



Heriot-Watt University
Research Gateway

PepTherDia

Citation for published version:

D'Aloisio, V, Dognini, P, Hutcheon, GA & Coxon, CR 2021, 'PepTherDia: database and structural composition analysis of approved peptide therapeutics and diagnostics', *Drug Discovery Today*, vol. 26, no. 6, pp. 1409-1419. <https://doi.org/10.1016/j.drudis.2021.02.019>

Digital Object Identifier (DOI):

[10.1016/j.drudis.2021.02.019](https://doi.org/10.1016/j.drudis.2021.02.019)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

Published In:

Drug Discovery Today

Publisher Rights Statement:

© 2021 Elsevier Ltd.

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

1 **PepTherDia: database and structural composition analysis of**
2 **approved peptide therapeutics and diagnostics (2020)**

3 **Vera D'Aloisio¹, Paolo Dognini¹, Gillian A. Hutcheon¹, and Christopher R. Coxon^{2*}**

4 ¹School of Pharmacy and Biomolecular Sciences, Faculty of Science, Liverpool John Moores
5 University, Liverpool, United Kingdom, L3 3AF

6 ²Institute of Chemical Sciences, School of Engineering and Physical Sciences, Heriot-Watt
7 University, Edinburgh, United Kingdom, EH14 4AS

8 *** Corresponding author:**

9 Christopher R. Coxon (email: c.coxon@hw.ac.uk; Tel: +44 131 451 8029)

10 **Keywords:** peptide therapeutics and diagnostics, peptide drug, peptide market, structural
11 analysis, approved peptides database

12 **Teaser:** We describe a freely-accessible database of approved peptide therapeutics and
13 diagnostics, providing an overview of some key structural and compositional trends to help
14 guide the design of future peptide medicines.

15 Abstract

16 As of 2020, there are over 100 approved peptides with therapeutic or diagnostic applications.
17 However, a complete database providing information on marketed peptides is not freely
18 available, making the peptide chemists' job of designing future peptide drug candidates
19 challenging. Unlike the rules for small molecule drugs, there is no general set of guidelines
20 for designing a successful peptide-based drug. This foundation review, together with our
21 freely available database (PepTherDia, <http://peptherdia.herokuapp.com>), provides some
22 insights into what a successful peptide therapeutic or diagnostic agent looks like and will lay
23 the foundation for establishing a set of rules to help future medicinal chemists to design
24 peptide candidates with increased approval rates.

25 Introduction

26 Following the approval of the first peptide therapeutic agent, the 51-amino acid hormone
27 insulin in 1923 [1,2], drug discovery has been progressively expanding into the chemical
28 space between small molecules and large proteins. Subsequently, a significant number of
29 peptides (and peptidomimetics) have received regulatory approval. Recently, peptides have
30 emerged as novel modalities for various applications in the therapeutic and diagnostic
31 markets, providing new opportunities for the modulation of difficult targets. As shown in
32 Figure 1, since the second half of the last century, the number of peptides on the therapeutics
33 and diagnostics market has steadily increased, reaching the milestone of over 100 approved
34 peptide drugs in 2020. These drugs represent a unique class of chemical compounds that
35 bridges the gap between small molecules (typically molar mass < 500 g/mol) and large
36 biologics (typically molar mass > 5000 g/mol). Occupying an intermediate region of
37 complexity and molar mass, they combine many of the benefits of the two abovementioned
38 categories. The main disease areas presently treated with peptide drugs are metabolic
39 disorders, cancer and cardiovascular disease, but other emerging therapeutic applications are
40 in the areas of infectious diseases, pain, urinary tract, gastrointestinal and respiratory
41 disorders [1,3]. As the incidence of metabolic disorders, cancer and cardiovascular diseases
42 in the Western world is increasing alarmingly [4–6] and the need for new effective medicines
43 to treat emerging health problems (e.g., SARS-CoV-2/COVID-19) is growing, it is highly
44 likely that the demand for peptide drugs will continue to grow.

45 Current state of peptide therapeutics and diagnostics

46 According to Transparency Market Research's latest report, the worldwide market for
47 peptide pharmaceuticals has been growing at a Compound Annual Growth Rate (CAGR) of
48 roughly 8%; this is expected to increase over time with the same trend, reaching a value of 50
49 billion USD in 2027 from 25 billion USD in 2018 [7].

50 As of November 2020, according to our database analysis, there are 105 peptide
51 pharmaceutical products (see definitions below) with regulatory approval in the main
52 pharmaceutical markets - North America, Europe and Japan – of which 89 are peptide drugs
53 and 16 are diagnostic agents. Moreover, nowadays, a great number of clinical studies involve
54 peptide agents – 4,859 in total, 468 of which are in phase 3 [8]. This suggests that the
55 pharmaceutical industry is committed to exploring the role of peptide therapeutics on
56 modulating previously 'undruggable' targets and addressing unmet medical needs. In Table
57 1, the main advantages of peptides over small molecules and proteins are illustrated. Despite
58 these important benefits, the drug development process for future therapeutic peptides from
59 laboratory to approval has traditionally presented many obstacles. In fact, unique challenges,
60 such as chemical and physical instability, short circulating half-life, high proteolytic

61 degradation, rapid renal clearance, poor membrane permeability, poor oral bioavailability,
62 and low solubility must be addressed in order to bring a peptide to clinical use [1,9].
63 Nonetheless, lately, advances in drug delivery and emerging medicinal chemistry strategies
64 have brought peptides to a significant renaissance, by overcoming their issues and eventually
65 improving pharmacokinetic profiles and oral bioavailability.

66 In the last decade, several research groups have tried to reach a better understanding of
67 approved peptides and their properties. In 2010, Vlieghe *et al* reviewed and listed the
68 synthetic therapeutic peptides that have reached the main pharmaceutical markets (US, EU,
69 and Japan) [10]. Six years later, Raghava and collaborators filed a repository (PEPlife) to
70 provide the scientific community with data on peptide half-lives [11], followed, one year
71 later, by a database containing 852 FDA-approved biologics, among which 28 were peptides
72 [12]. Recently, an extensive review on approved peptide therapeutics targeting G protein-
73 coupled receptors (GPCRs) have been published by Davenport and collaborators [13],
74 demonstrating their dominant presence in the market.

75 Nevertheless, to the best of the authors' knowledge, up to the present moment (November
76 2020), a complete database with structural analysis, production methods, pharmacokinetic
77 properties (i.e., terminal half-life and protein binding), indications, and routes of
78 administration of regulatory-approved peptides is not freely available online. In fact, the
79 information is scattered throughout the scientific literature and in various websites, making
80 the search for approved peptides very challenging. In addition to this, there is a critical lack
81 of rules that makes the medicinal chemists' job of designing entirely new potential peptide
82 drugs very difficult. In fact, as already pointed out by Tyagi and co. [14], the "drug-likeness"
83 criteria used for small molecules – principally, *Lipinski's Rule of Five, Ro5* [15], but also
84 models such as the central nervous system multiparameter optimization (CNS MPO) [16] -
85 are not applicable to peptides due to their entirely different intrinsic properties and
86 applications.

87 **PepTherDia**

88 In this context, we have developed and made accessible online **PepTherDia (Peptide**
89 **Therapeutics and Diagnostics:** <http://peptherdia.herokuapp.com>), a manually curated
90 database containing a searchable list of approved peptide drugs and diagnostic agents, with
91 information on their physicochemical and pharmacokinetic properties, as well as their routes
92 of administration and indications. Its purpose is to provide assistance to medicinal chemists
93 and scientists in the field of peptide drug discovery. On the compounds enumerated and
94 described in PepTherDia, we have performed a detailed analysis of the structural features and
95 collected information on their terminal half-life, plasma protein-binding, indication, route of
96 administration, production methodologies, marketing authorisation (year and agency of first
97 approval) and origin of their design. The information contained on this database will be
98 updated on a regular basis (e.g., yearly) with new approvals as well as new properties
99 investigated. For the first time, this review highlights some important trends in peptide
100 approvals, together with providing the reader with an insight into the features and
101 characteristics that are common in approved peptide agents. We envisage that this
102 information will aid the scientific community to more successfully design or pre-screen
103 candidates at an early stage of the peptide drug discovery process to increase longitudinal
104 approval rates.

105 *How is a "peptide" defined?*

106 A challenge in designing this study was the breadth and diversity in the so-called ‘peptide’
107 molecule due to differences in structure, size and composition. Therefore, the first thing the
108 authors felt necessary to clarify was: what is a peptide? The International Union of Pure and
109 Applied Chemistry (IUPAC) defines peptides as “amides derived from two or more amino
110 carboxylic acid molecules (the same or different) by formation of a covalent bond from the
111 carbonyl carbon of one to the nitrogen atom of another with formal loss of water” [17].
112 Whereas the recommendations of the IUPAC-IUB Joint Commission on Biochemical
113 Nomenclature (JCBN) defines a peptide as a chemical entity presenting from 2 to 50 amino
114 acid residues [18], the currently used regulatory FDA definition delineates a peptide as “any
115 polymer composed of 40 or fewer amino acids” and regardless their production method
116 [19,20]. Finally, the European Medicines Agency (EMA), instead of describing peptides
117 based on their size, considers them as small molecules if chemically synthesised, while treats
118 them as biological entities if they are extracted from natural sources or produced with
119 recombinant methodologies [21], which underlines that the definition of peptide is still
120 ambiguous.

121 At this stage, the reader might state that this definition can be considered as a philosophical
122 debate that may elicit controversial answers. Indeed, the scientific community differs greatly
123 on where to stop using the term peptide and start using the term protein. In the glossary
124 section of the “Glossary and Key Sources Box”, the definition of peptide that led our
125 research, together with other terminologies are explained.

126 *Data collection and calculations*

127 A repository of 105 compounds was obtained by searching in DrugBank [22], FDA and EMA
128 web pages [19,23], Pharmaceutical and Medical Devices Agency website [24], and Drug
129 Central website [25]. The key inclusion and exclusion criteria used in the curation of
130 PepTherDia are listed in Table 2. Examples of peptides not included and the reason for their
131 exclusion are shown in Table 3. With the aim of providing each peptide with a complete
132 profile comprising relevant information regarding terminal half-life, protein binding,
133 therapeutic indications, and routes of administration, specific searches were carried out in
134 DrugBank [22], National Centre for Advancing Translational Sciences web page [26],
135 Drugs.com [27], and pharmaceutical companies’ websites, using the generic name of the
136 individual peptide. References specific to each approved peptide as well as SMILES codes
137 used to calculate the peptide molar mass values can be found on our website PepTherDia.
138

139 *Structural composition analysis of approved peptide agents*

140 As discussed and defined earlier, peptide medicines are generally composed of natural amino
141 acids, unnatural amino acids and non-amino acidic modifications. Figure 2 provides an
142 example of how the peptide daptomycin may be divided into the above components.
143 However, in some cases, e.g., the glycopeptide antibiotics (dalbavancin, telavancin,
144 oritavancin, and teicoplanin), this is a very complex (if not impossible) task. Similarly, the
145 complexity of some multicyclic peptides does not allow the unambiguous identification of a
146 defined single macrocycle and, therefore, the members of each macrocycle were not counted
147 here.

148 Each constitutional member can be further classified as polar, acidic, basic, non-polar
149 aliphatic or aromatic based on its structural and physicochemical characteristics. For the
150 natural amino acidic residues, the designations polar, acidic, basic, non-polar aliphatic or
151 aromatic, derived from literature precedent [28], are generally ascribed by the nature of the

152 side chain. Since non-natural amino acid members form an amide backbone in the same way
153 as natural amino acids, they may be classified following the same principles used for natural
154 amino acids. In contrast, peptide modifications are a broad and varied structural class, and
155 their classification requires consideration of their complete structure and the way they are
156 conjugated to the peptide. For example, in daptomycin (Figure 2), decanoic acid can be
157 classified as a non-polar aliphatic modification as the carboxylic acid moiety becomes part of
158 an amide and its contribution to the final polarity predominantly increases lipophilicity.

159 The complete list of non-natural amino acids and modifications together with their polarity
160 classification can be found on PepTherDia website.

161 *Molar mass distribution and origin of peptide design*

162 Thanks to the improvements in synthetic and manufacturing technologies, nowadays it is
163 possible to synthesise ever-larger peptides in a short time, yielding high purities and
164 quantities. Nonetheless, from our study it emerged that the majority of approved peptides
165 (68%) are relatively ‘small’ peptides, composed of 2-16 constitutional members, with a
166 second minor cluster (27%) of larger size peptides with around 28-37 members (Figure 3A).
167 This is mirrored in a bimodal molar mass distribution in the ranges 300-1750 g/mol (major -
168 71%) and 2750-4250 g/mol (minor - 22%), with a remarkable lack of mid-length approved
169 peptides (Figure 3B). Hence, the data suggest that there are two main groups of peptides: low
170 molar mass and high molar mass with only a few examples in between (e.g., sinapultide and
171 ziconotide, with molar mass of 2469.45 and 2639.14 g/mol, respectively). To evaluate the
172 possibility of bias in the peptide design, which may have led to the deliberate development of
173 peptides of certain sizes to mimic specific biomolecules, it is necessary to analyse the origin
174 of the peptide design case by case (Figure 3E). In this context, it emerges that *natural*
175 *peptides* account for 30% of our sample. However, *peptide analogues* account for 54%.
176 Finally, *heterologous peptides* account for 16% of our sample. This underlines that
177 heterologous peptides are difficult to design *a priori* and are mostly discovered by library
178 screening. In light of these findings, we can state that, unsurprisingly, there is a clear trend (in
179 84% of the cases) towards following the route of inspiration from nature as a greater promise
180 of success and that the bimodal distribution may be attributed to the characteristics of the
181 natural molecules that have inspired the design. Examples that demonstrate this are the
182 natural nonapeptide oxytocin and the 32-membered peptide calcitonin. Here, in both cases,
183 their length is due to the size of the natural molecule of origin.

184 *Natural side chain amino acid occurrence*

185 Amongst all the constitutional members, the large majority (around 81%) is represented by
186 natural L-amino acids. The residual 19% comprises non-natural AAs and modifications. A
187 careful analysis of the amino acid residues contained in each approved peptide (Figure 3C)
188 showed that the most common amino acids found in the sequences are the non-polar aliphatic
189 leucine (L) and glycine (G), followed by the polar serine (S). On the contrary, the least
190 common residues are methionine (M), histidine (H) and isoleucine (I). This is largely in
191 agreement with the occurrence of natural amino acids in proteins: in nature, leucine (L)
192 accounts for 9.1%, serine (S) for 6.8%, glycine (G) for 7.2%, and alanine (A) for 7.8%, being
193 the most common amino acids. In contrast, methionine (M) and histidine (H) account each
194 only for 2.3%, cysteine (C) for 1.9%, and tryptophan (W) for 1.4%, being the least common
195 amino acid residues found in proteins [28]. The occurrence of cysteine (C) in pharmaceutical
196 peptides is higher than in proteins, which is due to the frequent use of disulfide bonds as a
197 tool for macrocyclisation (see section ‘*Conformational and shape properties*’). In general, it
198 is important to emphasise that the amino acid composition of proteins and peptides is highly

199 variable; some amino acids may occur only once or not at all in a given peptide and may be
200 repeated several times in another peptide sequence. An example of where this is seen is the
201 repetition of the moiety KL₄ in the peptide sinapultide, designed to mimic the C-terminal
202 domain of the surfactant protein B [29].

203 D-amino acids with natural side chains account only for a small percentage (around 4% of the
204 total amino acids with natural side chain) and are mainly represented by phenylalanine,
205 alanine tryptophan, and arginine. The selective replacement of L-amino acids by their
206 enantiomers (D-AAAs) can protect the molecule from protease degradation [30]. This is a
207 common technique to obtain proteolytic stabilisation by backbone modification, even if this
208 causes conformational changes that might affect biological activity. An example is the
209 somatostatin-like peptide octreotide, in which natural phenylalanine and tryptophan are
210 replaced with their mirror-image forms, leading to a 100-fold increase in the terminal half-
211 life [31,32].

212 *Non-natural amino acids occurrence*

213 In the field of peptide drug discovery, the use of non-natural amino acids as well as the
214 conjugation with non-amino acidic members are common techniques to overcome peptide
215 limitations and have been widely explored in both protein and peptide design [33,34]. In
216 Figure 3D, the most common non-natural amino acids and structural modifications in
217 approved peptides are reported.

218 Position-specific incorporation of non-natural amino acids bearing a variety of bespoke side
219 chains can provide improvements in peptide properties, activity and function [35]. In this
220 study, non-natural amino acids have emerged to be present in heterologous peptides (e.g.,
221 ACE inhibitors, macimorelin, argatroban) and peptide analogues (e.g., GnRH agonists,
222 buserelin, carbetocin, icatibant, pasireotide, carfilzomib) but also in one third of the natural
223 peptides. In fact, from our analysis, the percentage of proteinogenic amino acids in natural
224 peptide antibiotics (e.g., capreomycin, vancomycin, bleomycin, oritavancin), has been
225 estimated to be between 0 and 10% of all the constitutional members, a value that is
226 significantly lower than the average percentage of natural amino acids in the pool of 105
227 approved peptides (81%).

228 Ornithine (Orn), 2,4-Diaminobutyric acid (Dab), and 2,3-diaminopropionic acid (Dap) are
229 homologues of lysine (Lys), where the variability in structures is due to the difference in the
230 number of side chain carbons. They all contain an amino group on the side chain, which
231 exhibits a basic/ionisable contribution and, at the same time, allows opportunities for
232 cyclisation or conjugation. Moreover, it has been reported that the use of Dap and Dab in
233 antimicrobial peptides (AMPs) can prevent the hemolytic activity of positively charged
234 natural amino acids (i.e., Arg and Lys), solving a characteristic issue of AMPs [36]. AMPs
235 have also shown improved stability to trypsin while retaining their biological activity, when
236 Lys and Arg residues are replaced by Dab, Dap, or homoarginine [37], demonstrating that
237 Lys homologues are not suitable substrate of this hydrolytic enzyme. Likewise, lower
238 susceptibility to tryptic hydrolysis has been reported after Lys replacement with Orn [38]. In
239 our pool of approved peptides, Dap and Dab are only found in natural peptides – Dap in
240 capreomycin, viomycin, enviomycin, while Dab in colistin and polymyxin B. On the other
241 hand, Orn is used in a variety of peptide analogues (e.g., ornipressin, atosiban, anidulafungin,
242 caspofungin, and micafungin) as well as in natural peptides (e.g., daptomycin and bacitracin).
243 Overall, amino acid side chains bearing a primary amine, such as lysine and its homologues,
244 represent one of the most frequently used amino acid moieties (around 7% of the total amino

245 acids). Specifically, the Dab residue was encountered 12 times in total, but this amino acid
246 appears only in two peptide structures (colistin and polymyxin B), while Orn appears only
247 four times, each in different peptides (Daptomycin, Ornipressin, Atosiban, Bacitracin).

248 4-hydroxyproline (Hyp) is a proline containing a hydroxyl group on the pyrrolidine ring. In
249 general, the relatively frequent use of hydroxyproline derivatives (9 in total in the pool of
250 approved peptides, e.g., caspofungin, icatibant, voxilaprevir, to name but a few) may be
251 ascribable to the polarity enhancement, additional hydrogen-bonding ability and the
252 possibility of further conjugation gained with their introduction. Moreover, it has been
253 demonstrated that the presence of Hyp stabilises the triple-helical structure of collagen [39].
254 Similarly, this stabilisation may occur in peptide secondary structures.

255 Naphthyl-alanine (Nal) is often used to mimic tryptophan (Trp) and to explore potential
256 improvements in peptide pharmacological profiles [40–42]; however, it is not clear if 1-Nal
257 or 2-Nal adequately replicate the effects of Trp aromatic interactions. In fact, as with any
258 modification, the consequence of these replacements upon the peptide potency needs careful
259 assessment, because it has been demonstrated that substitution of Trp with 1-Nal or 2-Nal
260 decreases the potency of cholecystokinin analogues [43]. Nonetheless, this strategy has been
261 successfully used in the development of GnRH receptor (GnRH-R) antagonists (abarelix,
262 ganirelix, degarelix, and cetrorelix) and other peptides (lanreotide, nafarelin, pralmorelin).
263 Specifically, in GnRH-R blockers, the His – Trp motif of the natural hormone GnRH, has
264 been replaced with the three-amino acid motif 2-Nal – (4-Cl) Phe – 3-Pal, suggesting an
265 intention to improve the aromatic contribution at the peptide N-terminus, which may be
266 important for the binding and the antagonistic activity at the receptor.

267 *Non-amino acidic modifications*

268 An additional common strategy used to improve peptide drug-likeness is the introduction of
269 non-amino acid appendages to tune the pharmacokinetic properties. These are usually linked
270 to the main chain by only one functional group (e.g., -COOH for fatty acids or -OH for
271 sugars) and are generally attached to amino acids containing polar functional groups (e.g.,
272 OH, NH₂, COOH) or to the N- or C-termini. Enhancement of stability, protein binding, and
273 membrane permeability can be obtained through peptide lipid acylation, while improved
274 solubility and bioavailability can be achieved through glycosylation [44,45]. Indeed, lipid
275 acylation and glycosylation are the most common modifications encountered in the pool of
276 approved peptides.

277 Lipid acylation is a post-translational modification of proteins that has found applications in
278 peptide design, in order to improve pharmacokinetic and pharmacodynamic properties while
279 retaining the ability to bind the target receptor [44,46,47]. 13% of marketed peptides present
280 a lipophilic carbon chain attached to their structure and, in some cases, they demonstrated
281 prolonged terminal half-life (e.g., in oritavancin and dalbavancin that present a terminal half-
282 life of 245 and 346 hours, respectively,) and high protein binding (>90% in the majority of
283 lipidated approved peptides). The length of the carbon chain may influence the half-life
284 duration but, at the moment, we do not have enough data around approved lipidated peptides
285 to state so. Other examples of approved peptides presenting a lipophilic carbon chain, include
286 the popular diabetes medicines and GLP-1 receptor agonists, liraglutide (conjugated with
287 palmitic acid) and semaglutide (conjugated with an octadecanedioic acid). They both bind
288 with high affinity to plasma proteins (98-99% of the peptide bound), promoting greater
289 peptide stability that results in significantly extended half-lives of 13 and 168 hours,
290 respectively, compared with the parent GLP-1.

291 Carbohydrate groups are less frequent but still significant modifications, being found in 8 out
292 of 105 peptides on the market (e.g., bleomycin and vancomycin). Given the synthetic
293 challenges that glycochemistry presents, these are typically found in peptides of natural
294 origin. In these peptides, glycosyl units are attached to the main structure via an N-terminal
295 amine group or hydroxyl group on the side chain, similar to recombinant glycoprotein
296 therapeutics, in which carbohydrates are commonly N-linked to asparagine or O-linked to
297 serine and threonine [48]. It is probably not by chance that these 8 examples comprise a large
298 aromatic core or aliphatic chain, in which hydrophobicity is balanced by one or more sugars.
299 In fact, glycosylation improves the physicochemical and pharmacokinetics properties of
300 peptide drugs through an enhancement of solubility and an increase in bioavailability and oral
301 absorption [21,45].

302 Other modifications can include metal cation-chelating agents (DOTA in dotatate and DTPA
303 in pentetreotide), typically found in diagnostic agents or linkers (2-amino-4,6-dimethyl-3-
304 oxo-3H-phenoxazine-1,9-dicarbonyl in dactinomycin). Finally, pyroglutamic acid – cyclic
305 lactam of glutamic acid – is naturally found at the N-terminus of many neuronal peptides and
306 hormones but its function in living cells is still unclear [49]. In drug design, N-pyroglutamyl
307 formation is a common modification used to cap the N-terminus in order to modulate peptide
308 activity and increase resistance to degradation [50]. Indeed, in some cases, Pyr has shown to
309 be essential in order to achieve full biological activity [51]. This modification is found in 8
310 approved peptides, such as leuprolide and all GnRH agonists.

311 Notably, voxilaprevir is the first example of fluorinated peptide on the market, bearing four
312 fluorine atoms: a difluoromethylene adjacent to a benzopyrazine modification involved in
313 forming a macrocyclic structure, and a difluoromethyl group on an aliphatic non-natural
314 amino acid – 1-amino-2-(difluoromethyl)cyclopropane-1-carboxylic acid. We predict that the
315 exploitation of fluorine in peptides will follow what has already happened in small
316 molecules, by becoming a key medicinal chemistry tool, in which the judicious addition of a
317 small and highly electron withdrawing atom such as fluorine has been shown to play a key
318 role in improving pharmacokinetic and physicochemical properties [52]. Hence, in the
319 current peptide drug discovery pipeline, fluorination has been found to increase thermal
320 stability and proteolytic stability, without affecting biological activity [53]. This has been
321 applied to glucagon-like peptide-1 (GLP-1), in which the substitution with hexafluoroleucine
322 in different positions has been shown to improve both binding affinity and protease (DPP IV)
323 stability [54].

324 *Polarity trends*

325 Further examination of the peptide structures reveals that there is on average a balance
326 between polar and hydrophobic residues, when the polar contribution is derived by the
327 summation of the polar, basic and acidic constitutional members and the hydrophobic
328 contribution is the summation of the aromatic and aliphatic constitutional members. Figure
329 4A shows that the majority of the approved peptides contain from 35 to 75% of polar
330 residues, indicating that these molecules do not present a high excess of either hydrophilic or
331 lipophilic components. This perhaps should not be a surprise, given that both polar and
332 hydrophobic components are generally required for drugs with good pharmacokinetic
333 profiles. A small number of outliers that comprise 100% polar or 100% hydrophobic building
334 blocks are present. However, these exceptions are generally represented by a very small
335 number of building blocks (2, 3 or 5) and their small size makes these peptides more similar
336 to small molecules. Indeed, the hydrophilic dipeptide spaglamic acid as well as the
337 hydrophobic tripeptides ACE inhibitors (enalapril, perindopril, ramipril, quinapril, and

338 trandolapril) respect the Lipinski's rule of five tailored for small molecules (computed by
339 ChemAxon, Chemicalize [55]). Other outliers are represented by the growth hormone
340 secretagogue receptor agonist macimorelin composed of only 3 building blocks, and the
341 hydrophobic antiviral peptides telaprevir, boceprevir, and ombitasvir, composed of 4 or 5
342 members. Finally, cyclosporine is the only case in which a marketed peptide composed by
343 more than 5 members (11 in this case) completely lacks the polar component. This is
344 reflected in its very long half-life (19 hours) [56] and the fact that, after administration, 90%
345 is found to be bound to serum proteins, mainly lipoproteins [57]. In this respect, it must be
346 underlined that the classification of natural and non-natural amino acids is solely based on the
347 nature of their side chain. Assuming that the backbone impact on the final properties is
348 largely consistent with the size of the peptide, its polar contribution has not been considered.
349 This is the clearest way to distinguish between hydrophilic and hydrophobic amino acids,
350 without overshadowing the contribution of the side chain with that of the backbone. To
351 provide the reader with a visual overview of the relative physical property balance, the colour
352 map in Figure 4B shows the % composition of each approved peptide when the constitutional
353 members are colour-coded according to their side chain properties: polar, acidic, basic, non-
354 polar aliphatic, and aromatic.

355 *Conformational properties*

356 Another important aspect of structural composition is the proportion of peptide drugs and
357 diagnostics that are either linear or contain a macrocycle. In nature, cyclic peptides of various
358 sizes (from 8 to 50 AAs) occur in all the kingdoms of life. Their enhanced stability and
359 advantageous biopharmaceutical properties make their application in drug design very
360 common [58]. Indeed, as shown in Figure 3F, 53% of the marketed peptides are linear, while
361 47% present one or more macrocycles in their structure. Among the approved cyclic peptides,
362 39% are of natural origin, 55% are analogues and only 6% are heterologous – perhaps nature
363 has once again demonstrated how to develop stable, biocompatible peptides, and is
364 consequently a rich source of inspiration for candidates with optimal drug-like properties.
365 Interestingly, our analysis highlighted that, overwhelmingly, peptide macrocycles consist of 5
366 to 7 residues, with only a few exceptions (e.g., nesiritide and carperitide comprise a 17-
367 membered ring), as shown in Figure 3G. Examples of peptides with 5 to 7-membered
368 macrocycles include oxytocin, desmopressin, lanreotide, and eptifibatide, to name but a few.
369 In general, smaller macrocycles tend to have greater conformational stability due to physical
370 restraints and fewer rotatable bonds. As such, macrocyclisation is a very common medicinal
371 chemistry technique used to enhance peptide conformational stability and restrict the usual
372 peptide chain flexibility [59,60]. This may stabilise peptide conformation for optimal receptor
373 complexation and confer a protein-like secondary and tertiary structure [60–62].

374 Depending on the desired site of cyclisation, there are various strategies to generate cyclic
375 peptides [63]. These can involve the peptide head (peptide C-terminal moiety), the peptide tail
376 (peptide N-terminal moiety), or amino acid side chains. According to our findings, summarised
377 in Fig. 3H, the most common technique is side chain-to-side chain cyclisation (58% of all the
378 marketed cyclic peptides), with 25 out of 26 side chain-to-side chain macrocycles formed by a
379 disulfide bond between cysteine thiol pairs. The exception is bremelanotide, in which a lactam
380 is formed between the amine side chain of a lysine residue and the carboxylic acid side chain
381 of an aspartic acid residue. However, disulfide bridges are not always metabolically stable *in*
382 *vivo*, limiting their application [64]. Macrocycles within a peptide can be formed also by head-
383 to-side chain cyclisation (24%), mainly *via* lactamisation between the C-terminal carboxylic
384 acid and a side chain amine (e.g., lysine). The head-to-tail cyclisation between the N- and C-
385 termini (7%), generates an all-amide end-to-end cyclic lactam, thus abrogating exopeptidase

386 hydrolysis. Finally, another cyclisation strategy encountered in 2 out of 47 cyclic peptides (i.e.,
387 grazoprevir and elcatonin) is side chain-to-tail macrocyclisation. When the head and/or the tail
388 of the peptide are involved in the macrocycle, protease access to the backbone is reduced as
389 the cyclisation removes the free N- and C-termini that are targeted by amino- and carboxy-
390 peptidases, respectively [65]. This explanation bears out in the approved peptides, with the
391 mean experimental terminal half-life of cyclic peptides compared to linear peptides, being 27
392 h and 12 h, respectively.

393 Amongst marketed peptide therapeutics and diagnostics, 38% present an amidated C-
394 terminus, while 10% present an acetyl group at the N-terminus (Figures 3I and 3J).
395 Modifications of peptide N-terminus include also the addition of pyroglutamic acid (7%) or
396 deamination of the last AA (4%). Similarly to the head-to-tail cyclisation, this abrogates
397 exopeptidase hydrolysis by masking amino- and carboxyl- termini [66]. Moreover, N-
398 terminal acetylation or C-amidation precludes ionisation and hydrogen-bonding of NH₂ and
399 COOH groups, respectively [67,68], thus better mimicking natural proteins.

400 **Conclusions and future prospects**

401 For the first time, this foundation review together with the database PepTherDia, offers the
402 possibility of exploring common trends in approved peptide drugs and diagnostics. We
403 highlighted some of the strategies most commonly used in peptide drug design, which
404 successfully brought these peptides to the market. The trends underlined cannot be ascribed
405 to luck or coincidence. The majority of approved peptides (84%) follows the rules
406 “established” by nature over several millennia of evolution as they are naturally-derived or
407 analogues of natural compounds. It is open to debate among drug discovery scientists
408 whether it is important to follow the route of nature to obtain a successful peptide lead or
409 whether the exploration of completely new chemical space and compounds through rational
410 design and library screening is a fruitful solution. We believe that strategies provided by
411 nature have to be properly comprehended and taken into account when designing new drug
412 candidates.

413 On account of this meta-analysis, we can conclude that, at present, a peptide most likely to
414 become a drug will have a molar mass lower than 2000 g/mol and will present a balance
415 between hydrophobic and polar contributions. Furthermore, careful evaluation of the C- and
416 N-terminal modifications will be key for the pharmacokinetic properties as well as for the
417 desired activity. Moreover, if a cyclisation strategy is under evaluation, it is important for a
418 peptide chemist to consider a more stable small-size macrocycle (5 to 7 members). In
419 addition to the most frequently encountered natural amino acids, we have provided a list of
420 commonly employed non-natural amino acids that can effectively replace the former to
421 introduce enhanced properties or opportunities for further diversification. Indeed, strategies
422 to modify a peptide through conjugation with fatty acids or sugars are well represented in
423 various approvals, suggesting that they are promising tools to modify peptidic structures
424 while retaining favourable pharmacological properties. The highlighted common threads that
425 bring together the approved peptides does not have the presumption of being strict rules to
426 follow, but rather they should be seen as a means of lending a hand to peptide designers to
427 avoid certain pitfalls during early drug discovery stages. Indeed, some of the above trends
428 may be particularly subjective towards certain disease areas or routes of administration.

429 Forecasting a starring role for peptides in the next decades, we anticipate that new clear
430 trends – and perhaps rules – in structural composition will emerge beyond our observations,
431 leading to improved rational peptide drug design. Whilst the mainstay of bioactive peptide

432 discovery remains an analogue-based approach – modifying naturally-derived peptides or
433 protein epitopes – new medicinal chemistry strategies e.g., peptide-stapling, multicyclisation,
434 novel amino acid synthesis, selective-fluorination, peptoids and improved *in silico* design,
435 should accelerate the future approval rate of new designed or heterologous peptide
436 pharmaceutical agents. In the medium term, a combination of both naturally derived and
437 rationally designed strategies is likely to be the most successful route to fulfilling the
438 increasing need for chemical entities to treat new or previously untreatable diseases, from
439 cancer to infections, cardiovascular and neurodegenerative diseases.

440 **Acknowledgements**

441 The authors gratefully acknowledge Daniele Tomasi for the development of the website
442 PepTherDia.

443
444 This work was supported by a Liverpool John Moores University Vice Chancellor's PhD
445 Scholarship (VDA) and a University Alliance Doctoral Training Alliance COFUND/Marie
446 Skłodowska-Curie (European Union's Horizon 2020 research and innovation programme,
447 grant agreement No 801604) PhD Fellowship Programme in Applied Biosciences for Health
448 (PD). The funding sources had no involvement in the study design.

449 **Conflict of interest**

450 Dr Coxon is a Director of Pepmotec Ltd, a peptide synthesis spin-out company from Durham
451 University, UK.

452 **Supplementary Material**

453 The website PepTherDia contains the full list and classification of non-natural amino acids
454 and modifications together with the detailed methodologies used for data collection,
455 computational analysis and structural analysis.

456 **References**

- 457 1 Henninot, A., Collins, J.C. and Nuss, J.M. (2018) The Current State of Peptide Drug
458 Discovery: Back to the Future? *Journal of Medicinal Chemistry*, **61**, 1382–1414.
459 <https://doi.org/10.1021/acs.jmedchem.7b00318>.
- 460 2 Lau, J.L. and Dunn, M.K. (2018) Therapeutic Peptides: Historical Perspectives,
461 Current Development Trends, and Future Directions. *Bioorganic and Medicinal*
462 *Chemistry*, **26**, 2700–2707. <https://doi.org/10.1016/j.bmc.2017.06.052>.
- 463 3 Tsomaia, N. (2015) European Journal of Medicinal Chemistry Peptide Therapeutics:
464 Targeting the Undruggable Space. *European Journal of Medicinal Chemistry*, **94**,
465 459–470.
- 466 4 Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin,
467 D.M., Forman, D. and Bray, F. (2015) Cancer Incidence and Mortality Worldwide:
468 Sources, Methods and Major Patterns in GLOBOCAN 2012. *International Journal of*
469 *Cancer*, Wiley-Liss Inc., **136**, E359–E386. <https://doi.org/10.1002/ijc.29210>.
- 470 5 Daviglus, M.L., Lloyd-Jones, D.M. and Pirzada, A. (2006, September 20) Preventing
471 Cardiovascular Disease in the 21st Century: Therapeutic and Preventive Implications

- 472 of Current Evidence. *American Journal of Cardiovascular Drugs*, Springer, 87–101.
473 <https://doi.org/10.2165/00129784-200606020-00003>.
- 474 6 Zimmet, P.Z., Magliano, D.J., Herman, W.H. and Shaw, J.E. (2014, January 1)
475 Diabetes: A 21st Century Challenge. *The Lancet Diabetes and Endocrinology*,
476 Elsevier, 56–64. [https://doi.org/10.1016/S2213-8587\(13\)70112-8](https://doi.org/10.1016/S2213-8587(13)70112-8).
- 477 7 Peptide Therapeutics Market to Expand at a CAGR of 7.9% from 2019 to 2027.
478 Transparency Market Research, (2019)
479 www.transparencymarketresearch.com/peptide-therapeutics-market.html
480 [Accessed August 4, 2020]
- 481 8 ClinicalTrials.Gov
482 <https://www.clinicaltrials.gov/>
483 [Accessed November 20, 2020]
- 484 9 Craik, D.J., Fairlie, D.P., Liras, S. and Price, D. (2013) The Future of Peptide-Based
485 Drugs. *Chemical Biology and Drug Design*, **81**, 136–147.
486 <https://doi.org/10.1111/cbdd.12055>.
- 487 10 Vlieghe, P., Lisowski, V., Martinez, J. and Khrestchatisky, M. (2010) Synthetic
488 Therapeutic Peptides: Science and Market. *Drug Discovery Today*, **15**, 40–56.
489 <https://doi.org/10.1016/j.drudis.2009.10.009>.
- 490 11 Mathur, D., Prakash, S., Anand, P., Kaur, H., Agrawal, P., Mehta, A., Kumar, R.,
491 Singh, S. and Raghava, G.P.S. (2016) PEPLife : A Repository of the Half- Life of
492 Peptides. *Scientific Reports*, 1–7.
- 493 12 Usmani, S.S., Bedi, G., Samuel, J.S., Singh, S., Kalra, S., Kumar, P., Ahuja, A.A.,
494 Sharma, M., Gautam, A. and Raghava, G.P.S. (2017) THPdb: Database of FDA-
495 Approved Peptide and Protein Therapeutics. *PLoS ONE*, **12**, 1–12.
496 <https://doi.org/10.1371/journal.pone.0181748>.
- 497 13 Davenport, A.P., Scully, C.C.G., de Graaf, C., Brown, A.J.H. and Maguire, J.J. (2020)
498 Advances in Therapeutic Peptides Targeting G Protein-Coupled Receptors. *Nature*
499 *Reviews Drug Discovery*, **19**, 389–413. <https://doi.org/10.1038/s41573-020-0062-z>.
- 500 14 Tyagi, M., Beghini, F., Poongavanam, V., Doak, B.C. and Kihlberg, J. (2020) Drug
501 Syntheses Beyond the Rule of 5. *Chemistry - A European Journal*, **26**, 49–88.
502 <https://doi.org/10.1002/chem.201902716>.
- 503 15 Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J. (1997) Experimental
504 and Computational Approaches to Estimate Solubility and Permeability in Drug
505 Discovery and Development Settings. *Advanced drug delivery reviews*, 3–25.
- 506 16 Wager, T.T., Hou, X., Verhoest, P.R. and Villalobos, A. (2016) Central Nervous
507 System Multiparameter Optimization Desirability: Application in Drug Discovery.
508 *ACS Chemical Neuroscience*, **7**, 767–775.
509 <https://doi.org/10.1021/acschemneuro.6b00029>.
- 510 17 McNaught, A.D. and A. Wilkinson. (1997) IUPAC. Compendium of Chemical
511 Terminology, 2nd Ed. (the “Gold Book”). . . Blackwell Scientific Publications,

- 512 Oxford. <https://goldbook.iupac.org/terms/view/P04479>.
- 513 18 IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). (1984)
514 Nomenclature and Symbolism for Amino Acids and Peptides: Recommendations
515 1983. *European Journal of Biochemistry*, **138**, 9–37. <https://doi.org/10.1111/j.1432-1033.1984.tb07877.x>.
- 517 19 FDA.Gov
518 [www.fda.gov/drugs/regulatory-science-action/impact-story-developing-tools-evaluate-](http://www.fda.gov/drugs/regulatory-science-action/impact-story-developing-tools-evaluate-complex-drug-products-peptides)
519 [complex-drug-products-peptides](http://www.fda.gov/drugs/regulatory-science-action/impact-story-developing-tools-evaluate-complex-drug-products-peptides)
520 [Accessed August 10, 2020]
- 521 20 Wu, L.C. (2013) Regulatory Considerations for Peptide Drug Products. 1–30.
- 522 21 Uhlig, T., Kyprianou, T., Martinelli, F.G., Oppici, C.A., Heiligers, D., Hills, D., Calvo,
523 X.R. and Verhaert, P. (2014) The Emergence of Peptides in the Pharmaceutical
524 Business: From Exploration to Exploitation. *EuPA Open Proteomics*, **4**, 58–69.
525 <https://doi.org/10.1016/j.euprot.2014.05.003>.
- 526 22 Go.DrugBank
527 <https://go.drugbank.com/>
528 [Accessed May 20, 2020]
- 529 23 EMA-European Medicines Agency
530 www.ema.europa.eu/en
531 [Accessed May 20, 2020]
- 532 24 PMDA - Pharmaceuticals and Medical Devices Agency
533 www.pmda.go.jp/english/index.html
534 [Accessed May 10, 2020]
- 535 25 Drug Central
536 <http://drugcentral.org>
537 [Accessed August 20, 2020]
- 538 26 National Center for Advancing Translational Sciences NIH
539 <https://drugs.ncats.io>
540 [Accessed August 20, 2020]
- 541 27 Drugs.Com
542 <https://www.drugs.com>
543 [Accessed September 1, 2020]
- 544 28 Lehninger, A.L., Cox, M.M. and Nelson, D.L. (2000) *Lehninger Principles of*
545 *Biochemistry*. 6th ed., New York: Worth Publisher.
- 546 29 Walther, F.J., Gordon, L.M. and Waring, A.J. (2016) Design of Surfactant Protein B
547 Peptide Mimics Based on the Saposin Fold for Synthetic Lung Surfactants.
548 *Biomedicine Hub*, **1**, 1–21. <https://doi.org/10.1159/000451076>.
- 549 30 Yin, N., Brimble, M.A., Harris, P.W. and Wen, J. (2014) Enhancing the Oral
550 Bioavailability of Peptide Drugs by Using Chemical Modification and Other

- 551 Approaches. *Medicinal Chemistry*, **4**, 763–769. <https://doi.org/10.4172/2161->
552 0444.1000227.
- 553 31 Chanson, P., Tiit, J. and Harris, A.G. (1993) Clinical Pharmacokinetics of Octreotide:
554 Therapeutic Applications in Patients with Pituitary Tumours. *Clinical*
555 *Pharmacokinetics*, *Clin Pharmacokinet*, 375–391. <https://doi.org/10.2165/00003088->
556 199325050-00004.
- 557 32 O’Toole, T.J. and Sharma, S. (2019) Physiology, Somatostatin. StatPearls, StatPearls
558 Publishing. <http://www.ncbi.nlm.nih.gov/pubmed/30855911>.
- 559 33 Hodgson, D.R.W. and Sanderson, J.M. (2004) The Synthesis of Peptides and Proteins
560 Containing Non-Natural Amino Acids. *Chemical Society Reviews*, Royal Society of
561 Chemistry, **33**, 422–430. <https://doi.org/10.1039/b312953p>.
- 562 34 Hohsaka, T. and Sisido, M. (2002, December 1) Incorporation of Non-Natural Amino
563 Acids into Proteins. *Current Opinion in Chemical Biology*, Elsevier Ltd, 809–815.
564 [https://doi.org/10.1016/S1367-5931\(02\)00376-9](https://doi.org/10.1016/S1367-5931(02)00376-9).
- 565 35 Pollegioni, L. and Servi, S. (2012) Unnatural Amino Acids. *Methods in Molecular*
566 *Biology*, Press, Humana.
- 567 36 Mant, C.T., Jiang, Z., Gera, L., Davis, T., Nelson, K.L., Bevers, S. and Hodges, R.S.
568 (2019) De Novo Designed Amphipathic α -Helical Antimicrobial Peptides
569 Incorporating Dab and Dap Residues on the Polar Face to Treat the Gram-Negative
570 Pathogen, *Acinetobacter Baumannii*. *Journal of Medicinal Chemistry*, American
571 Chemical Society, **62**, 3354–3366. <https://doi.org/10.1021/acs.jmedchem.8b01785>.
- 572 37 Lu, J., Xu, H., Xia, J., Ma, J., Xu, J., Li, Y. and Feng, J. (2020) D- and Unnatural
573 Amino Acid Substituted Antimicrobial Peptides With Improved Proteolytic Resistance
574 and Their Proteolytic Degradation Characteristics. *Frontiers in Microbiology*,
575 Frontiers Media S.A., **11**, 563030. <https://doi.org/10.3389/fmicb.2020.563030>.
- 576 38 Witkowska, E., Orłowska, A., Sagan, B., Smoluch, M. and Izdebski, J. (2001) Tryptic
577 Hydrolysis of HGH-RH(1-29)-NH₂ Analogues Containing Lys or Orn in Positions 12
578 and 21. *Journal of Peptide Science*, *J Pept Sci*, **7**, 166–172.
579 <https://doi.org/10.1002/psc.316>.
- 580 39 Inouy, K., Sakakibara, S. and Prockop, D.J. (1976) Effects of the Stereo-Configuration
581 of the Hydroxyl Group in 4-Hydroxyproline on the Triple-Helical Structures Formed
582 by Homogeneous Peptides Resembling Collagen. *BBA - Protein Structure*, Elsevier,
583 **420**, 133–141. [https://doi.org/10.1016/0005-2795\(76\)90352-4](https://doi.org/10.1016/0005-2795(76)90352-4).
- 584 40 Meyer, D., Mutschler, C., Robertson, I., Batt, A. and Tatko, C. (2013) Aromatic
585 Interactions with Naphthylalanine in a β -Hairpin Peptide. *Journal of Peptide Science*,
586 John Wiley & Sons, Ltd, **19**, 277–282. <https://doi.org/10.1002/psc.2496>.
- 587 41 Sidransky, H., Verney, E. and Kurl, R. (1990) Comparison of Effects of L-Tryptophan
588 and a Tryptophan Analog, D,L- β -(1-Naphthyl)Alanine, on Processes Relating to
589 Hepatic Protein Synthesis in Rats. *Journal of Nutrition*, Oxford Academic, **120**, 1157–
590 1162. <https://doi.org/10.1093/jn/120.10.1157>.

- 591 42 Sidransky, H. and Verney, E. (1996) Influence of L-Alanine on Effects Induced by L-
592 Tryptophan on Rat Liver. *Journal of Nutritional Biochemistry*, Elsevier Inc., **7**, 200–
593 206. [https://doi.org/10.1016/0955-2863\(96\)00010-1](https://doi.org/10.1016/0955-2863(96)00010-1).
- 594 43 Rodriguez, M., Bernad, N., Galas, M., Lignon, M., Laur, J., Aumelas, A. and
595 Martinez, J. (1991) Synthesis and Biological Activities of Cholecystokinin Analogues
596 Substituted in Position 30 by 3-(1-Naphthyl)-l-Alanine [Nal(1)] or 3-(2-Naphthyl)-l-
597 Alanine [Nal(2)]. *European Journal of Medicinal Chemistry*, Elsevier Masson, **26**,
598 245–253. [https://doi.org/10.1016/0223-5234\(91\)90056-S](https://doi.org/10.1016/0223-5234(91)90056-S).
- 599 44 Zhang, L. and Bulaj, G. (2012) Converting Peptides into Drug Leads by Lipidation.
600 *Current Medicinal Chemistry*, Bentham Science Publishers Ltd., **19**, 1602–1618.
601 <https://doi.org/10.2174/092986712799945003>.
- 602 45 Moradi, S.V., Hussein, W.M., Varamini, P., Simerska, P. and Toth, I. (2016, April 1)
603 Glycosylation, an Effective Synthetic Strategy to Improve the Bioavailability of
604 Therapeutic Peptides. *Chemical Science*, Royal Society of Chemistry, 2492–2500.
605 <https://doi.org/10.1039/c5sc04392a>.
- 606 46 Kang, R., Wan, J., Arstikaitis, P., Takahashi, H., Huang, K., Bailey, A.O., Thompson,
607 J.X., Roth, A.F., Drisdell, R.C., Mastro, R., Green, W.N., Yates 3rd, J.R., Davis, N.G.
608 and El-Husseini, A. (2008) Neural Palmitoyl-Proteomics Reveals Dynamic Synaptic
609 Palmitoylation HHS Public Access. *Nature*, **456**, 904–909.
610 <https://doi.org/10.1038/nature07605>.
- 611 47 Ward, B.P., Ottaway, N.L., Perez-Tilve, D., Ma, D., Gelfanov, V.M., Tschöp, M.H.
612 and DiMarchi, R.D. (2013) Peptide Lipidation Stabilizes Structure to Enhance
613 Biological Function. *Molecular Metabolism*, **2**, 468–479.
614 <https://doi.org/10.1016/j.molmet.2013.08.008>.
- 615 48 DeFrees, S. (2007, July 23) Glycosylation of Peptides via O-Linked Glycosylation
616 Sequences.
- 617 49 Akhilesh Kumar and Anand K. Bachhawat. (2012) Pyroglutamic Acid: Throwing
618 Light on a Lightly Studied Metabolite on JSTOR. *Current Science*, 288–297.
619 <https://www.jstor.org/stable/24083854?seq=1>.
- 620 50 Garden, R.W., Moroz, T.P., Gleeson, J.M., Floyd, P.D., Li, L., Rubakhin, S.S. and
621 Sweedler, J. V. (1999) Formation of N-Pyroglutamyl Peptides from N-Glu and N-Gln
622 Precursors in Aplysia Neurons. *Journal of Neurochemistry*, John Wiley & Sons, Ltd,
623 **72**, 676–681. <https://doi.org/10.1046/j.1471-4159.1999.0720676.x>.
- 624 51 Kehlen, A., Haegele, M., Böhme, L., Cynis, H., Hoffmann, T. and Demuth, H.U.
625 (2017) N-Terminal Pyroglutamate Formation in CX3CL1 Is Essential for Its Full
626 Biologic Activity. *Bioscience Reports*, Portland Press Ltd, **37**, 20170712.
627 <https://doi.org/10.1042/BSR20170712>.
- 628 52 Shah, P. and Westwell, A.D. (2007) The Role of Fluorine in Medicinal Chemistry.
629 *Journal of Enzyme Inhibition and Medicinal Chemistry*, Taylor & Francis, **22**, 527–
630 540. <https://doi.org/10.1080/14756360701425014>.
- 631 53 Marsh, E.N.G., Buer, B.C. and Ramamoorthy, A. (2009) Fluorine - A New Element in

- 632 the Design of Membrane-Active Peptides. *Molecular BioSystems*, **5**, 1143–1147.
633 <https://doi.org/10.1039/b909864j>.
- 634 54 Meng, H., Krishnaji, S.T., Beinborn, M. and Kumar, K. (2008) Influence of Selective
635 Fluorination on the Biological Activity and Proteolytic Stability of Glucagon-like
636 Peptide-1. *Journal of Medicinal Chemistry*, **51**, 7303–7307.
637 <https://doi.org/10.1021/jm8008579>.
- 638 55 Chemicalize
639 <https://chemicalize.com/>
640 [Accessed September 4, 2020]
- 641 56 Faulds, D., Goa, K.L. and Benfield, P. (1993) Cyclosporin: A Review of Its
642 Pharmacodynamic and Pharmacokinetic Properties, and Therapeutic Use in
643 Immunoregulatory Disorders. *Drugs*, **45**, 953–1040. [https://doi.org/10.2165/00003495-](https://doi.org/10.2165/00003495-199345060-00007)
644 [199345060-00007](https://doi.org/10.2165/00003495-199345060-00007).
- 645 57 Novartis. (1988) Sandimmune. Intensive Therapy and Clinical Monitoring, 172.
- 646 58 Thorstholm, L. and Craik, D.J. (2012) Discovery and Applications of Naturally
647 Occurring Cyclic Peptides. *Drug Discovery Today: Technologies*, Elsevier Ltd, **9**,
648 e13–e21. <https://doi.org/10.1016/j.ddtec.2011.07.005>.
- 649 59 Qvit, N., Rubin, S.J.S., Urban, T.J., Mochly-Rosen, D. and Gross, E.R. (2017)
650 Peptidomimetic Therapeutics: Scientific Approaches and Opportunities. *Drug*
651 *Discovery Today*. <https://doi.org/10.1016/j.drudis.2016.11.003>.
- 652 60 Hill, T.A., Shepherd, N.E., Diness, F. and Fairlie, D.P. (2014) Constraining Cyclic
653 Peptides to Mimic Protein Structure Motifs. *Angewandte Chemie - International*
654 *Edition*, **53**, 13020–13041. <https://doi.org/10.1002/anie.201401058>.
- 655 61 Lenci, E. and Trabocchi, A. (2020) Peptidomimetic Toolbox for Drug Discovery.
656 *Chemical Society Reviews*, **49**, 3262–3277. <https://doi.org/10.1039/d0cs00102c>.
- 657 62 Lambert, J.N., Mitchell, J.P. and Roberts, K.D. (2001) The Synthesis of Cyclic
658 Peptides. *Journal of the Chemical Society. Perkin Transactions 1*, 471–484.
659 <https://doi.org/10.1039/b001942i>.
- 660 63 Tapeinou, A., Matsoukas, M.-T., Simal, C. and Tselios, T. (2015) Review Cyclic
661 Peptides on a Merry-Go-Round; towards Drug Design. *Biopolymers*, John Wiley &
662 Sons, Ltd, **104**, 453–461. <https://doi.org/10.1002/bip.22669>.
- 663 64 Botti, P., Pallin, T.D. and Tam, J.P. (1996) Cyclic Peptides from Linear Unprotected
664 Peptide Precursors through Thiazolidine Formation. *Journal of the American Chemical*
665 *Society*, **118**, 10018–10024. <https://doi.org/10.1021/ja954278g>.
- 666 65 Jing, X. and Jin, K. (2020) A Gold Mine for Drug Discovery: Strategies to Develop
667 Cyclic Peptides into Therapies. *Medicinal Research Reviews*, **40**, 753–810.
668 <https://doi.org/10.1002/med.21639>.
- 669 66 Wallace, R.J. (1992) Acetylation of Peptides Inhibits Their Degradation by Rumen
670 Micro-Organisms. *British Journal of Nutrition*, Cambridge University Press (CUP),

