; Cappello, M. *Am. J. Trop. Med. Hyg.* **2019,** *100,* 351–356. doi:10.4269/ajtmh.18-0727
36. Vercruysse, J.; Albonico, M.; Behnke, J. M.; Kotze, A. C.; Prichard, R. K.; McCarthy, J. S.; Montresor, A.; Levecke, B. *Int. J. Parasitol.: Drugs Drug Resist.* **2011,** *1,* 14–27. doi:10.1016/j.ijpddr.2011.09.002
37. Brooker, S.; Clements, A. C. A.; Bundy, D. A. P. Global Epidemiology, Ecology and Control of Soil-Transmitted Helminth Infections. In *Advances in Parasitology;* Hay, S. I.; Graham, A.; Rogers, D. J., Eds.; Global Mapping of Infectious Diseases: Methods, Examples and Emerging Applications, Vol. 62; Academic Press, 2006; pp 221–261. doi:10.1016/s0065-308x(05)62007-6
38. Ziegelbauer, K.; Speich, B.; Mäusezahl, D.; Bos, R.; Keiser, J.; Utzinger, J. *PLoS Med.* **2012,** *9,* e1001162. doi:10.1371/journal.pmed.1001162
39. Partridge, F. A.; Forman, R.; Willis, N. J.; Bataille, C. J. R.; Murphy, E. A.; Brown, A. E.; Heyer-Chauhan, N.; Marinič, B.; Sowood, D. J. C.; Wynne, G. M.; Else, K. J.; Russell, A. J.; Sattelle, D. B. *PLoS Neglected Trop. Dis.* **2018,** *12,* e0006487. doi:10.1371/journal.pntd.0006487
40. WHO Strategy: Intestinal worms. http://www.who.int/intestinal_worms/strategy/en/ (accessed Oct 29, 2019).
41. Olliaro, P.; Delgado-Romero, P.; Keiser, J. *J. Antimicrob. Chemother.* **2014,** *69,* 863–870. doi:10.1093/jac/dkt491
42. Bagchus, W. M.; Bezuidenhout, D.; Harrison-Moench, E.; Kourany-Lefoll, E.; Wolna, P.; Yalkinoglu, O. *Clin. Transl. Sci.* **2019,** *12,* 66–76. doi:10.1111/cts.12601
43. Doenhoff, M. J.; Cioli, D.; Utzinger, J. *Curr. Opin. Infect. Dis.* **2008,** *21,* 659–667. doi:10.1097/qco.0b013e328318978f
44. Zwang, J.; Olliaro, P. L. *PLoS Neglected Trop. Dis.* **2014,** *8,* e3286. doi:10.1371/journal.pntd.0003286
45. Crellen, T.; Walker, M.; Lamberton, P. H. L.; Kabatereine, N. B.; Tukahebwa, E. M.; Cotton, J. A.; Webster, J. P. *Clin. Infect. Dis.* **2016,** *63,* 1151–1159. doi:10.1093/cid/ciw506
46. Wang, W.; Wang, L.; Liang, Y.-S. *Parasitol. Res.* **2012,** *111,* 1871–1877. doi:10.1007/s00436-012-3151-z
47. McCreesh, N.; Nikulin, G.; Booth, M. *Parasites Vectors* **2015,** *8,* 4. doi:10.1186/s13071-014-0617-0
48. Toor, J.; Alsallaq, R.; Truscott, J. E.; Turner, H. C.; Werkman, M.; Gurarie, D.; King, C. H.; Anderson, R. M. *Clin. Infect. Dis.* **2018,** *66* (Suppl. 4), S245–S252. doi:10.1093/cid/ciy001
49. World Health Organization. *Wkly. Epidemiol. Rec.* **2019,** *94,* 457–470.
50. King, C. L.; Suamani, J.; Sanuku, N.; Cheng, Y.-C.; Satofan, S.; Mancuso, B.; Goss, C. W.; Robinson, L. J.; Siba, P. M.; Weil, G. J.; Kazura, J. W. *N. Engl. J. Med.* **2018,** *379,* 1801–1810. doi:10.1056/nejmoa1706854
51. Bjerum, C. M.; Ouattara, A. F.; Aboulaye, M.; Kouadio, O.; Marius, V. K.; Andersen, B. J.; Weil, G. J.; Koudou, B. G.; King, C. L. *Clin. Infect. Dis.*, in press. doi:10.1093/cid/ciz1050
52. *Guideline: Alternative Mass Drug Administration Regimens to Eliminate Lymphatic Filariasis;* World Health Organization, 2017.
53. Sharma, R.; Al Jayoussi, G.; Tyrer, H. E.; Gamble, J.; Hayward, L.; Guimaraes, A. F.; Davies, J.; Waterhouse, D.; Cook, D. A. N.; Myhill, L. J.; Clare, R. H.; Cassidy, A.; Steven, A.; Johnston, K. L.; Ford, L.; Turner, J. D.; Ward, S. A.; Taylor, M. J. *Sci. Rep.* **2016,** *6,* 23458. doi:10.1038/srep23458
54. Turner, J. D.; Tendongfor, N.; Esum, M.; Johnston, K. L.; Langley, R. S.; Ford, L.; Faragher, B.; Specht, S.; Mand, S.; Hoerauf, A.; Enyong, P.; Wanji, S.; Taylor, M. J. *PLoS Neglected Trop. Dis.* **2010,** *4,* e660. doi:10.1371/journal.pntd.0000660
55. World Health Organization. *Wkly. Epidemiol. Rec.* **2019,** *94,* 513–523.
56. Tekle, A. H.; Zouré, H. G. M.; Noma, M.; Boussinesq, M.; Coffeng, L. E.; Stolk, W. A.; Remme, J. H. F. *Infect. Dis. Poverty* **2016,** *5,* 66. doi:10.1186/s40249-016-0160-7
57. Nicholls, R. S.; Duque, S.; Olaya, L. A.; López, M. C.; Sánchez, S. B.; Morales, A. L.; Palma, G. I. *Parasites Vectors* **2018,** *11,* 237. doi:10.1186/s13071-018-2821-9
58. Guevara, Á.; Lovato, R.; Proaño, R.; Rodriguez-Perez, M. A.; Unnasch, T.; Cooper, P. J.; Guderian, R. H. *Parasites Vectors* **2018,** *11,* 265. doi:10.1186/s13071-018-2851-3
59. Opoku, N. O.; Bakajika, D. K.; Kanza, E. M.; Howard, H.; Mambandu, G. L.; Nyathirombo, A.; Nigo, M. M.; Kasonia, K.; Masembe, S. L.; Mumbere, M.; Kataliko, K.; Larbelee, J. P.; Kpawor, M.; Bolay, K. M.; Bolay, F.; Asare, S.; Attah, S. K.; Olipoh, G.; Vaillant, M.; Halleux, C. M.; Kuesel, A. C. *Lancet* **2018,** *392,* 1207–1216. doi:10.1016/s0140-6736(17)32844-1
60. Kuesel, A. C. *Int. J. Parasitol.: Drugs Drug Resist.* **2016,** *6,* 272–286. doi:10.1016/j.ijpddr.2016.04.002
61. Zhan, B.; Beaumier, C. M.; Briggs, N.; Jones, K. M.; Keegan, B. P.; Bottazzi, M. E.; Hotez, P. J. *Expert Rev. Vaccines* **2014,** *13,* 321–331. doi:10.1586/14760584.2014.872035
62. Tsuji, N.; Suzuki, K.; Kasuga-Aoki, H.; Matsumoto, Y.; Arakawa, T.; Ishiwata, K.; Isobe, T. *Infect. Immun.* **2001,** *69,* 7285–7292. doi:10.1128/iai.69.12.7285-7292.2001
63. Dixon, H.; Little, M. C.; Else, K. J. *Int. J. Parasitol.* **2010,** *40,* 683–693. doi:10.1016/j.ijpara.2009.11.008

64. Wei, J.; Versteeg, L.; Liu, Z.; Keegan, B.; Gazzinelli-Guimarães, A. C.; Fujiwara, R. T.; Briggs, N.; Jones, K. M.; Strych, U.; Beaumier, C. M.; Bottazzi, M. E.; Hotez, P. J.; Zhan, B. *PLoS Neglected Trop. Dis.* **2017,** *11,* e0005769. doi:10.1371/journal.pntd.0005769
65. Briggs, N.; Wei, J.; Versteeg, L.; Zhan, B.; Keegan, B.; Damania, A.; Pollet, J.; Hayes, K. S.; Beaumier, C.; Seid, C. A.; Leong, J.; Grencis, R. K.; Bottazzi, M. E.; Sastry, K. J.; Hotez, P. J. *PLoS Pathog.* **2018,** *14,* e1007273. doi:10.1371/journal.ppat.1007273
66. Loukas, A.; Hotez, P. J.; Diemert, D.; Yazdanbakhsh, M.; McCarthy, J. S.; Correa-Oliveira, R.; Croese, J.; Bethony, J. M. *Nat. Rev. Dis. Primers* **2016,** *2,* 16088. doi:10.1038/nrdp.2016.88
67. Riveau, G.; Deplanque, D.; Remoué, F.; Schacht, A.-M.; Vodougnon, H.; Capron, M.; Thiry, M.; Martial, J.; Libersa, C.; Capron, A. *PLoS Neglected Trop. Dis.* **2012,** *6,* e1704. doi:10.1371/journal.pntd.0001704
68. Lacerda, M. V. G.; Llanos-Cuentas, A.; Krudsood, S.; Lon, C.; Saunders, D. L.; Mohammed, R.; Yilma, D.; Batista Pereira, D.; Espino, F. E. J.; Mia, R. Z.; Chuquiyauri, R.; Val, F.; Casapía, M.; Monteiro, W. M.; Brito, M. A. M.; Costa, M. R. F.; Buathong, N.; Noedl, H.; Diro, E.; Getie, S.; Wubie, K. M.; Abdissa, A.; Zeynudin, A.; Abebe, C.; Tada, M. S.; Brand, F.; Beck, H.-P.; Angus, B.; Duparc, S.; Kleim, J.-P.; Kellam, L. M.; Rousell, V. M.; Jones, S. W.; Hardaker, E.; Mohamed, K.; Clover, D. D.; Fletcher, K.; Breton, J. J.; Ugwuegbulam, C. O.; Green, J. A.; Koh, G. C. K. W. *N. Engl. J. Med.* **2019,** *380,* 215–228. doi:10.1056/nejmoa1710775
69. Mesu, V. K. B. K.; Kalonji, W. M.; Bardonneau, C.; Mordt, O. V.; Blesson, S.; Simon, F.; Delhomme, S.; Bernhard, S.; Kuziena, W.; Lubaki, J.-P. F.; Vuvu, S. L.; Ngima, P. N.; Mbembo, H. M.; Ilunga, M.; Bonama, A. K.; Heradi, J. A.; Solomo, J. L. L.; Mandula, G.; Badibabi, L. K.; Dama, F. R.; Lukula, P. K.; Tete, D. N.; Lumbala, C.; Scherrer, B.; Strub-Wourgaft, N.; Tarral, A. *Lancet* **2018,** *391,* 144–154. doi:10.1016/s0140-6736(17)32758-7
70. Holden-Dye, L.; Walker, R. J. Anthelmintic drugs and nematicides: studies in *Caenorhabditis elegans*. *WormBook;* 2014. doi:10.1895/wormbook.1.143.2
71. Pedrique, B.; Strub-Wourgaft, N.; Some, C.; Olliaro, P.; Trouiller, P.; Ford, N.; Pécoul, B.; Bradol, J.-H. *Lancet Global Health* **2013,** *1,* e371–e379. doi:10.1016/s2214-109x(13)70078-0
72. Weng, H.-B.; Chen, H.-X.; Wang, M.-W. *Infect. Dis. Poverty* **2018,** *7,* 67. doi:10.1186/s40249-018-0444-1
73. Berman, J.; Radhakrishna, T. *Am. J. Trop. Med. Hyg.* **2017,** *96,* 11–13. doi:10.4269/ajtmh.16-0099
74. de Hostos, E. L.; Nguyen, T. Anthelmintic Drugs: Tools and Shortcuts for the Long Road from Discovery to Product. *Parasitic Helminths;* Wiley Blackwell, 2012; pp 217–232. doi:10.1002/9783527652969.ch13
75. Pelfrene, E.; Harvey Allchurch, M.; Ntamabyaliro, N.; Nambasa, V.; Ventura, F. V.; Nagercoil, N.; Cavaleri, M. *PLoS Neglected Trop. Dis.* **2019,** *13,* e0007381. doi:10.1371/journal.pntd.0007381
76. Taylor, M. J.; von Geldern, T. W.; Ford, L.; Hübner, M. P.; Marsh, K.; Johnston, K. L.; Sjoberg, H. T.; Specht, S.; Pionnier, N.; Tyrer, H. E.; Clare, R. H.; Cook, D. A. N.; Murphy, E.; Steven, A.; Archer, J.; Bloemker, D.; Lenz, F.; Koschel, M.; Ehrens, A.; Metuge, H. M.; Chunda, V. C.; Ndongmo Chounna, P. W.; Njouendou, A. J.; Fombad, F. F.; Carr, R.; Morton, H. E.; Aljayyoussi, G.; Hoerauf, A.; Wanji, S.; Kempf, D. J.; Turner, J. D.; Ward, S. A. *Sci. Transl. Med.* **2019,** *11,* eaau2086. doi:10.1126/scitranslmed.aau2086
77. Ramamoorthi, R.; Graef, K. M.; Dent, J. *Int. J. Parasitol.: Drugs Drug Resist.* **2014,** *4,* 220–225. doi:10.1016/j.ijpddr.2014.09.002
78. Todd, M. H. *ChemMedChem* **2019,** *14,* 1804–1809. doi:10.1002/cmdc.201900565
79. Mendez, D.; Gaulton, A.; Bento, A. P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M. P.; Mosquera, J. F.; Mutowo, P.; Nowotka, M.; Gordillo-Marañón, M.; Hunter, F.; Junco, L.; Mugumbate, G.; Rodriguez-Lopez, M.; Atkinson, F.; Bosc, N.; Radoux, C. J.; Segura-Cabrera, A.; Hersey, A.; Leach, A. R. *Nucleic Acids Res.* **2019,** *47,* D930–D940. doi:10.1093/nar/gky1075
80. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. *Nucleic Acids Res.* **2019,** *47,* D1102–D1109. doi:10.1093/nar/gky1033
81. OSM - Open Source Malaria. http://opensourcemalaria.org/ (accessed Nov 3, 2019).
82. Williamson, A. E.; Ylioja, P. M.; Robertson, M. N.; Antonova-Koch, Y.; Avery, V.; Baell, J. B.; Batchu, H.; Batra, S.; Burrows, J. N.; Bhattacharyya, S.; Calderon, F.; Charman, S. A.; Clark, J.; Crespo, B.; Dean, M.; Debbert, S. L.; Delves, M.; Dennis, A. S. M.; Deroose, F.; Duffy, S.; Fletcher, S.; Giaever, G.; Hallyburton, I.; Gamo, F.-J.; Gebbia, M.; Guy, R. K.; Hungerford, Z.; Kirk, K.; Lafuente-Monasterio, M. J.; Lee, A.; Meister, S.; Nislow, C.; Overington, J. P.; Papadatos, G.; Patiny, L.; Pham, J.; Ralph, S. A.; Ruecker, A.; Ryan, E.; Southan, C.; Srivastava, K.; Swain, C.; Tarnowski, M. J.; Thomson, P.; Turner, P.; Wallace, I. M.; Wells, T. N. C.; White, K.; White, L.; Willis, P.; Winzeler, E. A.; Wittlin, S.; Todd, M. H. *ACS Cent. Sci.* **2016,** *2,* 687–701. doi:10.1021/acscentsci.6b00086
83. Geary, T. G.; Sakanari, J. A.; Caffrey, C. R. *J. Parasitol.* **2015,** *101,* 125–133. doi:10.1645/14-703.1
84. Brenner, S. *Genetics* **1974,** *77,* 71–94.
85. The C. elegans Sequencing Consortium. *Science* **1998,** *282,* 2012–2018. doi:10.1126/science.282.5396.2012
86. Sulston, J. E.; Horvitz, H. R. *Dev. Biol. (Amsterdam, Neth.)* **1977,** *56,* 110–156. doi:10.1016/0012-1606(77)90158-0
87. White, J. G.; Southgate, E.; Thomson, J. N.; Brenner, S. *Philos. Trans. R. Soc., B* **1986,** *314,* 1–340. doi:10.1098/rstb.1986.0056
88. Culetto, E.; Sattelle, D. B. *Hum. Mol. Genet.* **2000,** *9,* 869–877. doi:10.1093/hmg/9.6.869
89. Jones, A. K.; Buckingham, S. D.; Sattelle, D. B. *Nat. Rev. Drug Discovery* **2005,** *4,* 321–330. doi:10.1038/nrd1692
90. Burns, A. R.; Wallace, I. M.; Wildenhain, J.; Tyers, M.; Giaever, G.; Bader, G. D.; Nislow, C.; Cutler, S. R.; Roy, P. J. *Nat. Chem. Biol.* **2010,** *6,* 549–557. doi:10.1038/nchembio.380
91. Caenorhabditis Genetics Center (CGC). https://cgc.umn.edu/ (accessed Nov 3, 2019).
92. Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature* **1998,** *391,* 806–811. doi:10.1038/35888
93. Fankhauser, R.; Cozzie, L. R.; Nare, B.; Powell, K.; Sluder, A. E.; Hammerland, L. G. Use of Rodent Models in the Discovery of Novel Anthelmintics. *Parasitic Helminths;* Wiley Blackwell, 2012; pp 181–199. doi:10.1002/9783527652969.ch11
94. Burns, A. R.; Luciani, G. M.; Musso, G.; Bagg, R.; Yeo, M.; Zhang, Y.; Rajendran, L.; Glavin, J.; Hunter, R.; Redman, E.; Stasiuk, S.; Schertzberg, M.; Angus McQuibban, G.; Caffrey, C. R.; Cutler, S. R.; Tyers, M.; Giaever, G.; Nislow, C.; Fraser, A. G.; MacRae, C. A.; Gilleard, J.; Roy, P. J. *Nat. Commun.* **2015,** *6,* 7485. doi:10.1038/ncomms8485

95. Glendinning, S. K.; Buckingham, S. D.; Sattelle, D. B.; Wonnacott, S.; Wolstenholme, A. J. *PLoS One* **2011,** *6,* e22390. doi:10.1371/journal.pone.0022390
96. Sloan, M. A.; Reaves, B. J.; Maclean, M. J.; Storey, B. E.; Wolstenholme, A. J. *Mol. Biochem. Parasitol.* **2015,** *204,* 44–50. doi:10.1016/j.molbiopara.2015.12.006
97. Hulme, S. E.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **2011,** *50,* 4774–4807. doi:10.1002/anie.201005461
98. Elfawal, M. A.; Savinov, S. N.; Aroian, R. V. *Sci. Rep.* **2019,** *9,* 12347. doi:10.1038/s41598-019-48720-1
99. Nigon, V. M.; Félix, M.-A. History of research on C. elegans and other free-living nematodes as model organisms. *WormBook;* 2017. doi:10.1895/wormbook.1.181.1
100. Ankeny, R. A. *Nat. Rev. Genet.* **2001,** *2,* 474–479. doi:10.1038/35076538
101. Corsi, A. K.; Wightman, B.; Chalfie, M. *Genetics* **2015,** *200,* 387–407. doi:10.1534/genetics.115.176099
102. Chalfie, M.; Tu, Y.; Prasher, D. Glow Worms - A New Method of Looking at *C. elegans* Gene Expression: Worm Breeder's Gazette. http://www.wormbook.org/wli/wbg13.1p19/ (accessed Aug 15, 2019).
103. Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994,** *263,* 802–805. doi:10.1126/science.8303295
104. Wilson, R. K. *Trends Genet.* **1999,** *15,* 51–58. doi:10.1016/s0168-9525(98)01666-7
105. Stein, L.; Sternberg, P.; Durbin, R.; Thierry-Mieg, J.; Spieth, J. *Nucleic Acids Res.* **2001,** *29,* 82–86. doi:10.1093/nar/29.1.82
106. WormBase: Nematode Information Resource. https://wormbase.org/ (accessed Sept 12, 2019).
107. WormBase ParaSite. https://parasite.wormbase.org/ (accessed Sept 12, 2019).
108. Howe, K. L.; Bolt, B. J.; Shafie, M.; Kersey, P.; Berriman, M. *Mol. Biochem. Parasitol.* **2017,** *215,* 2–10. doi:10.1016/j.molbiopara.2016.11.005
109. International, Helminth Genomes Consortium. *Nat. Genet.* **2019,** *51,* 163–174. doi:10.1038/s41588-018-0262-1
110. Ittiprasert, W.; Mann, V. H.; Karinshak, S. E.; Coghlan, A.; Rinaldi, G.; Sankaranarayanan, G.; Chaidee, A.; Tanno, T.; Kumkhaek, C.; Prangtaworn, P.; Mentink-Kane, M. M.; Cochran, C. J.; Driguez, P.; Holroyd, N.; Tracey, A.; Rodpai, R.; Everts, B.; Hokke, C. H.; Hoffmann, K. F.; Berriman, M.; Brindley, P. J. *eLife* **2019,** *8,* e41337. doi:10.7554/elife.41337
111. Gang, S. S.; Castelletto, M. L.; Bryant, A. S.; Yang, E.; Mancuso, N.; Lopez, J. B.; Pellegrini, M.; Hallem, E. A. *PLoS Pathog.* **2017,** *13,* e1006675. doi:10.1371/journal.ppat.1006675
112. Dulovic, A.; Streit, A. *PLoS Pathog.* **2019,** *15,* e1007705. doi:10.1371/journal.ppat.1007705
113. Au, V.; Li-Leger, E.; Raymant, G.; Flibotte, S.; Chen, G.; Martin, K.; Fernando, L.; Doell, C.; Rosell, F. I.; Wang, S.; Edgley, M. L.; Rougvie, A. E.; Hutter, H.; Moerman, D. G. *G3: Genes, Genomes, Genet.* **2019,** *9,* 135–144. doi:10.1534/g3.118.200778
114. Mitani, S. *Proc. Jpn. Acad., Ser. B* **2017,** *93,* 561–577. doi:10.2183/pjab.93.036
115. Javer, A.; Currie, M.; Lee, C. W.; Hokanson, J.; Li, K.; Martineau, C. N.; Yemini, E.; Grundy, L. J.; Li, C.; Ch'ng, Q.; Schafer, W. R.; Nollen, E. A. A.; Kerr, R.; Brown, A. E. X. *Nat. Methods* **2018,** *15,* 645–646. doi:10.1038/s41592-018-0112-1
116. Buckingham, S. D.; Partridge, F. A.; Sattelle, D. B. *Int. J. Parasitol.: Drugs Drug Resist.* **2014,** *4,* 226–232. doi:10.1016/j.ijpddr.2014.10.004
117. Mathew, M. D.; Mathew, N. D.; Ebert, P. R. *PLoS One* **2012,** *7,* e33483. doi:10.1371/journal.pone.0033483
118. Puckering, T.; Thompson, J.; Sathyamurthy, S.; Sukumar, S.; Shapira, T.; Ebert, P. *F1000Research* **2019,** *6,* 192. doi:10.12688/f1000research.10767.3
119. Mathew, M. D.; Mathew, N. D.; Miller, A.; Simpson, M.; Au, V.; Garland, S.; Gestin, M.; Edgley, M. L.; Flibotte, S.; Balgi, A.; Chiang, J.; Giaever, G.; Dean, P.; Tung, A.; Roberge, M.; Roskelley, C.; Forge, T.; Nislow, C.; Moerman, D. *PLoS Neglected Trop. Dis.* **2016,** *10,* e0005058. doi:10.1371/journal.pntd.0005058
120. Stroustrup, N.; Ulmschneider, B. E.; Nash, Z. M.; López-Moyado, I. F.; Apfeld, J.; Fontana, W. *Nat. Methods* **2013,** *10,* 665–670. doi:10.1038/nmeth.2475
121. Stroustrup, N. Lifespan Machine Source Code. https://github.com/nstroustrup/lifespan (accessed Aug 22, 2019).
122. Marcellino, C.; Gut, J.; Lim, K. C.; Singh, R.; McKerrow, J.; Sakanari, J. *PLoS Neglected Trop. Dis.* **2012,** *6,* e1494. doi:10.1371/journal.pntd.0001494
123. Storey, B.; Marcellino, C.; Miller, M.; Maclean, M.; Mostafa, E.; Howell, S.; Sakanari, J.; Wolstenholme, A.; Kaplan, R. *Int. J. Parasitol.: Drugs Drug Resist.* **2014,** *4,* 233–243. doi:10.1016/j.ijpddr.2014.08.003
124. GitHub - chrismarcellino/wormassay: automated whole plate screening of macroscopic parasites. https://github.com/chrismarcellino/wormassay (accessed Aug 22, 2019).
125. Ritler, D.; Rufener, R.; Sager, H.; Bouvier, J.; Hemphill, A.; Lundström-Stadelmann, B. *PLoS Neglected Trop. Dis.* **2017,** *11,* e0005618. doi:10.1371/journal.pntd.0005618
126. Denecke, S.; Nowell, C. J.; Fournier-Level, A.; Perry, T.; Batterham, P. *PLoS One* **2015,** *10,* e0145051. doi:10.1371/journal.pone.0145051
127. Preston, S.; Jabbar, A.; Nowell, C.; Joachim, A.; Ruttkowski, B.; Cardno, T.; Hofmann, A.; Gasser, R. B. *Mol. Cell. Probes* **2016,** *30,* 13–17. doi:10.1016/j.mcp.2015.08.005
128. Preston, S.; Jabbar, A.; Nowell, C.; Joachim, A.; Ruttkowski, B.; Baell, J.; Cardno, T.; Korhonen, P. K.; Piedrafita, D.; Ansell, B. R. E.; Jex, A. R.; Hofmann, A.; Gasser, R. B. *Int. J. Parasitol.* **2015,** *45,* 333–343. doi:10.1016/j.ijpara.2015.01.007
129. Preston, S.; Jiao, Y.; Jabbar, A.; McGee, S. L.; Laleu, B.; Willis, P.; Wells, T. N. C.; Gasser, R. B. *Int. J. Parasitol.: Drugs Drug Resist.* **2016,** *6,* 329–334. doi:10.1016/j.ijpddr.2016.07.004
130. Preston, S.; Jiao, Y.; Baell, J. B.; Keiser, J.; Crawford, S.; Koehler, A. V.; Wang, T.; Simpson, M. M.; Kaplan, R. M.; Cowley, K. J.; Simpson, K. J.; Hofmann, A.; Jabbar, A.; Gasser, R. B. *Int. J. Parasitol.: Drugs Drug Resist.* **2017,** *7,* 286–294. doi:10.1016/j.ijpddr.2017.05.004
131. Le, T. G.; Kundu, A.; Ghoshal, A.; Nguyen, N. H.; Preston, S.; Jiao, Y.; Ruan, B.; Xue, L.; Huang, F.; Keiser, J.; Hofmann, A.; Chang, B. C. H.; Garcia-Bustos, J.; Jabbar, A.; Wells, T. N. C.; Palmer, M. J.; Gasser, R. B.; Baell, J. B. *J. Med. Chem.* **2018,** *61,* 10875–10894. doi:10.1021/acs.jmedchem.8b01544
132. Jiao, Y.; Preston, S.; Koehler, A. V.; Stroehlein, A. J.; Chang, B. C. H.; Simpson, K. J.; Cowley, K. J.; Palmer, M. J.; Laleu, B.; Wells, T. N. C.; Jabbar, A.; Gasser, R. B. *Parasites Vectors* **2017,** *10,* 323. doi:10.1186/s13071-017-2246-x
133. Dilrukshi Herath, H. M. P.; Preston, S.; Hofmann, A.; Davis, R. A.; Koehler, A. V.; Chang, B. C. H.; Jabbar, A.; Gasser, R. B. *Vet. Parasitol.* **2017,** *244,* 172–175. doi:10.1016/j.vetpar.2017.07.005

134. Nguyen, L. T.; Kurz, T.; Preston, S.; Brueckmann, H.; Lungerich, B.; Herath, H. M. P. D.; Koehler, A. V.; Wang, T.; Skálová, L.; Jabbar, A.; Gasser, R. B. *Parasites Vectors* **2019,** *12,* 191. doi:10.1186/s13071-019-3426-7
135. Partridge, F. A.; Brown, A. E.; Buckingham, S. D.; Willis, N. J.; Wynne, G. M.; Forman, R.; Else, K. J.; Morrison, A. A.; Matthews, J. B.; Russell, A. J.; Lomas, D. A.; Sattelle, D. B. *Int. J. Parasitol.: Drugs Drug Resist.* **2018,** *8,* 8–21. doi:10.1016/j.ijpddr.2017.11.004
136. Partridge, F. A.; Murphy, E. A.; Willis, N. J.; Bataille, C. J. R.; Forman, R.; Heyer-Chauhan, N.; Marinič, B.; Sowood, D. J. C.; Wynne, G. M.; Else, K. J.; Russell, A. J.; Sattelle, D. B. *PLoS Neglected Trop. Dis.* **2017,** *11,* e0005359. doi:10.1371/journal.pntd.0005359
137. Hurst, R. J.; Hopwood, T.; Gallagher, A. L.; Partridge, F. A.; Burgis, T.; Sattelle, D. B.; Else, K. J. *BMC Infect. Dis.* **2014,** *14,* 520. doi:10.1186/1471-2334-14-520
138. GitHub - fpartridge/invapp-paragon: Analysis of movies of worms to score motility or growth. https://github.com/fpartridge/invapp-paragon (accessed Aug 22, 2019).
139. McQuin, C.; Goodman, A.; Chernyshev, V.; Kamentsky, L.; Cimini, B. A.; Karhohs, K. W.; Doan, M.; Ding, L.; Rafelski, S. M.; Thirstrup, D.; Wiegraebe, W.; Singh, S.; Becker, T.; Caicedo, J. C.; Carpenter, A. E. *PLoS Biol.* **2018,** *16,* e2005970. doi:10.1371/journal.pbio.2005970
140. Neves, B. J.; Dantas, R. F.; Senger, M. R.; Melo-Filho, C. C.; Valente, W. C. G.; de Almeida, A. C. M.; Rezende-Neto, J. M.; Lima, E. F. C.; Paveley, R.; Furnham, N.; Muratov, E.; Kamentsky, L.; Carpenter, A. E.; Braga, R. C.; Silva-Junior, F. P.; Andrade, C. H. *J. Med. Chem.* **2016,** *59,* 7075–7088. doi:10.1021/acs.jmedchem.5b02038
141. Wählby, C.; Kamentsky, L.; Liu, Z. H.; Riklin-Raviv, T.; Conery, A. L.; O'Rourke, E. J.; Sokolnicki, K. L.; Visvikis, O.; Ljosa, V.; Irazoqui, J. E.; Golland, P.; Ruvkun, G.; Ausubel, F. M.; Carpenter, A. E. *Nat. Methods* **2012,** *9,* 714–716. doi:10.1038/nmeth.1984
142. GitHub - CellProfiler/CellProfiler: An open-source application for biological image analysis. https://github.com/CellProfiler/CellProfiler (accessed Aug 22, 2019).
143. CellProfiler | Free open-source software for measuring and analyzing cell images. https://cellprofiler.org/examples/published_pipelines.html (accessed Aug 22, 2019).
144. Bulman, C. A.; Bidlow, C. M.; Lustigman, S.; Cho-Ngwa, F.; Williams, D.; Rascón, A. A., Jr.; Tricoche, N.; Samje, M.; Bell, A.; Suzuki, B.; Lim, K. C.; Supakorndej, N.; Supakorndej, P.; Wolfe, A. R.; Knudsen, G. M.; Chen, S.; Wilson, C.; Ang, K.-H.; Arkin, M.; Gut, J.; Franklin, C.; Marcellino, C.; McKerrow, J. H.; Debnath, A.; Sakanari, J. A. *PLoS Negl. Trop. Dis.* **2015,** *9,* e0003534. doi:10.1371/journal.pntd.0003534
145. Rueden, C. T.; Schindelin, J.; Hiner, M. C.; DeZonia, B. E.; Walter, A. E.; Arena, E. T.; Eliceiri, K. W. *BMC Bioinf.* **2017,** *18,* 529. doi:10.1186/s12859-017-1934-z
146. Mackenzie, C. D.; Geary, T. G. *Expert Rev. Anti-Infect. Ther.* **2013,** *11,* 539–541. doi:10.1586/eri.13.49
147. Wolstenholme, A. J.; Maclean, M. J.; Coates, R.; McCoy, C. J.; Reaves, B. J. *Invertebr. Neurosci.* **2016,** *16,* 7. doi:10.1007/s10158-016-0190-7
148. Ramot, D.; Johnson, B. E.; Berry, T. L., Jr.; Carnell, L.; Goodman, M. B. *PLoS One* **2008,** *3,* e2208. doi:10.1371/journal.pone.0002208
149. Swierczek, N. A.; Giles, A. C.; Rankin, C. H.; Kerr, R. A. *Nat. Methods* **2011,** *8,* 592–598. doi:10.1038/nmeth.1625
150. Kwon, N.; Pyo, J.; Lee, S.-J.; Je, J. H. *PLoS One* **2013,** *8,* e57484. doi:10.1371/journal.pone.0057484
151. Itskovits, E.; Levine, A.; Cohen, E.; Zaslaver, A. *BMC Biol.* **2017,** *15,* 29. doi:10.1186/s12915-017-0363-9
152. Restif, C.; Ibáñez-Ventoso, C.; Vora, M. M.; Guo, S.; Metaxas, D.; Driscoll, M. *PLoS Comput. Biol.* **2014,** *10,* e1003702. doi:10.1371/journal.pcbi.1003702
153. Banse, S. A.; Blue, B. W.; Robinson, K. J.; Jarrett, C. M.; Phillips, P. C. *PLoS One* **2019,** *14,* e0216283. doi:10.1371/journal.pone.0216283
154. Figshare: S1_File.dwg. https://figshare.com/s/d3f77ad407d80d545f89 (accessed Aug 23, 2019).
155. Huang, H.; Singh, K.; Hart, A. *Bio-Protoc.* **2017,** *7,* e2174. doi:10.21769/bioprotoc.2174
156. Gürkan, G.; Gürkan, K. *IEEE Access* **2019,** *7,* 58764–58779. doi:10.1109/access.2019.2914958
157. GitHub – GurayGurkan/Incu-Stream: (Published) An Open-Hardware Live-Cell Imaging System Based on Inverted Bright-field Microscopy and Automated Mechanical Scanning for Real-Time and Long-Term Imaging of Microplates in Incubator. https://github.com/GurayGurkan/Incu-Stream (accessed Aug 23, 2019).
158. The Pathogen Box – Medicines for Malaria Venture. https://www.mmv.org/mmv-open/pathogen-box (accessed Aug 20, 2019).
159. Veale, C. G. L. *ChemMedChem* **2019,** *14,* 386–453. doi:10.1002/cmdc.201800755

160. Van Voorhis, W. C.; Adams, J. H.; Adelfio, R.; Ahyong, V.; Akabas, M. H.; Alano, P.; Alday, A.; Resto, Y. A.; Alsibaee, A.; Alzualde, A.; Andrews, K. T.; Avery, S. V.; Avery, V. M.; Ayong, L.; Baker, M.; Baker, S.; Mamoun, C. B.; Bhatia, S.; Bickle, Q.; Bounaadja, L.; Bowling, T.; Bosch, J.; Boucher, L. E.; Boyom, F. F.; Brea, J.; Brennan, M.; Burton, A.; Caffrey, C. R.; Camarda, G.; Carrasquilla, M.; Carter, D.; Cassera, M. B.; Cheng, K. C.-C.; Chindaudomsate, W.; Chubb, A.; Colon, B. L.; Colón-López, D. D.; Corbett, Y.; Crowther, G. J.; Cowan, N.; D'Alessandro, S.; Dang, N. L.; Delves, M.; DeRisi, J. L.; Du, A. Y.; Duffy, S.; El-Sayed, S. A. E.-S.; Ferdig, M. T.; Robledo, J. A. F.; Fidock, D. A.; Florent, I.; Fokou, P. V. T.; Galstian, A.; Gamo, F. J.; Gokool, S.; Gold, B.; Golub, T.; Goldgof, G. M.; Guha, R.; Guiguemde, W. A.; Gural, N.; Guy, R. K.; Hansen, M. A. E.; Hanson, K. K.; Hemphill, A.; van Huijsduijnen, R. H.; Horii, T.; Horrocks, P.; Hughes, T. B.; Huston, C.; Igarashi, I.; Ingram-Sieber, K.; Itoe, M. A.; Jadhav, A.; Jensen, A. N.; Jensen, L. T.; Jiang, R. H. Y.; Kaiser, A.; Keiser, J.; Ketas, T.; Kicka, S.; Kim, S.; Kirk, K.; Kumar, V. P.; Kyle, D. E.; Lafuente, M. J.; Landfear, S.; Lee, N.; Lee, S.; Lehane, A. M.; Li, F.; Little, D.; Liu, L.; Llinás, M.; Loza, M. I.; Lubar, A.; Lucantoni, L.; Lucet, I.; Maes, L.; Mancama, D.; Mansour, N. R.; March, S.; McGowan, S.; Vera, I. M.; Meister, S.; Mercer, L.; Mestres, J.; Mfopa, A. N.; Misra, R. N.; Moon, S.; Moore, J. P.; da Costa, F. M. R.; Müller, J.; Muriana, A.; Hewitt, S. N.; Nare, B.; Nathan, C.; Narraidoo, N.; Nawaratna, S.; Ojo, K. K.; Ortiz, D.; Panic, G.; Papadatos, G.; Parapini, S.; Patra, K.; Pham, N.; Prats, S.; Plouffe, D. M.; Poulsen, S.-A.; Pradhan, A.; Quevedo, C.; Quinn, R. J.; Rice, C. A.; Rizk, M. A.; Ruecker, A.; Onge, R. S.; Ferreira, R. S.; Samra, J.; Robinett, N. G.; Schlecht, U.; Schmitt, M.; Villela, F. S.; Silvestrini, F.; Sinden, R.; Smith, D. A.; Soldati, T.; Spitzmüller, A.; Stamm, S. M.; Sullivan, D. J.; Sullivan, W.; Suresh, S.; Suzuki, B. M.; Suzuki, Y.; Swamidass, S. J.; Taramelli, D.; Tchokouaha, L. R. Y.; Theron, A.; Thomas, D.; Tonissen, K. F.; Townson, S.; Tripathi, A. K.; Trofimov, V.; Udenze, K. O.; Ullah, I.; Vallieres, C.; Vigil, E.; Vinetz, J. M.; Vinh, P. V.; Vu, H.; Watanabe, N.; Weatherby, K.; White, P. M.; Wilks, A. F.; Winzeler, E. A.; Wojcik, E.; Wree, M.; Wu, W.; Yokoyama, N.; Zollo, P. H. A.; Abla, N.; Blasco, B.; Burrows, J.; Laleu, B.; Leroy, D.; Spangenberg, T.; Wells, T.; Willis, P. A. *PLoS Pathog.* **2016,** *12,* e1005763. doi:10.1371/journal.ppat.1005763
161. Pasche, V.; Laleu, B.; Keiser, J. *ACS Infect. Dis.* **2019,** *5,* 102–110. doi:10.1021/acsinfecdis.8b00220
162. Rufener, R.; Dick, L.; D'Ascoli, L.; Ritler, D.; Hizem, A.; Wells, T. N. C.; Hemphill, A.; Lundström-Stadelmann, B. *Int. J. Parasitol.: Drugs Drug Resist.* **2018,** *8,* 440–450. doi:10.1016/j.ijpddr.2018.10.011
163. Risi, G.; Aguilera, E.; Ladós, E.; Suárez, G.; Carrera, I.; Álvarez, G.; Salinas, G. *Vet. Sci.* **2019,** *6,* 29. doi:10.3390/vetsci6010029
164. Cintra, G. A. S.; Neto, B. A. D.; Carvalho, P. H. P. R.; Moraes, C. B.; Freitas-Junior, L. H. *SLAS Discovery* **2019,** *24,* 755–765. doi:10.1177/2472555219851130
165. Machicado, C.; Soto, M. P.; Timoteo, O.; Vaisberg, A.; Pajuelo, M.; Ortiz, P.; Marcos, L. A. *Antimicrob. Agents Chemother.* **2019,** *63,* e02373-18. doi:10.1128/aac.02373-18
166. Jiao, Y.; Preston, S.; Song, H.; Jabbar, A.; Liu, Y.; Baell, J.; Hofmann, A.; Hutchinson, D.; Wang, T.; Koehler, A. V.; Fisher, G. M.; Andrews, K. T.; Laleu, B.; Palmer, M. J.; Burrows, J. N.; Wells, T. N. C.; Wang, Q.; Gasser, R. B. *Parasites Vectors* **2017,** *10,* 272. doi:10.1186/s13071-017-2191-8
167. Song, H.; Liu, Y.; Xiong, L.; Li, Y.; Yang, N.; Wang, Q. *J. Agric. Food Chem.* **2013,** *61,* 8730–8736. doi:10.1021/jf402719z
168. Martin, R. J.; Robertson, A. P.; Bjorn, H. *Parasitology* **1997,** *114,* 111–124. doi:10.1017/s0031182097001029
169. Le, T. G.; Kundu, A.; Ghoshal, A.; Nguyen, N. H.; Preston, S.; Jiao, Y.; Ruan, B.; Xue, L.; Huang, F.; Keiser, J.; Hofmann, A.; Chang, B. C. H.; Garcia-Bustos, J.; Wells, T. N. C.; Palmer, M. J.; Jabbar, A.; Gasser, R. B.; Baell, J. B. *J. Med. Chem.* **2019,** *62,* 1036–1053. doi:10.1021/acs.jmedchem.8b01789
170. Le, T. G.; Kundu, A.; Ghoshal, A.; Nguyen, N. H.; Preston, S.; Jiao, Y.; Ruan, B.; Xue, L.; Huang, F.; Keiser, J.; Hofmann, A.; Chang, B. C. H.; Garcia-Bustos, J.; Wells, T. N. C.; Palmer, M. J.; Jabbar, A.; Gasser, R. B.; Baell, J. B. *J. Med. Chem.* **2019,** *62,* 3367–3380. doi:10.1021/acs.jmedchem.8b01790
171. Hotez, P. J.; Engels, D.; Fenwick, A.; Savioli, L. *Lancet* **2010,** *376,* 496–498. doi:10.1016/s0140-6736(10)60879-3
172. Schistosomiasis Control Initiative. https://schistosomiasiscontrolinitiative.org (accessed Jan 24, 2020).
173. Woelfle, M.; Seerden, J.-P.; de Gooijer, J.; Pouwer, K.; Olliaro, P.; Todd, M. H. *PLoS Neglected Trop. Dis.* **2011,** *5,* e1260. doi:10.1371/journal.pntd.0001260
174. Balasegaram, M.; Kolb, P.; McKew, J.; Menon, J.; Olliaro, P.; Sablinski, T.; Thomas, Z.; Todd, M. H.; Torreele, E.; Wilbanks, J. *PLoS Med.* **2017,** *14,* e1002276. doi:10.1371/journal.pmed.1002276

License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (https://www.beilstein-journals.org/bjoc)

The definitive version of this article is the electronic one which can be found at: doi:10.3762/bjoc.16.105