Drug Discovery Today • Volume xxx, Number xx • xxxx 2022 REVIEWS

KEYNOTE

Trends in oral small-molecule drug discovery and product development based on product launches before and after the Rule of Five

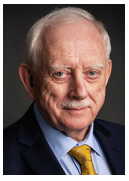
Sven Stegemann [1,*], Chris Moreton [3], Sami Svanbäck [4], Karl Box [5], Geneviève Motte [6], Amrit Paudel [1,2]

[1] Institute for Process and Particle Engineering, Graz University of Technology, Inffeldgasse 13, 8010 Graz, Austria
[2] Research Center Pharmaceutical Engineering GmbH, Inffeldgasse 13, 8010 Graz, Austria
[3] FinnBrit Consulting, Waltham, MA, USA
[4] The Solubility Company Ltd, Viikinkaari 4, 00790 Helsinki, Finland
[5] Pion Inc. (UK) Ltd, Forest Row, UK
[6] JEN Pharma Consulting, 182 Rue Henri Latour, 1450 Chastre, Belgium

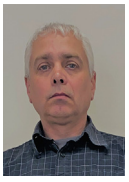
In 1997, the 'Rule of Five' (Ro5) suggested physicochemical limitations for orally administered drugs, based on the analysis of chemical libraries from the early 1990s. In this review, we report on the trends in oral drug product development by analyzing products launched between 1994 and 1997 and between 2013 and 2019. Our analysis confirmed that most new oral drugs are within the Ro5 descriptors; however, the number of new drug products of drugs with molecular weight (MW) and calculated partition coefficient (clogP) beyond the Ro5 has slightly increased. Analysis revealed that there is no single scientific or technological reason for this trend, but that it likely results from incremental advances are being made in molecular biology, target diversity, drug design, medicinal chemistry, predictive modeling, drug metabolism and pharmacokinetics, and drug delivery.

Keywords: chemical trends; Rule of Five; poor solubility; druggability; medicinal chemistry; oral drug delivery

Sven Stegemann Sven Stegemann is professor of patient-centric drug design and manufacturing at the Graz University of Technology and consultant to the pharmaceutical industry. Over the course of his 30-year career in the pharmaceutical industry, he has worked as an advisor to major pharmaceutical companies on ways to improve the formulation design, development, and manufacture of pharmaceutical products, including advanced drug delivery and manufacturing technologies and controls. In his academic role, he focuses his research on the rational development of patient-centric drug products and their associated manufacturing technologies, as well as education and training of students and young scientists.



Chris Moreton Chris Moreton is a partner at FinnBrit Consulting. Before joining FinnBrit Consulting, he spent ~35 years as a formulation scientist in large and small innovator companies, and in generic companies in the UK, Sweden, Canada, and USA, and also worked in quality assurance, quality control, and regulatory affairs for an excipient and drug delivery company. He is a visiting tutor at Manchester University (UK) on their PIAT program covering oral solid dosage forms. Chris holds a BPharm (Nottingham University, UK), an MSc in pharmaceutical analysis (University of Strathclyde, UK) and a PhD in pharmaceutics (Cardiff University, UK).



Karl Box Karl Box has been the Chief Scientific Officer (Europe) at Pion Inc. (UK) Ltd, since March 2020 and is involved in scientific and chemistry-related functions within the company, as well as supporting commercial activities and business development. His expertise is in the field of physicochemical measurements, where he has forged a career in the development of new instrumentation and assays for supporting drug discovery and development.

Geneviève Motte Geneviève Motte provides expertise and support to biotechs and pharmaceutical companies in the field of drug substance development, drug product development, CMC regulatory strategy and management of external contract development and manufacturing companies. She has over 30 years' experience ranging from drug design and preclinical to clinical operations and registration, with in-depth knowledge of small molecules and peptides, oral and injectable forms from early stage to commercial manufacturing.

* *Corresponding author:Stegemann, S. (sven.stegemann@tugraz.at)*

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

KEYNOTE

Introduction

Twenty-five years ago, the poor water solubility of small molecules referred to as new chemical entities (NCEs) as a result of the advent of combinatorial chemistry and high-throughput screening (HTS) procedures was recognized as a problem in drug development. The physicochemical descriptors were revealed by searching and analyzing chemical libraries of pharmaceutical companies. As a result, Lipinski *et al.* proposed the Ro5, which is based on limit values for MW ≤500 Da, clogP ≤5, hydrogen bond acceptors (HBAs) ≤10, and hydrogen bond donors (HBDs) ≤5. If not more than one parameter exceeds the Ro5 parameter, the compound can still be considered generally within the Ro5 [1]. The Ro5 not only brought the poor solubility of NCEs into scientific focus, but also sparked considerable debate among scientists regarding the proposed Ro5 and its valid limits. Regardless of the different scientific opinions and subsequently proposed rules, the problem of poor water solubility of NCEs was evident, and required appropriate responses from pharmaceutical scientific community. Here, we review the trends in oral drug product development according to the Ro5 to provide insights into the scientific advances as well as approaches the pharmaceutical companies have taken, as far as are publicly known, to solve the problem of poor water solubility and, hence, the bioavailability of NCE throughout the drug development process.

Amrit Paudel Amrit Paudel is an associate professor at the Graz University of Technology and is also a formulation team leader at the Research Center for Pharmaceutical Engineering (RCPE), Graz, Austria. His research exploits the design, engineering, and characterization of drug formulations and delivery systems intended for oral, inhalation, and implantable routes. The emphasis is given to a rational functionalization/manipulation of molecular, solid-state, surface, and particulate properties of active pharmaceutical ingredients, excipients, and formulations utilizing integrated physical, chemical, and engineering principles.

Materials and methods

Two data sets of drug product approvals containing a NCE by the US Food and Drug Administration (FDA) were generated. The data covered the years 2013–2019 and 1994–1997 as a reference data set. The new product launches involving a NCE (type 1 according to the submission classification) were retrieved from the FDA home page [2]. The NCEs were analyzed regarding their route of administration and the orally administered drug products formulated in a solid oral dosage form were selected for further analysis (Fig. 1).

The physicochemical characteristics were retrieved through the Drugbank Online database [3]. The data were analyzed according to the Ro5 listed characteristics and correlations were established between MW, clogP, HBD, and HBA to evaluate potential trending across all oral drug product launches including a NCE.

For the NCE with more than one physicochemical descriptors outside the Ro5, further analysis was performed on published information on the final drug product. Product-related information, such as European Public Assessment Reports, FDA Drug approval package, and prescription information, was retrieved

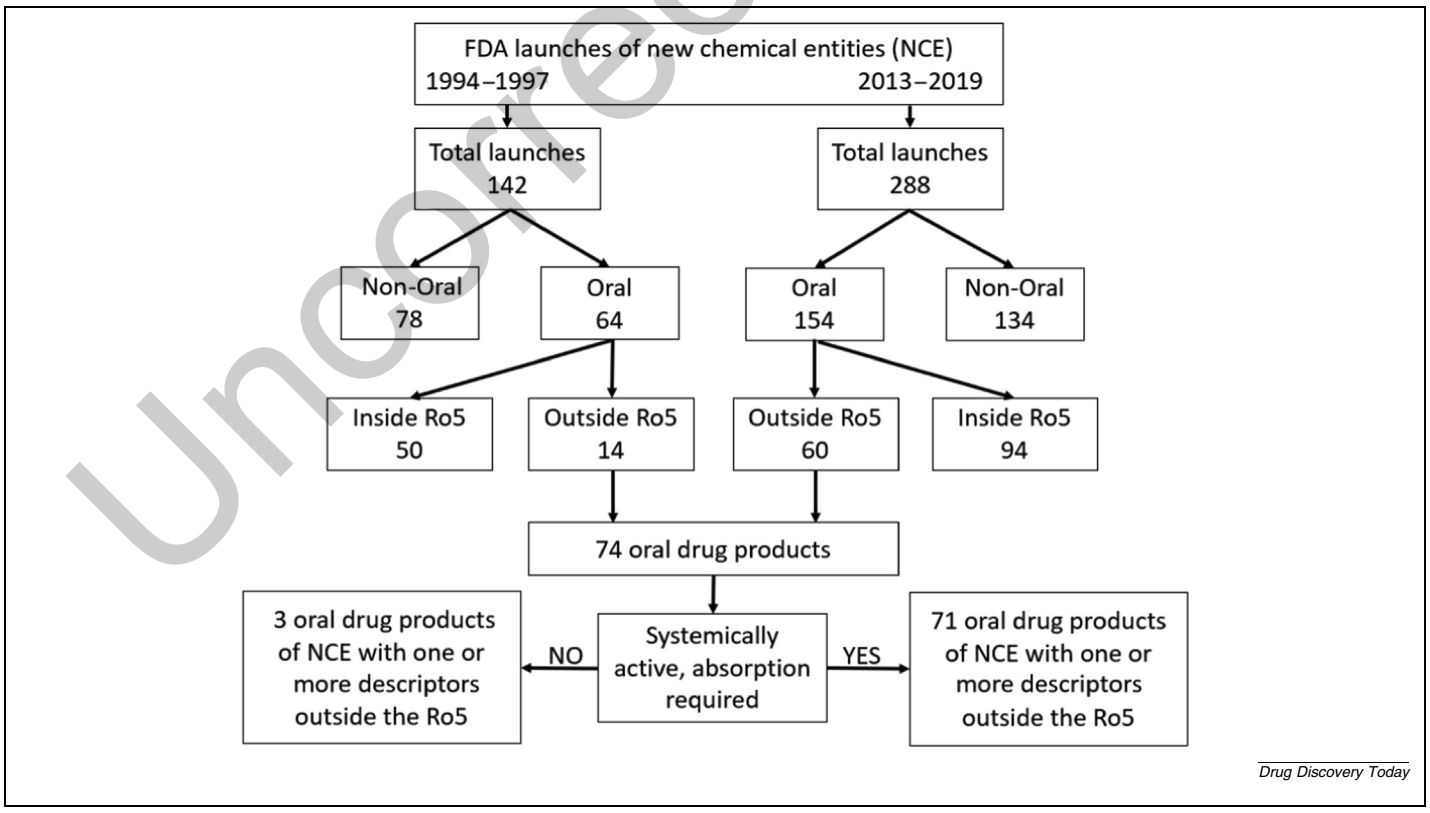


FIG. 1

New chemical entity (NCE) and product search strategy. Abbreviation: Ro5, Rule of 5.

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

TABLE 1

Orally administered drug products containing a NCE with one or more physicochemical descriptors outside the Ro5 launched between 1994 and 1997.[a]

Year	Trade name	API	Salt	MW	clogP	HBA	HBD	Highest formulated dose	BCS
1994	Prograf	Tacrolimus		804.0	3.19	11	3	5 mg	II
1995	Dynabac	Dirithromycin		835.1	2.90	15	5	250 mg	IV
1996	Norvir	Ritonavir		720.9	4.24	6	4	100 mg	IV
	Crixivan	Indinavir	Sulfate	613.8	3.26	7	4	400 mg	II
	Estrovis	Quinestrol		364.5	5.19	2	1	0.2 mg	?
	Allegra	Fenofexadine	Hydrochloride	501.6	2.49	5	3	180 mg	III
	Accolate	Zafirlukast		575.7	4.84	6	2	250 mg	II
	Elmiron	Pentosan-polysulfate		3836.0	–2.50	17	6	100 mg	III
	Stromectol	Ivermectin		1736.2	4.37	13	3	6 mg	II
	Lipitor	Atorvastatin	Calcium	558.6	4.24	5	4	80 mg	II
1997	Viracept	Nelfinavir	Mesylate	567.8	4.72	5	4	625 mg	IV
	Fareston	Toremifene	Citrate	405.9	6.27	2	0	60 mg	I
	Merida	Sibutramine	Hydrochloride	279.8	5.20	1	0	15 mg	I
	Evisla	Raloxifene	Hydrochloride	473.6	5.46	5	2	60 mg	II

[a] Descriptors exceeding the Ro5 are highlighted in gray.

from the European Medicines Agency (EMA) [4] and FDA websites [2]. The observed trends were analyzed and discussed within the broader context of the scientific and technological progress contributing to their development, such as in the area of predictive modeling, lead optimization, and solubility or bioavailability-enhancing drug delivery technologies, made since the publication of the Ro5 to address poor drug solubility and bioavailability.

Results

Trends in the chemical space of NCE launches between 1994 and 1997 and between 2013 and 2019

From 1994 to 1997, 64 NCEs delivered through the oral route were launched, of which 14 had one or more descriptors outside the Ro5 (21.9%). By comparison, from 2013 to 2019, 154 NCE delivered orally were identified, of which 60 had at least one physicochemical descriptor outside the Ro5 (39.9%) (Tables 1 and 2).

The comparison of the average MW, clog P, HBA, and HBD between the NCEs launched between 1994 and 1997, and those launched between 2013 and 2019, revealed an increase in MW and clogP, but not in HBA and HBD (Table 3). Moreover, the NCEs that breached at least one of the Ro5 descriptors increased from 21.9% to 39.9%.

Over the 4-year (1994–1997) and 7-year periods (2013–2019), the average chemical characteristics MW, clogP, HBD and HBA did not reveal a consistently monotonic trend, although they remained overall within the Ro5 space (Fig. 2). The increase in the average MW in 1996 can mainly be attributed to ivermectin (MW 1736.18) and the increase in 2016 to the low number of NCEs (ten), of which four exceeded the MW significantly [average MW 849.75 (range 766.9–882.0)].

Of the 154 NCEs launched between 2013 and 2019, 102 had a MW ≤500 Da (66.2%), 32 had a MW of 500–600 Da (20.7%), five had a MW of 600–700 Da (3.2%), six had a MW of 700–800 Da (3.9%) and seven were in the MW range of 800–900 Da (4.5%). Although no NCEs were identified in the MW 900–1100 Da range, two (pibrentasvir and tenapanor) were identified with a MW of 1100–1200 Da (1.3%).

An examination of the clogP values revealed that 19 NCEs had a clogP between 5 and 6 and four were within between 6 and 7. Seventeen of the NCE with a clogP >5 also had a MW >500 Da.

Of the 154 NCEs, four had HBA values abo>e 10, of which one compound had 11, two had 12 and one had 13 HBAs. Similarly, four NCEs had HBD values of 6 and were outside the suggested limit from the Ro5 of HBD ≤5.

Of the 154 NCEs, 60 (38.9%) had at least one characteristic outside the Ro5 limits. Of these 60 NCEs, 39 had one characteristic outside the limits (which, based on the Ro5 definition, can still be a candidate for oral delivery), whereas 19 NCEs had two and two NCEs had three characteristics outside the Ro5 limits. Higher MW accounted for 51 NCEs (33.3%) outside the Ro5, and 30 NCEs (19.5%) had clogP values outside the range.

The two NCEs identified with three characteristics outside the Ro5 limits were tenapanor and rifamycin. The 19 NCE identified with two characteristics outside the Ro5 limits were avatrombopag, brigatinib, ceritinib, dabrafenib, elbasvir, entrectinib, fostamatinib, ledipasvir, lusutrombopag, moxidectin, naloxegol, netupitant, omadacycline, ombitasvir, piprentasir, siponimod, telotristat ethyl velpatasvir, and venetolax.

From the NCE launches during the time frame of 1994–1997, 64 new products were launched, of which 14 (21.9%) violated at least one rule of the Ro5. Of these 14 products, ten NCEs exceeded one descriptor, three NCEs exceeded two (tacrolimus, dirithromycin, and ivermectin) and one NCE three (pentosanpolysulfate). Ten NCEs had a MW >500 (fenofexadine, atorvastatin, zafirlukast, pentosanpolysulfate, indinavir, ritonavir, tacrolimus, dirithromycin, ivermectin, and nelfinavir), of which three also had a HBA >10 (tacrolimus, dirithromycin, and ivermectin). Four NCEs had a clogP >5 (quinesterol, toremifene, sibutramine, and raloxifene).

Four of the NCEs exceeding Ro5 in MW had a MW of 500–600, one NCE between 600 and 700, one between 600 and 800, and two had a MW between 800 and 900. The other two NCEs had MW of 1736.18 (ivermectin) and 3836 (pentosanpolysulfate).

Three NCEs (pentosanpolysulfate, tenapanor, and rifamycin) were identified in the two data sets that exceeded the Ro5 in three characteristics: MW, HBA, and HBD. Even though all three

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

KEYNOTE

TABLE 2

Orally administered drug products containing a NCE with one or more physicochemical descriptors outside the Ro5 launched between 2013 and 2019.[a]

Year	Trade name	API	Salt	MW	clogP	HBA	HBD	Highest formulated dose	BCS
2013	Osphena	Ospemifene	–	378.9	5.56	2	1	60 mg	II
	Duavee	Bazedoxifene	Acetate	470.6	6.10	4	2	20 mg	I
	Tafinlar	Dabrafenib	Mesylate	519.6	5.46	6	2	75 mg	II
	Sovaldi	Sofosbuvir	–	529.5	1.28	6	3	400 mg	II
	Opsumit	Macitentan	–	588.3	3.69	9	2	10 mg	III
	Mekinist	Trametinib	Dimethylsulfoxide	615.4	3.18	5	2	2 mg	IV
	Olysio	Simeprevir	Sodium	749.9	4.56	9	2	150 mg	III
2014	Ofev	Nintedanib	Esylate	539.6	3.70	6	2	150 mg	I
	Zykadia	Ceritinib	–	558.1	5.23	8	3	150 mg	II
	Akynzeo	Netupitant	–	578.6	5.48	4	0	300 mg	IV
	Movantik	Naloxegol	Oxalate	651.8	1.73	12	2	25 mg	II
	Viekira	Paritaprevir	Dihydrate	765.9	3.50	10	3	75 mg	IV
	Viekira	Ombitasvir	Monohydrate	894.1	5.72	7	4	12.5 mg	IV
	Harvoni	Ledipasvir	–	889.0	5.98	6	4	90 mg	IV
2015	Rexulti	Brexpiprazole	–	433.6	5.38	4	1	4 mg	III
	Alecensa	Alectinib	Hydrochloride	482.6	5.59	5	1	150 mg	IV
	Odomzo	Sonidegib	Phosphate	485.5	5.64	5	1	200 mg	III
	Varubi	Rolapitant	Hydrochloride	500.5	4.07		2	90 mg	III
	Cotellic	Cobimetinib	Fumerate	531.3	3.35	4	3	20 mg	II
	Savaysa	Edoxaban	Tosylate	548.1	1.61	7	3	60 mg	IV
	Viberzi	Eluxadoline	-	569.7	1.08	7	4	100 mg	I
	Cresemba	Isavuconazonium	Sulfate	717.8	1.73	9	2	186 mg	I
	Daklinza	Daclatasvir	Dihydrochloride	738.9	4.67	6	4	90 mg	II
2016	Zepatir	Grazoprevir	Monohydrate	766.9	3.26	10	3	100 mg	I
	Venclexta	Venetoclax	–	868.5	6.76	10	3	100 mg	III
	Epclusa	Velpatasvir	–	882.0	5.93	8	4	100mg	IV
	Zepatir	Elbasvir	–	882.0	5.60	7	4	50 mg	II/IV
2017	Macrilen	Macimorelin	Acetate	474.6	1.77	4	6	60 mg	II
	Verzenio	Abemaciclib	Mesylate	506.6	4.25	7	1	200 mg	II
	Nerlynx	Neratinib	Maleate	557.1	4.72	8	2	40 mg	II
	Symproic	Naldemedine	Tosylate	570.6	3.14	8	4	0.2 mg	IV
	Rydapta	Midostaurin	–	570.6	4.52	4	1	25 mg	IV
	Prevymis	Letermovir	–	572.6	4.58	8	1	480 mg	II
	Xermelo	Telotristat ethyl	Hippurate	575.0	5.35	7	2	250 mg	IV
	Alunbrig	Brigatinib	–	584.1	5.11	9	2	180 mg	II
	Vosevi	Voxilaprevir	–	868.9	4.90	10	3	100 mg	II
	Mavyret	Pibrentasvir	–	1113.2	5.95	10	4	40 mg	IV
		Glecaprevir	–	838.9	4.26	10	3	100 mg	IV
2018	Epidiolex	Cannabidiol	–	314.5	6.10	2	2	20 mg	II
	Krintafel	Tafenoquine	Succinate	463.5	5.07	5	2	150 mg	I
	Symdeko	Tezacaftor	–	520.5	2.97	6	4	100 mg	II
	Braftovi	Encorafenib	–	540.0	4.16	7	3	75 mg	II
	Xospata	Gilteritinib	Fumarate	552.7	3.51	10	3	40 mg	II
	Nuzyra	Omadacycline	Tosylate	556.7	0.94	10	6	150 mg	I
	Xofluza	Baloxavir marboxil	–	571.6	2.12	8	0	40 mg	II/IV
	Tavalisse	Fostamatinib	Disodium	580.5	2.78	13	4	150 mg	II/IV
	Tibsovo	Ivosidenib	–	583.0	2.52	6	1	250 mg	IV
	Mulpleta	Lusutrombopag	–	591.5	6.04	6	2	3 mg	II/IV
	Orilissa	Elagolix	Sodium	631.6	4.68	6	2	200 mg	II
	Moxidectin	Moxidectin	–	639.8	5.30	8	2	2 mg	II
	Doptelet	Avatrombopag	Maleate	649.7	5.97	8	2	20 mg	II
	Aemcolo	Rifamycin	Sodium	719.7	4.15	11	6	194 mg	II
2019	Xenleta	Lefamulin	Acetate	507.7	3.72	5	3	600 mg	III
	Maysent	Siponimod	Fumarate	516.6	5.85	5	1	2 mg	II
	Trikafta	Tezacaftor	–	520.5	2.97	6	4	100 mg	II
	Inrebic	Fedratinib	Dihydrochloride	524.7	4.27	8	3	100 mg	II
	Ubrelvy	Ubrogepant	–	549.6	3.07	5	5	100 mg	IV
	Rozlytrek	Entrectinib	–	560.6	5.03	6	3	200 mg	II
	Trikafta	Elexacaftor	–	597.7	4.45	8	1	100 mg	IV
	Ibsrela	Tenapanor	Hydrochloride	1145.0	4.55	12	6	52 mg	IV

[a] Descriptors exceeding the Ro5 are highlighted in gray.

NCEs are orally delivered, they are either intended to treat intestinal targets without being absorbed (tenapanor and rifamycin) or the drug absorption and bioavailability <1% is sufficient for the therapeutic effect (pentosanpolysulfate). Taking this into account, they are consistent with the Ro5 because their bioavailability in negligible and blood concentrations remains below the detection limit. They were not included in the further analysis. Consequently, 58 NCEs were considered for further analysis of the oral space for the 2013–2019 data set and 13 for the 1994–1997 data set.

An analysis of NCEs exceeding two characteristics revealed that, between 2013 and 2019, 16 NCEs had a combination of MW >500 and clogP >5 (avatrombopag, brigatinib, ceritinib, dabrafenib, elbasvir, entrectinib, ledipasvir, lusutrombopag, moxidectin, netupitant, ombitasvir, pibrentasvir, siponimod, telotristat ethyl, velpatasvir, venetoclax); and two had higher

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

TABLE 3

Average values of MW, clogP, HBA, and HBD of drug products containing a NCE launched between 1994 and 1997 and between 2013 and 2019.

	1994–1997	2013–2019
MW	397.3	479.0
clog P	2.38	3.24
HBA	5.2	5.5
HBD	1.9	2.2
%bRo5	21.9% (*N* = 64)	39.9% (*N* = 154)

MW and HBA (fostamatinib and naloxegol). However, no NCE was identified in the 1994–1997 data set that had a MW >500 and clogP >5.

Breaching the Ro5 on MW and HBD was only observed for one NCE (omadacycline) in the 2013–1019 data set. By contrast, two NCEs out of the 154 approved between 2013–2019 were identified that had a MW >500 and HBA >10 (fostamatinib and naloxegol), compared with three from the 64 NCEs approved between 1994 –1997 (tacrolimus, dirithromycin, and ivermectin).

Analysis of the drug product launches exceeding two descriptors

NCEs exceeding MW and clogP

NCEs exceeding MW and clogP were only identified in the 2013–2019 data set. Brigatinib is a Biopharmaceutics Classification System (BCS) class I substance exhibiting high solubility and permeability. The product is a tablet formulation manufactured by direct compression. Netupitant is a new BCS class II compound developed in a tablet form and combined with the existing compound palonosetron formulated in a soft capsule filled into a single hard capsule. The tablet is manufactured by a wet granulation process using micronized netupitant to achieve consistent dissolution and a bioavailability of 60%. Avatrombopag maleate is practically insoluble in water in the pH range o1–9. A specific polymorph is maintained during the manufacturing by mixing, milling, mixing, dry granulation, sizing, lubrication, tableting, and coating. Velpatasvir is a BCS class IV compound with a pH-dependent solubility profile whereby it is soluble at pH 1.2 and practically insoluble at pH >5. In the drug product, velpatasvir is present in its amorphous form achieved by spray-drying technology. A dose of 100 mg velpatasvir is formulated as a fixed-dose combination product together with 400 mg sofosbuvir. Ledipasvir is a new compound launched as a fixed-dose combination with sofosbuvir. It is a BCS class II compound formulated as an amorphous spray-dried dispersion, with a pH-dependent dissolution and an oral bioavailability expected to be ≤30%. Siponimod fumarate is a BCS class II compound with good absorption and slightly increasing solubility at low pH or above pH 6.8. Despite the low dose, the compound achieved an estimated absolute bioavailability of 84%. The product is manufactured by a direct compression process. Pibrentasvir, which has a MW of 1113 Da, is formulated together with glecaprevir, another poorly soluble compound with a MW of 839 Da. Both drugs are individually formulated as amorphous solid dispersions (ASDs) containing vitamin E (tocopherol) polyethylene glycol (PEG) succinate (TPGS), and propylene glycol monocaprylate (PGMC II) in quantities that comply with the permitted upper intake level. The bioavailability of pibrentasvir was <10% in rodents, 29.8% in dogs, and 14.1% in monkeys, with a T_{max} of 3.7–9 h. Given that glecaprevir is an inhibitor of P-glycoprotein (P-gp) and breast cancer-resistant protein (BCRP) transporters, the co-administration increases the bioavailability of pibrentasvir approximately threefold.

Moxidectin (Moxidectin®) is a BCS class II compound administered in a dose of 4–8 mg. Given this low dose and its long half-life (20–43 days), the compound is sufficiently absorbed as a standard tablet formulation [5]. Lusutrombopag is practically insoluble in aqueous media below pH 9 and slightly soluble at higher pH. The product is manufactured in a nonstandard process of nine main steps: pre-mixing, screening, wet granulation, drying, screening and milling, blending with extra-granular excipients, lubrication, compression, and film-coating. The formulation contains magnesium oxide as a pH modifier. Entrectinib is poorly soluble compound with a pH-dependent dissolution. The formulation includes tartaric acid as a pH modifier and is manufactured by a dry granulation and encapsulation process. Dabrafenib is a BCS class II compound used in its micronized form in a HPMC capsule, whereby the HPMC supports dissolution (supersaturation) and reaches an absolute bioavailability of 95% [6]. Venetoclax is considered a BCS IV compound with unknown absolute oral bioavailability. One study suggested that the bioavailability was 5.4% in a solid-dosage form [7]. A solid dispersion using copovidone is used and includes polysorbate to increase the apparent aqueous solubility and bioavailability of venetoclax. Ombitasvir is a BCS class IV compound in a fixed-dose combination product containing 12.5 mg/75 mg/50 mg of ombitasvir, paritaprevir, and ritonavir respectively. All three components are converted into their amorphous form via hot melt extrusion. In addition, the formulation contains vitamin E TPGS, a known P-gp/cytochrome (CYP) P450 3A4 inhibitor. Telotristat ethyl is a BCS class IV compound that is rapidly absorbed and metabolized to its active form, telotristat. The product is formulated with HPMC and manufactured by roller compaction [8]. Elbasvir is considered a BCS class II compound. It is co-formulated at 50 mg dose with 100 mg of grazoprevir and vitamin E TPGS as a PGP/CYC 3A4 inhibitor. Both NCEs are converted and maintained in their amorphous form using spray drying. Ceritinib has good solubility in very acidic aqueous media and is absorbed by ≥25% based on the metabolites excreted. The product is formulated as an HPMC blend-filled gelatin capsule.

NCEs exceeding MW and HBA

Two NCEs, fostamatinib and naloxegol, had MW >500 and HBA >10. Fostamatinib is a 580-Da prodrug that is rapidly converted via enzymes in the gut to its active metabolite R406 (tamatinib) with a MW of 470 Da. Given the higher solubility of fostamatinib compared with tamatinib, the absolute bioavailability of the active moiety tamatinib increased to 55% but is highly variable (range 30–85%). The product is manufactured using a standard wet granulation process. An alkalizing diluent, sodium bicarbonate, was selected to modify local pH, produce effervescence, and increase the ionic concentration to interfere with

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

KEYNOTE

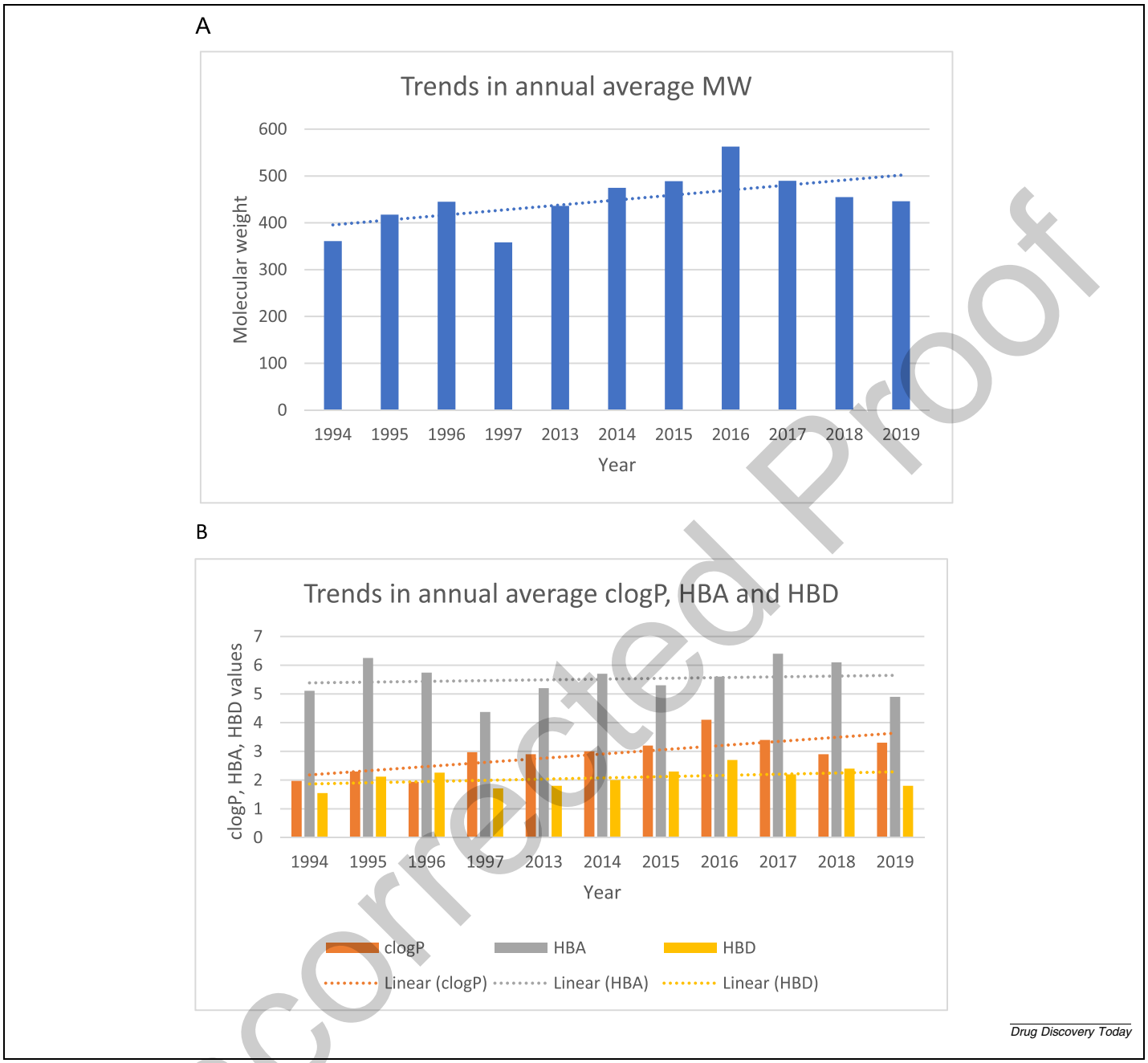


FIG. 2

Trends in the mean values and trends of **(a)** molecular weight (MW) and **(b)** calculated partition coefficient (clogP), H-bond acceptors (HBA), and H-bond donors (HDB) of US Food and Drug Administration (FDA)-approved oral drug products over the years 1994–1997 and 2013–2019.

drug–drug molecular interactions to prevent gel formation. Naloxegol is a pegylated naproxen with aqueous solubility between pH 1 and 7.5. The pegylation leads to low intestinal permeability (BCS class III), which is a desired characteristic of the drug to achieve slow drug absorption.

Three NCEs were identified from 1994 to 1997. Tacrolimus (MW 804 Da) is a low-dose BCS class II compound (0.5, 1.0 and 5 mg). It is absorbed along the gastrointestinal tract with an estimated bioavailability of 20–25% and high intrasubject variability and food effect. Ivermectin has a MW of 1736.2 Da and HBA 13 and is used in humans as an anthelmintic agent. It is derived by fermentation and contains at least 90% 5-*O*-demethyl-22,23-dihydroavermectin A_{1a} and <10% 5-*O*-demethyl-25-de(1-methylpropyl)-22,23-dihydro 25-(1- ethyl)avermectin A_{1a}. Ivermectin is a highly lipophilic compound with variable absorption and pharmacokinetics, although the latter are still not completely understood [9]. Dirithromycin was a lipid-soluble prodrug derivative of 9S-erythromycyclamine with a bioavailability calculated as 10% and manufactured as an enteric coated tablet to prevent drug degradation in gastric environment. However, the product was withdrawn from the market in 1997.

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

NCEs exceeding MW and HBD

From both data sets, only one NCE was identified with MW >500 and HBD >5. Omadacycline is absorbed by passive diffusion with a calculated absolute oral bioavailability of 34%. However, food, especially high-fat meals and the presence of divalent cations in the food, have a substantial negative effect on its bioavailability. The film-coated tablets are to be taken after fasting for 4 h and, after taking the tablet, no food or drink (except water) is to be consumed for 2 h and no dairy products, antacids, or multivitamins for 4 h. Therefore, omadacycline is also provided as an injectable form.

Comparative analysis of the two data sets

Comparison of the maximum oral dose delivered

From all 58 NCEs launched between 2013 and 2019 with at least one characteristic outside the Ro5, 53 NCEs (89.1%) were delivered in an oral dose of 200 mg or below (Table 4). Within the 58 NCEs, 20 had a MW of ≥600 Da, of which 16 are dosed at ≤100 mg and 4 at ≤200 mg. None of these 20 NCEs required a dose above 200 mg. There is no significant difference with regard to the dose between the NCEs outside the Ro5 and with the entire data set of 154 NCEs. Even though the number of new launched products between 1994 and 1997 only included 64 NCEs, of which 13 have at least one characteristic outside the Ro5, a higher percentage of NCEs tend to require a higher dose (>200 mg) compared with the 2013–2019 drug launches. In particular, three NCEs between 1994 and 1997 exceeding the MW required a dose between 250 mg and 400 mg and one compound required 625 mg. However, neither across all NCEs nor within the NCE with a descriptor outside the Ro5. Thus, there is no clear trend toward lower or higher oral doses over the two time periods.

Drug delivery technologies applied to products outside the Ro5 descriptors

From the products launched in the period 2013–2019, 17 products were identified that used bioavailability-enhancing formulations, with 14 using ASDs and three using lipid-based drug delivery systems. Of these 14 products, four were drugs within the Ro5 (suvorexant, ivacaftor apalutamide, and olaparib), five products exceeded the Ro5 in MW in the range of 500–840 Da (elagolix, elexacaftor, tezacaftor, encorafenib, and glecaprevir) and six products violated the Ro5 with a MW >750 Da and another rule (logP, HBA, or HBD) (paritaprevir, ombitasvir, grazoprevir, venetolax, velpatasvir, ledipasvir, and pibrentasvir). Three products used a lipid-based drug delivery approach (nintedanib, midostaurin, and tafamidis). Midostaurin (MW 570.6 Da) is a formulated ion mixture of lipophilic and hydrophilic solvents with a surfactant forming a microemulsion in water. Nintedanib (MW 539.6 Da) is formulated as a suspension in a lipophilic matrix. Tafamidis is a product within the Ro5 and formulated as a suspension in a mixture of PEG 400, sorbitan monooleate, and polysorbate 80.

From the products launched period from 1994 to 1997, two could be identified that apply bioavailability-enhancing drug delivery technologies. Tacrolimus is manufactured by a wet granulation process to achieve a defined solid microdispersion. Ritonavir is formulated as a semi-solid amorphous dispersion comprising a mixture of caprylic/capric medium-chain triglycerides, polyoxyl 35 castor oil, citric acid, ethanol, polyglycolized glycerides, polysorbate 80, and propylene glycol [10]. The product was discontinued in 1998 because of the formation of a new polymorph during its shelf-life. The product was reformulated and marketed as a soft gelatin capsule (butylated hydroxytoluene, ethanol, oleic acid, and polyoxyl 35 castor oil) and, from 2010 onwards, as a 100-mg ASD tablet formulation manufactured by HME (copovidone, sorbitan laurate, anhydrous calcium hydrogen phosphate, colloidal anhydrous silica, and sodium stearyl fumarate).

Troglitazone (Rezulin®) is a solid dispersion based on polyvinylpyrrolidone (PVP) and PEG, in which the drug is maintained in amorphous form [11]. Zafirlukast is a poorly water-soluble compound used in its amorphous form. The tablet formulation contains hydroxypropyl methylcellulose (HPMC) and PVP as a crystallization inhibitor to maintain supersaturation of zafirlukast in aqueous media [12].

Discussion

Poor aqueous solubility and bioavailability were identified as accounting for ~40% of drug compound attrition during drug development in 1991, which had reduced to 10% by 2000 [13] and declined further to 4% in 2010 [14]. To address this issue,

TABLE 4

Trends in the maximum dose delivered of the 58 and 13 drug products containing an NCE launched in 2013–2019 and 1994–1997, respectively exceeding at least one characteristic of the Ro5 compared with the entire data set.[a]

Maximum oral dose delivered	Number (%) of NCE with at least one characteristic outside the Ro5 versus all launches			
	2013–2019 (*N* = 58)	2013–2019 (*N* = 154)	1994–1997 (*N* = 13)	1994–1997 (*N* = 64)
<1 mg	1 (1.7)	3 (0.6)	1 (7.7)	3 (4.7)
1–10 mg	6 (10.4)	20 (13.0)	2 (15.35)	13 (20.4)
10–50 mg	12 (20.7)	37 (24.2)	1 (7.7)	9 (14.0)
50–100 mg	20 (34.5)	39 (25.5)	4 (30.8)	9 (14.0)
100–150 mg	7 (12.0)	16 (9.8)	0	2 (3.1)
150–200 mg	7 (12.0)	17 (11.1)	1 (7.7)	9 (14.0)
200–300 mg	2 (3.5)	12 (7.8)	2 (15.35)	8 (12.5)
300–600 mg	3 (5.2)	8 (5.2)	1 (7.7)	7 (11.0)
>600 mg	0	1 (0.6)	1 (7.7)	4 (6.3)

[a] Total NCE: 2013–2019 = 5153; 1994–1997 = 564.

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

KEYNOTE

the Ro5 was developed from a chemical perspective based on the physicochemical properties of 2245 NCEs from the United States Adopted Names (USAN) data base that had entered Phase II clinical trials. The Ro5 limits were based on the 90th percentile prediction, which considers that 90% of NCEs would fall into this class. According to Lipinski´s paper from 2001 [15], 90% of NCEs fall within the Ro5, whereas Brown and Boström found 17% of NCEs were outside the Ro5 [16]. In their original paper, Lipinski *et al.* [1] stated that some drugs will lie outside the set rule limits because the Ro5 is derived from the distribution of calculated properties within a data set of several thousand drugs considering that a certain subset of chemical NCEs outside the Ro5 might have structural features that allow the drugs to act as substrates for naturally occurring transporters, such as antibiotics, antifungals, glycosides, or vitamins among others.

To review the validity of the Ro5 25 years after its publication, NCEs of FDA new drug approvals between 2013 and 2019 were evaluated according to the descriptors of the Ro5 and their trends in the chemical space compared with FDA drug approvals between 1994 and 1997. The analysis of the data and trends obtained from these two data sets led to a review of the scientific and technological advancements to address the solubility and bioavailability issues in drug discovery and development over the past 25 years.

The two data sets revealed an increase in MW and clogP of NCEs launched between 1994 and 1997 compared with 2013–2019. The descriptors HBA and HBD did not show any changes between the two data sets. Our analysis of orally delivered NCEs revealed an increase in average MW from 397.3 to 479.0 Da and in clogP from 2.38 to 3.24 from 1994–1996 to 2013–2019. These findings were in line with a previous analysis showing that NCEs exceeded the Ro5 mainly as a result of MW, attributed to chemical optimization [17]. Another more in-depth analysis compared the MW, clogP, number of HBDs, number of HBAs, topological polar surface area (TPSA), rotatable bonds (NROTs), fraction of sp3 carbon atoms (Fsp3), and number of aromatic rings (#ArRNG) of NCEs launched between 1900 and 1997 with those launched between 1998 and 2007 and between 2008 and 2017, revealing an increase in the average MW to ~600 Da and also an increase in the average clogP to ~6 [18].

The analysis revealed that the MW and clogP descriptors continue to represent a frontier, whereby the drug products of NCEs with descriptors outside the Ro5 suggested that these were derived from chemical and pharmaceutical optimization. Given that no single scientific approach or technology was apparent for the increase in NCEs and their products with a MW >500 and clogP >5, the data suggest a continuous evolution in the different areas of drug discovery and formulation technologies.

Drug discovery platforms

Since the beginning of the 21st century, different drug discovery platforms have emerged to improve lead generation. Besides HTS, these major drug discovery platforms include: fragment-based lead generation (FBLG); structure-based drug design (SBDD); utilization of known literature, such as fast-follower or knowledge-based programs; and, more recently, DNA-encoded library screening (DEL) [16]; combinations thereof to overcome some of their inherent limitations [19]; or the integration of a series of different *in silico* tools to support the follow-on lead discovery and optimization [20]. For example, phenotypic approaches provide advantages over target-based approaches because they are not solely based on the receptor target [21]. Further improvements in drug screening are being achieved by 3D-engineered cell constructs forming organoid structures. Such *in vitro* systems are superior to traditionally used 2D cell-based systems [22,23]. For HTS, microfluidic-based 2D and 3D screening platforms have been developed, which also provide additional advantages, such as improved simulation of the environment, fluid control, and maintenance of cell morphology [24]. Additionally, the degree of automation in drug discovery, lead generation and synthesis, as well as testing, is increasing, which is expected to considerably reduce the number of lead compounds required by an integrated approach, resulting in an unbiased and rational approach and adaptive molecular design [25]. An open-source tool for structure-based virtual screening (SBVS) has been introduced that combines large compound databases with ultra-large ligand library screening. The open-source platform was validated using the Kelch-like ECH-associated protein 1 (KEAP1) as a target [26]. Consequently, the drug discovery and lead identification processes have been expanded by additional screening tools to improve drug design, drug-like properties, and lead candidate selection.

Physicochemical descriptors

Although the resulting lead NCEs from the various drug discovery platforms have not changed in terms of their general structural properties [16], the Ro5 led to increasing attention being paid to lead optimization to account for the physicochemical properties of the molecules for therapeutic application. The focus was on leveraging technological advances to better design NCEs in the druggability or drug-like space. This included more emphasize on additional chemical characteristics required to achieve the desired pharmacokinetic, pharmacodynamic, safety, and quality requirements to be approved as a pharmaceutical drug product. Veber *et al.* evaluated more than 1100 early drug compounds from Glaxo SmithKline, and suggested that MW can be further specified by the NROT and the TPSA of a new compound. Ten or fewer ROT and TPSA <140 $Å^2$ appear to achieve good bioavailability in rat animal models [27]. Moreover, additional chemical characteristics and parameters were validated for their predictive nature [e.g., Chrom $logD_{7.5}$,Property Forecast Index (PFI), calculated molar refractivity (cmr), etc] [18,28]. For example, for oral drug candidates, the distribution coefficient at pH 7.4 ($log D_{7.4}$) is another valuable descriptor of lipophilicity at physiological pH, which appears to be favorable in the range ~1–3 [29]. Chrom logD can be measured at different pH by reverse phase HPLC [30]. PFI is a combination of Chrom logD7,4 and the aromatic ring count. Permeability has been successfully predicted by correlating lipophilicity determined by $clogD_{7.4}$ versus size calculated by cmr [31]. These include numerous surrogate assays that are predictive of *in vivo* absorption, distribution, metabolism, excretion, and toxicity (ADMET). Sutherland *et al.* [32] found that surrogate pairs are most predictive, such as rat primary hepatocyte (RPH) cytolethality and volume of distribution (Vd) for *in vivo* toxicology outcomes, scaled microsome metabolism and clogP for *in vivo* unbound clearance,

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

and calculated logD and kinetic aqueous solubility for thermodynamic solubility. Such predictive surrogates can be used to predict the influence of certain chemical substituents on ADMET drug properties *in vivo*, allowing rational lead optimization [32]. With the increasing digital possibilities, the use of large data sets from companies or chemical libraries provides a deeper understanding of physicochemical characteristics of ADMET. For example, Waring investigated the *in vitro* permeability of 9571 NCEs with regards to logD, clogP, PSA, HBD, HBA, MW and ROTB. The results showed a linear relationship between the descriptors and permeability, providing good estimates (50% probability) [33]. Using logD and MW to predict permeability and metabolism/clearance (human liver microsomes) based on a data set of 47 018 preclinical NCEs provided further guidance for early drug discovery as well as lead optimization [34]. Another review found that a multiparametric scoring function (AB-MPS) based on ΔlogD, number of aromatic rings (NAR), and ROTB could be correlated with oral bioavailability in rats [35].

There is growing evidence that, for some molecular structures, their physicochemical characteristics depend on the environment they are exposed to. Such types of chameleonic molecule have a good balance between solubility and permeability despite a larger MW. For example, the difference between TPSA and molecular PSA calculated on a conformation of minimum energy in a low dielectric medium (MPSAnp) provided information on the environmental impact of modulating the polarity of a NCE [36]. The chameleonic properties are driven by intramolecular hydrogen bond interactions and intramolecular effects on steric conformation in different environments favoring either solubility or permeability [37,38]. Ermondi *et al.* suggested Δlog $P_{oct\text{-}tol}$ and chameleonicity (ChameLogD) as additional physicochemical descriptors of permeability [36]. These newly introduced physicochemical descriptors provide an additional set of tools for lead candidate selection, including those with descriptors exceeding the Ro5 limits.

Predictive modeling to predict solubility and pharmacokinetics

Predictive modeling, especially structure-based (aqueous) solubility prediction, of drug-like molecules is a Holy Grail and has been fertile area of research over the past two decades. Predictive models range from empirical thermodynamic models, equation of states, statistical and machine-learning models to high-fidelity molecular dynamics simulations [39]. Well-established models include Yalkowsky's classical General Solubility Equation (GSE) [40], ABSOLV by Abraham [41], and Breiman's Random Forest Regression (RFR) models [42]. In the context of structural parameters for solubility prediction, it was recently shown that a model combining the molecular flexibility parameter (NROTB) and HBAs (a Ro5 parameter) can generate the best solubility prediction for a wide range of drug-like molecules, even outperforming RFR-based models [43]. Such a model was used to predict the solubility of drug molecules approved between 2016 and2020 [44]. This highlighted the need to further rationalize the parameters that are included in, and associated with, the Ro5 for drug developability prediction. Similarly, predictive models have been developed to predict pharmacokinetics and ADMET and are applied in drug discovery and lead optimization [45,46].

Lead optimization

Large data sets of clinical and marketed NCE have been used to correlate different single chemical structures or substitutions with the *in vivo* behavior of a molecule. Young *et al.* [29] provided an excellent summary of such lead optimization programs from hit to drug candidate. The examples show the trajectory of the same pharmacophore hit through the chemical space of optimization by different companies, such as with regard to the clogP and heavy atom counts on the predicted 50% inhibitory concentration (pIC_{50}), ligand efficiency (LE), and lipophilic ligand efficiency (LEE). In addition, increasing evidence for crucial chemical structures (e.g., benzylic C-H bond, the allylic methyl, and *O*-, *N*-, and *S*-methyl groups) as substrates of cytochrome P450-mediated oxidative and reductive metabolism reactions is being addressed by the introduction of carbohydrates or fluorine on the aromatic moiety or deuterium to hinder metabolism. Recent trends observed in the chemical design of NCEs launched between 2015 and June 2020 revealed the increasing use of nitrogen heterocycles as pharmacophores and the rational introduction of nitro groups to reduce cytotoxicity and of boron to enhance reactivity toward nucleophiles of enzymes [47]. In 2014, an analysis was performed of all known NCEs in clinical trials and on the market to determine those that were outside the Ro5 to review their physicochemical properties and *in vitro* cell permeability [48]. The analysis found that the majority of orally administered drug NCEs with a MW >700 Da were natural-like NCEs belonging to the class of macrolide antibiotics or antiviral NCEs, such as for HIV or hepatitis C virus (HCV). Such chemical structures mimicking endogenous structures could be considered a specific group of chemical NCEs. In addition, certain structures, such as natural products (e.g., macrolides), peptidomimetic structures (e.g., antiviral NCEs), or an increased proportion of saturated carbons and number of chiral centers make the molecules less flat compared with typical *de novo* synthesized NCEs, and increase their solubility characteristics.

The importance of optimization of the beyond Ro5 space constituting macrocyclic and flexible molecules for oral drugs was recently re-emphasized [49]. Intramolecular hydrogen bond formation and *N*-methylation of amide bonds are being used as structural features to improve cell permeability and oral bioavailability. By contrast, increases in intramolecular hydrogen bond propensity generate a 'molecular chameleon' with erratic aqueous solubility behavior [50]. Such molecules have a negative enthalpy of solution and form crystal structures with large voids that can accommodate a significant amount of water when in contact with the liquid medium.

Lead optimization also includes an increased understanding of the target–receptor interaction and receptor affinity. Computational and predictive models have gained increasing importance and validity over the past 20 years. Quantitative structure–activity relationship (QSAR) models have evolved as valid tools to increase our knowledge and optimization of NCEs with regard to compound affinity and toxicology [51]. QSAR methods have been developed further by various dimensional QSAR approaches and their combination with chemo-

KEYNOTE

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

KEYNOTE

metric methods [52]. Additional computational tools are available based on crystallographic techniques and computer-aided molecular modeling to optimize the ligand–receptor interaction [53]. Improved predictive models for *in vivo* performance suitable for HTS are considered using the binding/unbinding kinetics of the protein–ligand interaction [54]. This has led to important understanding of the compound characteristics of NCEs that exceed Ro5 limits. Structural analysis of the receptors targeted by NCEs beyond the Ro5 have one of two distinct hot spot structures ('complex' and 'simple'), which both differ from the hot spot structures targeted by marketed drugs within the Ro5 [55]. The optimization of the receptor interaction has led to an increasing number of NCEs outside the Ro5 that are effective at oral doses of ≤50 mg [48]. For such beyond Ro5 NCEs, stereochemistry is also being regarded as the means to improve not only drug potency and selectivity, but also its solubility [56].

The evolution of beyond Ro5 molecules appears to be connected to the advancement and finding of specific drug targets. For example, a potent drug molecule targeted to a neurotransmitter and lipid metabolism pathway or to the target accessible via lipid transport is expected to be highly lipophilic [57]. Protein kinase inhibitors are extremely liphophilic [58,59], and account for most drugs using bio-enabling formulations, such as ASDs, between 2013 and 2019. Protease inhibitors and other peptidomimetic small-molecule drugs are large, with several macrocycles present, as an indication of violating multiple Ro5. Likewise, this is expected of drug candidates such as proteolysis-targeting chimeras (PROTACs), because their structures contain lengthy linkers and flexible bonds [60]. Lead optimization has continuously intensified over the past decades and, combined with predictive models, is a source for rapid oral drug development [61].

Salt screening and selection

An important aspect not thoroughly analyzed and presented herein is the landscape of salt forms of drug molecules. Free acid or base molecules with ionizable groups can be made into salts using a polar/hydrophilic counter-ion to enhance aqueous solubility and, in other instances, to improve stability and manufacturability [62]. In this context, around one-third of FDA-approved drugs are clinically used as their salt forms [63], with nearly 50 different types of counter-ion [64]. Among the drugs approved by FDA between 2015 and 2019, 61 were salt forms [65]. Although salt forms have proven to be important for augmenting biopharmaceutical, stability, and processability-related attributes of drugs, they also have inherent risk of disproportionation [66]. In this context, it would be interesting to investigate the trend in the presentation of approved beyond Ro5 drugs as salt forms during and after the advent of Ro5.

Medicinal chemistry and synthesis

Over the past two decades, advances in medicinal chemistry have contributed to the drug discovery and lead optimization process. Advances in chemical synthesis, such as asymmetric organocatalysis [67] and continuous flow technology [68], enable more sophisticated synthesis to produce more complex molecules for target screening compared with traditional combinatorial chemistry. In addition, new chemical synthesis tools and methods are being applied to facilitate the synthesis of more complex and natural-like compounds as well as the incorporation of chemical substituents into molecules that are able to increase metabolic stability, permeability, and absorption [69,70]. In particular, beyond Ro5 NCEs are based on natural compounds and their analogs, which are required for increasingly difficult clinical targets [71]. One of the major classes of NCE launched in this space over the past two decades are kinase inhibitors, with 61 launched between 2001 and 2021 [72].

Drug delivery technologies

Over the past two decades, bioavailability-enhancing drug delivery technologies have been increasingly applied to drug molecules, especially those outside the Ro5. The most commercially applied drug delivery technologies to enhance bioavailability are particle size reduction and conversion of the drug into its amorphous form. Both approaches were developed during the mid-1990s and became commercially viable at the start of the 21st century [73,74]. This is in accordance with the product launches between 1994 and 1996, where only one product was developed used a semi-solid amorphous dispersion approach (ritonavir), although it had to be withdrawn 2 years later because of the formation and precipitation of a new polymorph altering its dissolution [75]. Through intensive basic research focusing on the formulation principles of ASDs and on spray drying and hot-melt extrusion (HME) approaches to produce them, there is now both a mechanistic and predictive basis for derisking the development of such products. The application of analytical technology over the past few decades has enabled more in-depth characterization of solid dispersions. These also provided further insights into drug–polymer interactions and the physical state properties of the components, their mixtures, and stability [76]. The development of predictive models and process simulation tools additionally supports the development and process control [77,78]. From this increasing knowledge, the manufacturing processes and equipment could be further developed, so that a wide variety of processing and machine geometries are now available for product optimization [79–81]. For example, Norvir® was reformulated into an ASD by HME in 2010, similar to its follow-on product Kaletra® (lopinavir/ritonavir), which was introduced to the market in 2000 as a soft gelatin capsule and later redeveloped as a HME tablet formulation approved in 2005 by the FDA.

Within the course of advancing drug delivery technologies, novel and advanced excipients have reached the market, tailored to specific application. For example, HPMC-AS was introduced as an enteric coating polymer ~30 years ago and was further developed as a polymer for ASD. It displays good solubilization and miscibility with many drug NCEs for spray drying and has been applied successfully in pharmaceutical drug products [82,83].

Despite intensive research on lipid-based and liquid [self-emulsifying drug delivery systems (SEDDS), self-microemulsifying DDS (SMEDDS)] DDSs [84,85], they have not found their way into the commercial development of new products. Of the three approved products, only two have applications for molecules outside of Ro5 and, in both cases, the MW is only slightly exceeded (539.6 Da and 570.1 Da, respectively). Accord-

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

ing to Savla *et al.*, five drug molecules (sirolimus, saquinavir, tipranavir, dutasteride, and cyclosporine) violated the Ro5, with all of them included in FDA-approved lipid-based formulations up until 2016 [86]. Of these, all violated MW and clogP, and two additionally violated the HBA criterion.

The advancement of computational and experimental tools applicable to drug design, formulation, and processing over past decades has led to an encouraging clinical success in several beyond Ro5 molecules. Nevertheless, there are several untapped opportunities for molecular structure-based decision making in terms of early formulation and drug delivery design for beyond Ro5 NCEs. For example, several parameters that can be decisive for the amorphization tendency of beyond Ro5 molecules, the lattice energy of their crystals, and the associated hydrophobicity, can also be directly or indirectly linked to Ro5 parameters. For example, the role of active pharmaceutical ingredient lipophilicity in its ASD manufacturability is seldom investigated. The classification by Friesen *et al.* includes the ratio of melting point (T_m) to glass transition temperature (T_g) versus logP to categorize the amorphous formation propensity [83]. For molecules with T_m/T_g <1.25, amorphous formulation can form irrespective of logP; by contrast, molecules with logP >6 are problematic in forming amorphous formulation irrespective of T_m/T_g. Furthermore, a MW cutoff for successful amorphization is proposed to be 300 g/mol [86]. In addition, the relative trend in amorphization method (melt versus solvent based) is argued in terms of MW. Likewise, several molecular and material descriptors of beyond Ro5 molecules have been compared and contrasted for their ability to predict the lipid-based formulation success for these molecules. Of these, drug solubility in lipid excipient media is crucial and molecules with T_m <423 K as well as clogP >4 are proposed to show reasonable solubility in glycerides. By contrast, Salva *et al.* suggested that experimental drug solubility in octanol is a better predictor of lipid formulation success compared with calculated clogP [86]. Beyond this, recently developed approaches, such as structural mimic-based poorly soluble molecular solid optimization as well as various QSAR, molecular dynamics and machine learning-based predictions and optimizations of beyond Ro5 drug formulations, are promising [57,87,88]. While these are still early days, it is noteworthy that the perception of a rational and molecular level understanding of solubility limitations is growing in academic and industrial areas. It is evident that such understanding can expand the applicability of Ro5 from candidate selection to science-based developability assessment and earlier decisions as to formulation approaches and process selection.

Concluding remarks

The original 1997 paper by Lipinski *et al.* addressed the poor aqueous solubility of NCEs as a major challenge for drug product development. In response to the publication, the pharmaceutical industry recognized the importance of placing more and earlier emphasis on the chemical optimization, pharmacokinetic, and druggability characteristics in candidate selection and lead optimization programs to enhance the drug-like properties. This triggered efforts to better integrate the different disciplines and their expertise to improve the drug candidate selection and product development process by interdisciplinary collaboration [89–91].

The trends in NCEs approved by the FDA between 1994 and 1997 and between 2013 and 2019 revealed that, although the Ro5 descriptors continue to have a certain relevance, the percentage of new drug products with a MW and clog P beyond the Ro5 is increasing. While excellent work has appeared on trends in medicinal chemistry, less is known about advances in other areas of pharmaceutical science that have contributed to the successful development of products outside of Ro5. Over the past two decades, our knowledge of pharmacology, molecular biology, medicinal chemistry, formulation science, data science, and predictive approaches has evolved tremendously. Closer collaboration between stakeholders in the drug candidate selection, lead optimization, and drug product development process increases the probability of success. The review of the drug products exceeding at least two descriptors of the Ro5 (beyond the Ro5) indicates that the drug candidates passed through intensive chemical optimization using emerging descriptors and predictive models as well as more sophisticated chemical synthesis especially of flexible and chameleonic compounds. Sufficient bioavailability is being achieved by combined drug delivery technologies, such as stable amorphous drug delivery formulation approaches. Although the Ro5 is based on four state-of-the-art descriptors developed during the mid-1990s, they still represent crucial chemical descriptors. However, from today's perspective, a variety of tools are available and continue to emerging to develop an increasing number of drug products in the chemical space beyond the Ro5. Consequently, the issue of poor aqueous solubility and bioavailability has increasingly become an interdisciplinary task that has gradually expanded the boundaries of oral product development. This trend analysis and other published work can yield further useful guidance for developing a holistic approach from drug discovery to a drug product with a desired bioavailability and could eventually trigger further interdisciplinary discussion and collaboration.

Declaration of interests

The authors declare that they do not have any conflict of interest.

Data availability

No data was used for the research described in the article.

References

[1] Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23, 3–25.

[2] FDA (2022). Drugs@FDA: FDA-Approved Drugs. www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event5reportsSearch.process [Accessed October 24, 2022].

[3] Drugbank (2022). Drugbank Online. www.go.drugbank.com [Accessed October 24, 2022].

[4] EMA (2022). European Public Assessment Reports. www.ema.europe.eu [Accessed October 24, 2022].

[5] Milton, P. *et al.* (2020) Moxidectin: an oral treatment for human onchocerciasis. *Expert Rev. Anti Infect. Ther.* 18, 1067–1081.

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

KEYNOTE

[6] Ouellet, D. *et al.* (2013) Effects particle size, food, and capsule shell composition on the oral bioavailability of dabrafenib, a BRAF inhibitor, in patients with BRAF mutation-positive tumors. *J. Pharm. Sci.* 102, 3100–3109.
[7] Kala, S.G. and Chinni, S. (2021) Development and characterization of venetoclax nanocrystals for oral bioavailability enhancement. *AAPS PharmSciTech* 22, 92.
[8] Dillon, J.S. and Chandrasekharan, C. (2018) Telotristat ethyl: a novel agent for the therapy of carcinoid syndrome diarrhea. *Future Oncol.* 14, 1155–1164.
[9] Canga, A.G. *et al.* (2008) The pharmacokinetics and interactions of ivermectin in humans - a mini-review. *AAPS J.* 10, 42–46.
[10] Strickley, R.G. (2007) Currently marketed oral lipid-based dosage forms: drug products and excipients. In *Oral Lipid-Based Formulations: Enhancing the Bioavailability of Poorly Water-Soluble Drugs* Hauss D.J., ed. pp. 1–32, CRC PressBoca Raton.
[11] Nicolaides, E. *et al.* (1999) Forecasting the *in vivo* performance of four low soluble drugs from their in vitro dissolution data. *Pharm. Res.* 16, 1876–1882.
[12] Madsen, C.M. *et al.* (2016) Supersaturation of zafirlukast in fasted and fed state intestinal media with and without precipitation inhibitors. *Eur. J. Pharm. Sci.* 91, 31–39.
[13] Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rate. *Nat. Rev. Drug Discov.* 3, 711–715.
[14] Waring, M.J. *et al.* (2015) An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discov.* 14, 475–486.
[15] Lipinski, C.A. *et al.* (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development. *Adv. Drug Deliv. Rev.* 46, 3–26.
[16] Brown, D.G. and Boström, J. (2018) Where do recent small molecule clinical development candidates come from? *J. Med. Chem.* 61, 9442–9468.
[17] Brown, D.G. and Wobst, H.J. (2021) A decade of FDA-approved drugs (2010–2019): trends and future directions. *J. Med. Chem.* 64, 2312–2338.
[18] Shultz, M.D. (2019) Two decades under the influence of the Rule of Five and the changing properties of approved oral drugs. *J. Med. Chem.* 62, 1701–1714.
[19] Leveridge, M. *et al.* (2018) Integration of lead discovery tactics and the evolution of the lead discovery toolbox. *SLAS Discov.* 23, 881–897.
[20] de Souza Neto, L.R. *et al.* (2020) In silico strategies to support fragment-to-lead optimization in drug discovery. *Front. Chem.* 8, 93.
[21] Swinney, D.C. (2013) Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin. Pharmacol. Ther.* 93, 299–301.
[22] Mainardi, S. and Bernards, R. (2022) A large-scale organoid-based screening platform to advance drug repurposing in pancreatic cancer. *Cell Genomics* 2 100100.
[23] Marei, I. *et al.* (2022) 3D tissue-engineered vascular drug screening platforms: promise and considerations. *Front. Cardiov. Med.* 9 847554.
[24] Sun, D. *et al.* (2022) Why 90% of clinical drug development fails and how to improve it? *Acta Pharm. Sin. B* 12, 3049–3062.
[25] Schneider, G. (2018) Automating drug discovery. *Nat. Rev. Drug Discov.* 17, 97–113.
[26] Gorgulla, C. *et al.* (2020) An open-source drug discovery platform enables ultra-large virtual screens. *Nature* 580, 663–668.
[27] Veber, D.F. *et al.* (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45, 2615–2623.
[28] Tinworth, C.P. and Young, R.J. (2020) Facts, patterns, and principles in drug discovery: appraising the rule of 5 with measured physicochemical data. *J. Med. Chem.* 63, 10091–10108.
[29] Young, R.J. and Leeson, P.D. (2018) Mapping the efficiency and physicochemical trajectories of successful optimizations. *J. Med. Chem.* 61, 6421–6467.
[30] Bunally, S.B. *et al.* (2019) Using physicochemical measurements to influence better compound design. *SLAS Discov.* 24, 791–801.
[31] Fullam, E. and Young, R.J. (2021) Physicochemical properties and *Mycobacterium tuberculosis* transporters: keys to efficacious antitubercular drugs? *RSC Med. Chem.* 12, 43–56.
[32] Sutherland, J.J. *et al.* (2012) Relating molecular properties and *in vitro* assay results to *in vivo* drug disposition and toxicity outcomes. *J. Med. Chem.* 55, 6455–6466.
[33] Waring, M.J. (2009) Defining optimum lipophilicity and molecular weight ranges for drug candidates-molecular weight dependent lower log D limits based on permeability. *Bioorg. Med. Chem. Lett.* 19, 2844–2851.
[34] Johnson, T.W. *et al.* (2009) Using the Golden Triangle to optimize clearance and oral absorption. *Bioorg. Med. Chem. Lett.* 19, 5560–5564.
[35] Degoey, D.A. *et al.* (2018) Beyond the Rule of 5: lessons learned from AbbVie's drugs and compound collection. *J. Med. Chem.* 61, 2636–2651.
[36] Caron, G. and Ermondi, G. (2017) Updating molecular properties during early drug discovery. *Drug Discov. Today* 22, 835–840.
[37] Whitty, A. *et al.* (2016) Quantifying the chameleonic properties of macrocycles and other high-molecular-weight drugs. *Drug Discov. Today* 21, 712–717.
[38] Sheikh, A.Y. *et al.* (2021) Implications of the conformationally flexible, macrocyclic structure of the first-generation, direct-acting anti-viral paritaprevir on its solid form complexity and chameleonic behavior. *J. Am. Chem. Soc.* 143, 17479–17491.
[39] Kuentz, M. and Bergström, C.A.S. (2021) Synergistic computational modeling approaches as team players in the game of solubility predictions. *J. Pharm. Sci.* 110, 22–34.
[40] Yalkowsky, S.H. and Valvani, S.C. (1980) Solubility and partitioning I: solubility of nonelectrolytes in water. *J. Pharm. Sci.* 69, 912–922.
[41] Abraham, M.H. and Le, J. (1999) The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship. *J. Pharm. Sci.* 88, 868–880.
[42] Breiman, L. (2001) Random forests. *Mach. Learn.* 45, 5–32.
[43] Avdeef, A. and Kansy, M. (2020) 'Flexible-acceptor' general solubility equation for beyond rule of = drugs. *Mol. Pharm.* 17, 3930–3940.
[44] Avdeef, A. and Kansy, M. (2022) Predicting solubility of newly-approved drugs (2016–2020) with a simple ABSOLV and GSE (flexible-acceptor) consensus model outperforming random forest regression. *J. Solution Chem.* 51, 1020–1055.
[45] Miller, N.A. *et al.* (2019) Physiologically based pharmacokinetic modelling for first-in-human predictions: an updated model building strategy illustrated with challenging industry case studies. *Clin. Pharmacokinet.* 58, 727–746.
[46] Göller, A.H. *et al.* (2020) Bayer's in silico ADMET platform: a journey of machine learning over the past two decades. *Drug Discov. Today* 25, 1702–1709.
[47] Bhutani, P. *et al.* (2021) U.S. FDA approved drugs from 2015-June 2020: a perspective. *J. Med. Chem.* 64, 2339–2381.
[48] Doak, B.C. *et al.* (2014) Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chem. Biol.* 21, 1115–1142.
[49] Caron, G. *et al.* (2021) Steering new drug discovery campaigns: permeability, solubility, and physicochemical properties in the bRo5 chemical space. *ACS Med. Chem. Lett.* 12, 13–23.
[50] Avdeef, A. and Kansy, M. (2020) Can small drugs predict the intrinsic aqueous solubility of 'beyond Rule of 5' big drugs? *ADMET DMPK* 8, 180–206.
[51] Perkins, R. *et al.* (2003) Quantitative structure-activity relationship methods: perspectives on drug discovery and toxicology. *Ann. Rev.* 22.
[52] Patel, H.M. *et al.* (2014) Quantitative structure-activity relationship (QSAR) studies as strategic approach in drug discovery. *Med. Chem. Res.* 23, 4991–5007.
[53] Jakubík, J. *et al.* (2015) Towards predictive docking at aminergic G-protein coupled receptors. *J. Mol. Model.* 21, 1–18.
[54] Chiu, S.H. and Xie, L. (2016) Toward high-throughput predictive modeling of protein binding/unbinding kinetics. *J. Chem. Inf. Model.* 56, 1164–1174.
[55] Egbert, M. *et al.* (2019) Why some targets benefit from beyond Rule of Five drugs. *J. Med. Chem.* 62, 10005–10025.
[56] Saal, C. *et al.* (2022) Atropisomerism – a neglected way to escape out of solubility flatlands. *J. Pharm. Sci.* 111, 206–213.
[57] Bergström, C.A.S. *et al.* (2016) Computational prediction of formulation strategies for beyond-rule-of-5 compounds. *Adv. Drug Deliv. Rev.* 101, 6–21.
[58] Rowland, A. *et al.* (2017) Kinase inhibitor pharmacokinetics: comprehensive summary and roadmap for addressing inter-individual variability in exposure. *Expert Opin. Drug Metab. Toxicol.* 13, 31–49.
[59] Remko, M. *et al.* (2011) Molecular structure, pKa, lipophilicity, solubility, absorption, polar surface area, and blood brain barrier penetration of some antiangiogenic agents. *Struct. Chem.* 22, 635–648.
[60] Weerakoon, D. *et al.* (2022) Impact of PROTAC linker plasticity on the solution conformations and dissociation of the ternary complex. *J. Chem. Inf. Model.* 62, 340–349.
[61] Owen, D.R. *et al.* (2021) An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19. *Science* 374, 1586–1593.
[62] Hossain Mithu, M.S. *et al.* (2021) Advanced methodologies for pharmaceutical salt synthesis. *Cryst. Growth Des.* 21, 1358–1374.
[63] Bharate, S.S. (2021) Carboxylic acid counterions in FDA-approved pharmaceutical salts. *Pharm. Res.* 38, 1307–1326.
[64] Bharate, S.S. (2021) Modulation of biopharmaceutical properties of drugs using sulfonate counterions: a critical analysis of FDA-approved pharmaceutical salts. *J. Drug Delivery Sci. Technol.* 66 120993.
[65] Bharate, S.S. (2021) Recent developments in pharmaceutical salts: FDA approvals from 2015 to 2019. *Drug Discov. Today* 26, 384–398.

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344

[66] Patel, M.A. *et al*. (2018) Assessing the risk of salt disproportionation using crystal structure and surface topography analysis. *Cryst. Growth Des*. 18, 7027–7040.
[67] Xiang, S.H. and Tan, B. (2020) Advances in asymmetric organocatalysis over the last 10 years. *Nat. Commun*. 11, 3786.
[68] Gioiello, A. *et al*. (2020) The medicinal chemistry in the era of machines and automation: recent advances in continuous flow technology. *J. Med. Chem*. 63, 6624–6647.
[69] Voight, E.A. *et al*. (2021) Fueling the pipeline via innovations in organic synthesis. *ACS Med. Chem. Lett*. 12, 1365–1373.
[70] Schiesser, S. *et al*. (2021) The powerful symbiosis between synthetic and medicinal chemistry. *Future Med. Chem*. 13, 941–944.
[71] Atanasov, A.G. *et al*. (2021) International Natural Product Sciences Taskforce, Supuran CT. Natural products in drug discovery: advances and opportunities. *Nat. Rev. Drug Discov*. 20, 200–216.
[72] Attwood, M.M. *et al*. (2021) Trends in kinase drug discovery: targets, indications and inhibitor design. *Nat. Rev. Drug Discov*. 20, 839–861.
[73] Möschwitzer, J.P. (2013) Drug nanocrystals in the commercial pharmaceutical development process. *Int. J. Pharm*. 453, 142–156.
[74] Iyer, R. *et al*. (2021) Amorphous solid dispersions (ASDs): the influence of material properties, manufacturing processes and analytical technologies in drug product development. *Pharmaceutics* 13, 1682.
[75] Bauer, J. *et al*. (2001) Ritonavir: an extraordinary example of conformational polymorphism. *Pharm. Res*. 18, 859–866.
[76] Ma, X. and Williams, R.O. (2019) Characterization of amorphous solid dispersions: An update. *J. Drug Delivery Sci. Technol*. 50, 113–124.
[77] Vehring, R. (2008) Pharmaceutical particle engineering via spray drying. *Pharm. Res*. 25, 999–1022.
[78] Simões, M.F. *et al*. (2019) Hot-melt extrusion in the pharmaceutical industry: toward filing a new drug application. *Drug Discov. Today* 24, 1749–1768.
[79] Patil, H. *et al*. (2016) Hot-melt extrusion: from theory to application in pharmaceutical formulation. *AAPS PharmSciTech* 17, 20–42.
[80] Vasconcelos, T. *et al*. (2016) Amorphous solid dispersions: rational selection of a manufacturing process. *Adv. Drug Deliv. Rev*. 100, 85–101.
[81] Bhujbal, S.V. *et al*. (2021) Pharmaceutical amorphous solid dispersion: a review of manufacturing strategies. *Acta Pharm. Sin. B* 11, 2505–2536.
[82] Dong, Z. and Choi, D.S. (2008) Hydroxypropyl methylcellulose acetate succinate: potential drug - excipient incompatibility. *AAPS PharmSciTech* 9, 991–997.
[83] Friesen, D.T. *et al*. (2008) Hydroxypropyl methylcellulose acetate succinate-based spray-dried dispersions: an overview. *Mol. Pharm*. 5, 1003–1019.
[84] Bennett-Lenane, H. *et al*. (2020) A retrospective biopharmaceutical analysis of >800 approved oral drug products: are drug properties of solid dispersions and lipid-based formulations distinctive? *J. Pharm. Sci*. 109, 3248–3261.
[85] Feeney, O.M. *et al*. (2016) 50 years of oral lipid-based formulations: provenance, progress and future perspectives. *Adv. Drug Deliv. Rev*. 101, 167–194.
[86] Savla, R. *et al*. (2017) Review and analysis of FDA approved drugs using lipid-based formulations. *Drug Dev. Ind. Pharm*. 43, 1743–1758.
[87] Perera, M.D. *et al*. (2018) Using structural mimics for accessing and exploring structural landscapes of poorly soluble molecular solids. *Acta Crystallogr. Sect. B Struct. Sci., Cryst. Eng. Mater*. 74, 42–48.
[88] Kuentz, M. *et al*. (2021) Rational selection of bio-enabling oral drug formulations – a PEARRL commentary. *J. Pharm. Sci*. 110, 1921–1930.
[89] Fauman, E.B. *et al*. (2011) Structure-based druggability assessment-identifying suitable targets for small molecule therapeutics. *Curr. Opin. Chem. Biol*. 15, 463–468.
[90] Landis, M.S. *et al*. (2018) Commentary: why pharmaceutical scientists in early drug discovery are critical for influencing the design and selection of optimal drug candidates. *AAPS PharmSciTech* 19, 1–10.
[91] Maas, J. *et al*. (2007) An integrated early formulation strategy - from hit evaluation to preclinical candidate profiling. *Eur. J. Pharm. Biopharm*. 66, 1–10.

Please cite this article in press as: Stegemann, S. et al., Drug Discovery Today (2022), https://doi.org/10.1016/j.drudis.2022.103344