

Mgr Katarzyna Witoszka

**Synteza bibliotek pochodnych fenyloalaniny
i tryptofanu oraz ich wykorzystanie do opracowania
funkcji oceniającej specyficznnej dla receptora
tachykininowego NK1**

Rozprawa na stopień naukowy doktora
w dziedzinie nauk medycznych i nauk o zdrowiu
w dyscyplinie nauki medyczne

Promotor: prof. dr hab. Aleksandra Misicka-Kęsik

Promotor: dr hab. Piotr F.J. Lipiński



Postępowanie w sprawie nadania stopnia doktora
w Instytucie Medycyny Doświadczalnej i Klinicznej
im. Mirosława Mossakowskiego PAN

Warszawa, 2024

Podziękowania

Szczególne podziękowania kieruję do mojej Promotor, **prof. dr hab. Aleksandry Misickiej-Kęsik**. Jestem ogromnie wdzięczna za zaufanie, jakim zostałam obdarzona, umożliwienie mi realizacji tej pracy doktorskiej oraz wsparcie w moim rozwoju naukowym przez wiele lat. Dziękuję za cenne wskazówki i uwagi merytoryczne, które znacząco przyczyniły się do jakości i wartości mojej pracy, motywację, wyrozumiałość i życzliwość oraz nieślabnący entuzjazm.

Najserdeczniejsze wyrazy podziękowania składam również mojemu drugiemu Promotorowi, **dr. hab. Piotrowi F.J. Lipińskiemu**, bez którego zaangażowania ta praca by nie powstała. Dziękuję za poświęcony czas, nieocenione wsparcie przy badaniach i przygotowaniu tej rozprawy, inspirujące rozmowy oraz cenne uwagi merytoryczne. Jestem niezmiernie wdzięczna za przekazaną wiedzę, niezliczone pokłady cierpliwości, życzliwości i wyrozumiałości oraz wiarę w sukces tego projektu.

Osobne podziękowania kieruję do Promotor moich prac licencjackiej i magisterskiej, **dr Ewy Witkowskiej**, za wprowadzenie mnie w świat chemii peptydów, przekazaną wiedzę, ukształtowanie mojej pasji do pracy naukowej i warsztatu pracy laboratoryjnej.

Serdecznie dziękuję Członkom Zakładu Neuropeptydów, w szczególności **dr Jolancie Dyniewicz** i **dr Patrycji Redkiewicz** oraz **dr. Piotrowi Kossonowi** za pomoc w rozwiązywaniu problemów naukowych i nie tylko, za troskę, wsparcie, motywujące rozmowy oraz stworzenie milej atmosfery pracy.

Dziękuję **dr Joannie Matalińskiej** (ZN IMDiK PAN) za wykonanie ewaluacji biologicznej zsyntezowanych przeze mnie związków. Dziękuję również **dr. Krzysztofowi Łyczko** (ICH TJ) za wykonanie pomiarów strukturalnych dla moich związków.

Z głębi serca dziękuję **dr Beacie Wileńskiej**, **dr Joannie Turkowskiej**, **dr. Adrianowi Konopko**, **Annie Marczyk**, **Magdzie Popławskiej** za wszelaką pomoc, dzielenie się wiedzą i doświadczeniem chemicznym, inspirujące rozmowy, motywację i zrozumienie.

Najgłębszą wdzięczność składam moim **Rodzicom**, za poświęcenie, troskę, zrozumienie, nieustanne wsparcie na każdym etapie mojego życia, możliwość rozwijania pasji naukowych, za wszystko. Dziękuję również mojemu **Rodzeństwu**, **Rodzinie** i **Przyjaciołom** za wyrozumiałość, słowa otuchy w trudnych momentach oraz wiarę w moje możliwości.

Największe podziękowania kieruję do mojego **Męża Damiana**, bez którego pomocy, wiary we mnie, wyrozumiałości, cierpliwości, nieustającego wsparcia i miłości ukończenie tej pracy doktorskiej byłoby niemożliwe. Twoja obecność była dla mnie niezastąpiona.

Niektóre z wyników przedstawionych w prezentowanej rozprawie zostały opublikowane w pracach [1] i [2]:

[1] Matalińska, J.¹; **Kosińska, K.^{1*}**; Halik, P.K.; Koźmiński, P.; Lipiński, P.F.J.; Gniazdowska, E.; Misicka, A. Novel NK1R-Targeted 68Ga-/177Lu-Radioconjugates with Potential Application against Glioblastoma Multiforme: Preliminary Exploration of Structure–Activity Relationships. *Int. J. Mol. Sci.* 2022, 23, 1214.

[2] **Witoszka, K.¹**; Matalińska, J.; Misicka, A.; Lipiński, P.F.J. Moving out of CF₃-Land: Synthesis, Receptor Affinity, and in silico Studies of NK1 Receptor Ligands Containing a Pentafluorosulfanyl (SF₅) Group. *ChemMedChem* 2023, 18.

oraz w patencie [3]:

[3] Lipiński, P.F.J.; Matalińska, J.; **Kosińska, K.^{*}**; Dyniewicz, J.; Misicka, A. Patent PL244031B1: Nowe pochodne peptydowe, zwłaszcza stosowania w terapii nowotworów 2023.

Dodatkowo, wybrane związki były przedmiotem szczegółowych badań strukturalnych (struktura krystalograficzna) i spektroskopowych (widma VCD w ciele stałym), wykonanych przez współpracowników i opisanych w publikacji [4]:

[4] Rode, J.E.¹; Lyczko, K.¹; **Kosińska, K.^{*}**; Matalińska, J.; Dyniewicz, J.; Misicka, A.; Dobrowolski, J.C.; Lipiński, P.F.J. The solid state VCD of a novel N-acylhydrazono trifluoroacetate. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 2022, 269, 120761.

Przeprowadzone badania finansowano ze środków:

1. Narodowego Centrum Nauki w ramach grantów 2016/23/D/NZ7/0363 oraz 2017/25/B/NZ7/01896
2. Programu Operacyjnego Wiedza Edukacja Rozwój 2014-2020 współfinansowanego ze środków Europejskiego Funduszu Społecznego w ramach grantu POWR.03.02.00-00- 1028/17-00.

* Kosińska – nazwisko panieńskie

Streszczenie

Receptor tachykininowy NK1 (NK1R) wraz ze swoim endogennym ligandem, substancją P (SP), uczestniczy w regulacji wielu procesów fizjologicznych i patologicznych. Ten szeroki udział w fizjologii człowieka, czyni receptor NK1 ważnym terapeutycznym celem molekularnym. Nowsze podejścia do wykorzystania NK1R jako punktu uchwytu dla terapii (np. wielocelowe analgetyki; celowane dostarczanie ładunku terapeutycznego, celowana terapia radionuklidowa, antagonizm endosomowego sygnałowania przez NK1R) wymaga poszukiwania nowych ligandów tego receptora lub stosownych modyfikacji związków już istniejących.

Celem niniejszej pracy doktorskiej było zaprojektowanie i synteza biblioteki nowych pochodnych fenyloalaniny i tryptofanu jako potencjalnych ligandów NK1R. W tym kontekście, zaprojektowano serie związków, które a) zawierały kluczowe dla powinowactwa do NK1R elementy osadzone na różnych szkieletach organicznych (serie W/F-71–W/F-74), b) stanowiły rozbudowanie znanego antagonisty NK1R L732,138 (serie W-80–W-84) lub c) były jego modyfikacją (seria W-90). Zaprojektowane związki otrzymano w ramach kilkietapowych syntez, stosując znane techniki syntezy organicznej oraz syntezy peptydów w roztworze. Otrzymano 94 związki, w tym 93 nowe pochodne.

Otrzymane związki poddano ewaluacji biologicznej (przeprowadzonej przez współpracowników). Badania powinowactwa receptorowego wobec NK1R posłużyły analizie zależności struktura–aktywność w otrzymanych bibliotekach w kontekście receptora NK1. Pomimo uwzględnienia w strukturze związków z serii W/F-70–W/F-74 kluczowych dla powinowactwa do receptora NK1 elementów, związki te nie wykazywały znaczącej aktywności wobec NK1R. W ramach przeprowadzonych badań powinowactwa wykazano, że możliwe jest rozbudowanie (na sposób umożliwiający tworzenie koniugatów i związków wielofunkcyjnych) znanego antagonisty NK1R L-732,138 o dodatkowe elementy od N-końca bez utraty powinowactwa do NK1R (seria związków W-80–W-84). Dodatkowo, na podstawie analizy zależności struktura–aktywność w serii związków W-90 stwierdzono, że możliwe jest zastąpienie w strukturze L-732,138 kluczowej dla powinowactwa grupy CF₃ grupą pentafluorosulfanylową (SF₅, nazywaną też „super-CF₃”) bez znaczącego spadku powinowactwa.

Innym rodzajem badania *in vitro*, w którym testowana była seria związków W/F-70–W/F-74 było badanie wpływu na żywotność komórek (cytotoksyczność). Przebadane związki cechowały się zróżnicowaną aktywnością wobec komórek nowotworowych COLO679 i prawidłowych BJ. Spośród nich dwa związki charakteryzowały się szczególnym profilem aktywności, wykazując silne działanie cytotoksyczne i selektywne wobec linii komórkowej COLO679, jednocześnie nie wykazując znaczącego powinowactwa wobec NK1R. Korzystając z metod *in silico* (przewidywania SwissTarget oraz dokowanie molekularne) opracowano hipotezę sugerującą możliwość potencjalnego działania tych związków przez receptor CCK2.

Niezależnym celem mojej pracy było opracowanie funkcji oceniających specyficznych dla receptora NK1. W dzisiejszych czasach, kluczową rolę w poszukiwaniu nowych substancji czynnych pełnią metody *in silico*, w tym dokowanie molekularne. Funkcje oceniające w dokowaniu molekularnym służą do prognozowania siły wiązania. Niestety, w wielu przypadkach dokowanie, szczególnie pod względem przewidywania energii wiązania, jest niedokładne, co utrudnia praktyczne zastosowanie tych technik. Aby poprawić skuteczność dokowania, wykorzystuje się różnorodne strategie, w tym opracowywanie funkcji oceniających specyficznych dla konkretnych celów molekularnych.

W swoich badaniach obliczeniowych, w pierwszym etapie, wykazałam, że oryginalna funkcja oceniająca programu AutoDock Vina stosunkowo słabo radzi sobie w szacowaniu energii wiązania ligandów z NK1R i odróżnianiu związków aktywnych od nieaktywnych. Kolejno, uzyskane dane eksperymentalne o powinowactwie receptorowym do NK1R nowosyntezyzowanych związków oraz dane literaturowe wykorzystano do wytrenowania funkcji specyficznych dla NK1R. Otrzymano nowe funkcje oceniające, które w przesiewie wirtualnym były skuteczniejsze niż oryginalna funkcja oceniająca programu AutoDock Vina.

Abstract

The neurokinin NK1 receptor (NK1R) along with its endogenous ligand, substance P (SP), takes part in the regulation of many physiological and pathological processes. This vast participation in the human physiology renders the NK1 receptor an important molecular target for therapies. Newer approaches to the application of NK1R as a therapeutic target (*e.g.*, multitarget analgesics; targeted delivery of therapeutic cargo, targeted radionuclide therapy, antagonism of NK1R endosome signalling) require searching for new ligands of this receptor or appropriate modifications of the already existing compounds.

The aim of this doctoral project was the design and synthesis of a library of novel phenylalanine and tryptophan derivatives as potential ligands of NK1R. In this context, designed were a few series of compounds a) that contained structural elements important for the NK1R affinity embedded upon various organic scaffolds (W/F-71–W/F-74 series), or b) that were expanded analogues of the known antagonist NK1R L732,138 (W-80–W-84 series) or c) that were modification of that antagonist (W-90 series). The designed compounds were obtained via a-few-steps synthetic routes, using typical techniques of organic synthesis and those of peptide synthesis in solution. In total, 94 compounds were obtained, including 93 novel derivatives.

The synthesized compounds were evaluated biologically (with the help of collaborators). Data on the NK1R receptor affinity served for the analysis of the structure-activity relationships in the obtained libraries. Despite including structural elements crucial for the NK1R affinity, the compounds of W/F-70–W/F-74 series did not exhibit any significant activity towards NK1R. Furthermore, it was found that it is possible to expand (in a way which enables the formation of conjugates and multifunctional compounds) the known NK1R antagonist, called L-732,138, by including additional elements from as the N-terminal piece, without losing NK1R affinity (W-80–W-84 series). Moreover, on the basis of the analysis of the structure-activity relationships in the compound series W-90, it was established that it is possible to substitute the trifluoromethyl (CF₃) group, essential for the NK1R affinity, in the L-732,138 structure with a pentafluorosulphonyl group (SF₅, also called „the super-CF₃”) without significant affinity decreases.

Another type of *in vitro* studies, in which the compounds of W/F-70–W/F-74 series were tested, was the study of influence on the cell viability (cytotoxicity). The investigated

compounds were characterized by varied activity towards COLO679 cancer cells and normal BJ cells. Among them, two compounds displayed a particular activity profile, exhibiting strong cytotoxic activity and selectivity towards the COLO679 cell line, simultaneously not exhibiting a significant affinity towards NK1R. Based on the *in silico* methods (SwissTarget predictions and molecular docking) a hypothesis was developed suggesting the possibility that these compounds are active *via* CCK2 receptor.

An independent goal of my project was the development of scoring functions specific for NK1R. Nowadays, a key role in searching for new active substances is played by *in silico* methods, including molecular docking. Scoring functions in molecular docking are used for predicting the free energy of binding (binding strength). Unfortunately, in many cases docking is inaccurate (especially when it comes to predicting binding strength) which makes the practical application of these techniques less efficient. In order to improve the effectiveness of docking, various strategies are employed, including the development of scoring functions specific for particular molecular targets.

In the first stage of my computational work, I found that the original scoring function of AutoDock Vina program has a relatively poor efficacy in evaluating the binding energies of NK1R ligands and distinguishing active and nonactive compounds. Next, the obtained experimental data on NK1R affinity of my newly synthesized compounds as well as the literature data were used to train scoring functions specific for NK1R. The new scoring turned out to be more effective in virtual screening compared to the original scoring function of the AutoDock Vina program.

Innowacyjność rozprawy

- 1) Otrzymanie biblioteki nowych 64 pochodnych fenyloalaniny i tryptofanu, jako potencjalnych ligandów NK1R.
- 2) Otrzymanie 14 nowych pochodnych tryptofanu poprzez rozbudowanie (na sposób umożliwiający tworzenie koniugatów i związków wielofunkcyjnych) o dodatkowe elementy od N-końca znanego antagonisty NK1R L-732,138 bez utraty powinowactwa do NK1R.
- 3) Otrzymanie 10 nowych pochodnych tryptofanu (analogów L-732,138), w których kluczową dla powinowactwa grupę trifluorometylową (CF₃) zastąpiono grupą pentafluorosulfanylową (SF₅), bez znaczącego spadku powinowactwa.
- 4) Wykazanie umiarkowanej skuteczności funkcji oceniającej programu AutoDock Vina w przesiewaniu wobec receptora NK1.
- 5) Opracowanie nowych funkcji oceniających (do *rescoringu*) specyficznych dla NK1R, które w testach walidacyjnych wykazywały systematyczną i statystycznie istotną poprawę skuteczności w porównaniu do oryginalnej funkcji oceniającej programu AutoDock Vina.

Spis treści

Podziękowania.....	I
Streszczenie.....	V
Abstract.....	VII
Innowacyjność rozprawy.....	IX
Spis treści	1
Spis rysunków.....	4
Spis schematów.....	5
Spis tabel.....	6
Wykaz skrótów.....	10
I. Wstęp.....	13
II. Część literaturowa	15
II.1. Substancja P i receptor NK1	15
II.1.1. Rodzina peptydów tachykininowych.....	15
II.1.2. Receptor NK1	16
II.1.3. Rola SP i NK1R w procesach fizjologicznych i patologicznych	23
II.2. Chemia medyczna receptora NK1	27
II.2.1. Receptor NK1R jako cel molekularny terapii	27
II.2.2. Chemotypy antagonistów NK1R.....	28
II.2.2.1. Peptydowe antagonisty receptora NK1	29
II.2.2.2. Małocząsteczkowe antagonisty NK1R.....	32
II.2.2.2.1 Imidazochinoksaliny i heterosteroidy.....	32
II.2.2.2.2 Perhydroizoindolony	33
II.2.2.2.3 3-(<i>N</i> -benzyloamino)chinuklidyny	34
II.2.2.2.4 3-(<i>N</i> -benzyloamino)piperydyny	36
II.2.2.2.5 Etery benzytowe chinuklidyny, piperydyny oraz morfoliny	37
II.2.2.2.6 Antagonisty oparte na strukturze piperydyny.....	39
II.2.2.2.7 Pochodne tryptofanu.....	41
II.2.3. Nowsze koncepcje w chemii medycznej receptora NK1	42
II.2.3.1. Związki hybrydowe działające przez receptory opioidowe i NK1R.....	43
II.2.3.2. Antagonizm endosomalnego sygnałowania przez NK1R	46
II.2.3.3. Antagonisty NK1R o działaniu przeciwnowotworowym.....	49
II.2.3.4. Receptor NK1R w radiofarmacji	52
II.2.3.5. Receptor NK1R w celowanym dostarczaniu elementu (cargo) terapeutycznego (innego niż radionuklid).....	55
II.3. Funkcje oceniające.....	59

II.3.1.	Dokowanie molekularne.....	59
II.3.1.1.	Rola dokowania we współczesnym poszukiwaniu leków.....	59
II.3.1.2.	Proces dokowania.....	60
II.3.2.	Funkcje oceniające.....	63
II.3.3.	Funkcje oceniające specyficzne dla celu molekularnego.....	68
Badania własne.....		77
III.	Cel pracy.....	77
IV.	Wyniki.....	78
IV.1.	Projektowanie związków.....	78
IV.2.	Syntezy.....	83
IV.3.	Aktywność biologiczna.....	91
IV.3.1.	Powinowactwo do receptora NK1R.....	91
IV.3.1.1.	Serie W/F-70–W/F-74.....	93
IV.3.1.2.	Serie W-80 – W-84.....	94
IV.3.1.3.	Seria W-90.....	95
IV.3.2.	Oddziaływanie W-90 z NK1R.....	96
IV.3.3.	Cytotoksyczność.....	100
IV.3.4.	Hipoteza celu molekularnego dla związków z serii W-71.....	103
IV.4.	Badanie skuteczności funkcji oceniającej programu AutoDock Vina.....	107
IV.5.	Opracowanie funkcji oceniających specyficznych dla receptora NK1.....	118
V.	Dyskusja.....	152
VI.	Wnioski.....	163
VII.	Materiały i metody.....	165
VII.1.	Chemia: uwagi ogólne.....	165
VII.2.	Procedury syntetyczne.....	166
VII.2.1.	Otrzymywanie hydrazydów (Procedura A).....	166
VII.2.2.	Sprzęganie grupy karboksylowej i aminowej (Procedura B).....	167
VII.2.3.	Sprzęganie grupy karboksylowej i aminowej z wykorzystaniem preformowanego estru aktywnego (Procedura C).....	167
VII.2.4.	Otrzymywanie hydrazonów (Procedura D).....	167
VII.2.5.	Usuwanie grupy ochronnej Boc (Procedura E).....	168
VII.2.6.	Acetylowanie wolnej grupy aminowej (Procedura F).....	168
VII.2.7.	Alkilowanie bromkami benzyłowymi (Procedura G).....	168
VII.3.	Syntezy związków.....	169
VII.3.1.	Seria W-71.....	169
VII.3.2.	Seria F-71.....	173
VII.3.3.	Seria W-72.....	177

VII.3.4.	Seria F-72	181
VII.3.5.	Seria W-73	184
VII.3.6.	Seria F-73	187
VII.3.7.	Seria W-74	190
VII.3.8.	Seria F-74	194
VII.3.9.	Seria W-80	197
VII.3.10.	Seria W-81	199
VII.3.11.	Seria W-82	201
VII.3.12.	Seria W-83	204
VII.3.13.	Seria W-84	206
VII.3.14.	Seria W-90	208
VII.4.	Dokowanie do receptora NK1	213
VII.5.	Zestawy danych do oceny skuteczności i trenowania funkcji oceniających	214
VII.6.	Przygotowanie ligandów z bazy ChEMBL do dokowania	215
VII.7.	Parametry oceny skuteczności przesiewu wirtualnego	215
VII.8.	Trenowanie funkcji oceniających	217
VII.9.	Predykcja celów molekularnych za pomocą metody SwissTarget	218
VII.10.	Dokowanie do receptora CCK2	219
VII.11.	Badania przeprowadzone przez współpracowników	219
VIII.	Literatura.....	221

Spis rysunków

Rysunek 1. Schematyczny model budowy receptora NK1.....	18
Rysunek 2. Szlaki sygnałowe związane z receptorem NK1.....	20
Rysunek 3. Wybrane peptydowe/peptydomimetyczne antagonisty NK1R.....	32
Rysunek 4. Wybrane antagonisty NK1R oparte na strukturze imidazochinoksaliny i heterosteroidów.....	33
Rysunek 5. Wybrane antagonisty NK1R oparte na strukturze perhydroizoindolonu.....	34
Rysunek 6. Wybrane antagonisty NK1R oparte na strukturze 3-(N-benzyloamino)chinuklidyny.....	35
Rysunek 7. Wybrane antagonisty NK1R oparte na strukturze 3-(N-benzyloamino)piperydyny.....	37
Rysunek 8. Wybrane antagonisty NK1R oparte na strukturze eterów benzytowych chinuklidyny, piperydyny oraz morfoliny.....	39
Rysunek 9. Wybrane antagonisty NK1R oparte na strukturze piperydyny.....	41
Rysunek 10. Wybrane antagonisty NK1R oparte na strukturze tryptofanu.....	42
Rysunek 11. Wybrane cząsteczki oparte o antagonistów NK1R w endosomalnym sygnałowaniu.....	49
Rysunek 12. Wybrane przykłady z serii pochodnych 5-arylotio-2-amino-4H-piranów.....	51
Rysunek 13. Nowe antagonisty NK1R oparte na strukturze węglowodanów.....	51
Rysunek 14. Wybrane radiofarmaceutyki oparte na strukturze antagonisty NK1R.....	55
Rysunek 15. Struktury niektórych koniugatów lub nośników opartych na strukturze SP.....	56
Rysunek 16. Struktury chemiczne niektórych znanych antagonistów NK1R.....	78
Rysunek 17. Wzór ogólny związków serii 71-74.....	79
Rysunek 18. Wzór ogólny serii związków W-80–W-84.....	81
Rysunek 19. Wzór ogólny serii związków W-90.....	82
Rysunek 20. Struktura krystaliczna związku H-F-71 (CCDC2096461).....	85
Rysunek 21. Struktura krystaliczna związku Boc-F-74.....	89
Rysunek 22. Wyniki dokowania do NK1R dla związku Ac-W-90-Br.....	99
Rysunek 23. Wyniki dokowania do CCK2R dla związku Z-dW-71.....	106

Spis schematów

Schemat 1. Schemat otrzymywania związków z serii 71.	84
Schemat 2. Schemat otrzymywania związków z serii 72.	86
Schemat 3. Schemat otrzymywania związków z serii 73.	87
Schemat 4. Schemat otrzymywania związków z serii 74.	88
Schemat 5. Schemat otrzymywania związków z serii W-80–84.	90
Schemat 6. Schemat otrzymywania związków z serii W-90.	91

Spis tabel

Tabela 1. Przykładowe peptydy tachykininowe i ich sekwencje.....	16
Tabela 2. Sekwencje wybranych peptydowych antagonistów NK1R w porównaniu do Substancji P, na podstawie [82].....	29
Tabela 3. Wagi poszczególnych wkładów energetycznych do siły oddziaływania w modelu domyślnej funkcji oceniającej programu AutoDock Vina.....	66
Tabela 4. Struktury związków 1–64.	80
Tabela 5. Struktury związków 65–79.	82
Tabela 6. Struktury związków 80–94.	83
Tabela 7. Powinowactwo zsyntezowanych związków do receptora NK1 (IC_{50}).	92
Tabela 8. Badanie wpływu związków z serii W/F-71 – W/F-74 na żywotność (cytotoksyczność) linii komórkowych COLO679 oraz BJ (IC_{50}).....	102
Tabela 9. Wybrane związki, na podstawie których SwissTarget przewiduje prawdopodobieństwo oddziaływania Z-dW-71 z receptorem CCK2.....	105
Tabela 10. Scenariusze eksperymentu wirtualnego przesiewu.....	109
Tabela 11. Średnie wartości parametru logAUC i rozkład wartości EF_{frac} dla dokowań w Scenariuszu A.	112
Tabela 12. Średnie wartości parametru logAUC i rozkład wartości EF_{frac} dla dokowań w Scenariuszu B.....	113
Tabela 13. Średnie wartości parametru logAUC i rozkład wartości EF_{frac} dla dokowań w Scenariuszu C.....	115
Tabela 14. Średnie wartości parametru logAUC i rozkład wartości EF_{frac} dla dokowań w Scenariuszu D.	116
Tabela 15. Scenariusze trenowania i walidacji funkcji rescoreingowych (trenowanie z wykorzystaniem wyłącznie Zestawu I)	122
Tabela 16. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W1 (T1W1). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	123
Tabela 17. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W2 (T1W2). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	124

Tabela 18. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W3 (T1W3). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	125
Tabela 19. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W4 (T1W4). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	126
Tabela 20. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W1 (T2W1). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	127
Tabela 21. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W2 (T2W2). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	128
Tabela 22. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W3 (T2W3). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	129
Tabela 23. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W4 (T2W4). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	130
Tabela 24. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W1 (T3W1). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	131
Tabela 25. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W2 (T3W2). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	132
Tabela 26. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W3 (T3W3). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	133
Tabela 27. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W4 (T3W4). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	134
Tabela 28. Scenariusze trenowania i walidacji funkcji rescoreingowych (trenowanie z wykorzystaniem podzbioru Zestawów I i II)	139

Tabela 29. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W5 (T4W5). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	140
Tabela 30. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W6 (T4W6). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	141
Tabela 31. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W7 (T4W7). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	142
Tabela 32. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W8 (T4W8). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	143
Tabela 33. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W5 (T5W5). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	144
Tabela 34. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W6 (T5W6). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	145
Tabela 35. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W7 (T5W7). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	146
Tabela 36. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W8 (T5W8). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	147
Tabela 37. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W5 (T6W5). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	148
Tabela 38. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W6 (T6W6). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	149
Tabela 39. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W7 (T6W7). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	150

Tabela 40. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W8 (T6W8). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.	151
Tabela 41. Wzory strukturalne związków z serii W-71.....	169
Tabela 42. Wzory strukturalne związków z serii F-71.	173
Tabela 43. Wzory strukturalne związków z serii W-72.....	177
Tabela 44. Wzory strukturalne związków z serii F-72.	181
Tabela 45. Wzory strukturalne związków z serii W-73.....	184
Tabela 46. Wzory strukturalne związków z serii F-73.	187
Tabela 47. Wzory strukturalne związków z serii W-74.....	190
Tabela 48. Wzory strukturalne związków z serii F-74.	194
Tabela 49. Wzory strukturalne związków z serii W-80.....	197
Tabela 50. Wzory strukturalne związków z serii W-81.....	199
Tabela 51. Wzory strukturalne związków z serii W-82.....	201
Tabela 52. Wzory strukturalne związków z serii W-83.....	204
Tabela 53. Wzory strukturalne związków z serii W-84.....	206
Tabela 54. Wzory strukturalne związków z serii W-90.....	208

Wykaz skrótów

Ac	grupa acetylowa
Ac₂O	bezwodnik octowy
ACN	acetonitryl
AP-2	białko adaptorowe-2
BBB	bariera krew-mózg (ang. <i>blood-brain barrier</i>)
Boc	grupa <i>tert</i> -butyloksykarbonylowa
Bzl	grupa benzylowa
cAMP	cykliczny monofosforan adenozyiny
Cbz	grupa benzyloksykarbonylowa
CCK2R	receptor cholecystokininowy 2
CPT	kamptotecyna
DAG	diacyloglicerol
DCC	<i>N,N'</i> -dicykloheksylokarbodiimid
DCM	dichlorometan (chlorek metylenu)
DCU	<i>N,N'</i> -dicykloheksylomocznik
D-Lys(Nic)	<i>N</i> ^ε -nikotynoilo-D-lizyna
DMF	<i>N,N</i> -dimetyloformamid
Dmt	2',6'-dimetylotyrozyna
DOTA	kwas 1,4,7,10-tetraazacyklododekan-1,4,7,10-tetraoctowy
DOTAGA	kwas 1,4,7,10-tetraazacyklododekan-4,7,10-trioctowo-1-[2-glutarowy]
D-Phe(Cl₂)	3,4-dichloro-D-feniloalanina
DTPA	kwas dietyleno-triamino-pentaoctowy
ECE-1	enzym konwertujący endotelinę 1
ECL	pętla zewnątrzkomórkowa (ang. <i>extracellular loop</i>)
EF	współczynnik wzbogacenia (ang. <i>enrichment factor</i>)
EF_{frac}	znormalizowany współczynnik wzbogacenia
EGFR	nabłonkowy czynnik wzrostu (ang. <i>epidermal growth factor receptor</i>)
EK	endokinina
ekw.	ekwiwalent (równoważnik molowy, gramorównoważnik)
ERK	kinaza regulowana sygnałem zewnątrzkomórkowym (ang. <i>extracellular signal regulated kinase</i>)
ESI	elektrorozpylanie (ang. <i>electrospray</i>)
Et₂O	eter dietylowy
EtOAc	octan etylu (ang. <i>ethyl acetate</i>)
FA	kwas mrówkowy (ang. <i>formic acid</i>)
GABA	kwas γ -aminomasłowy (ang. <i>γ-aminobutyric acid</i>)
GDP	guanozyno-5-difosforan
GRK	kinazy białek sprzężonych z białkiem G
GTP	guanozyno-5-trifosforan
GPCR	receptor sprzężony z białkiem G (ang. <i>G-protein coupled receptor</i>)
HK-1	hemokinina-1

HOSu	<i>N</i> -hydroksysukcynoimid
IBS	zespół jelita drażliwego (ang. <i>irritable bowel syndrome</i>)
ICH TJ	Instytut Chemii i Techniki Jądrowej w Warszawie
ICL	pętla wewnątrzkomórkowa (ang. <i>intracellular loop</i>)
IP₃	1,4,5-trifosforan inozytolu
iPrOH	izopropanol
LC-MS	chromatografia cieczowa sprzężona ze spektrometrią mas (ang. <i>liquid chromatography mass spectrometry</i>)
logAUC	pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej
MAPK	kinazy aktywowane mitogenami (ang. <i>mitogen activated kinases</i>)
MAE	średni błąd bezwzględny (ang. <i>Mean Absolute Error</i>)
MeOH	metanol
MLSF	funkcje oceniające uczone maszynowo (ang. <i>machine-learned scoring function</i>)
ML-TSSF	funkcja/funkcje oceniające specyficzne dla celu molekularnego uczone maszynowo (ang. <i>machine-learned target specific scoring function</i>)
NF-κB	jądrowy czynnik transkrypcyjny NF kappa B
NLS	sygnał lokalizacji jądrowej (ang. <i>nuclear localization signal</i>)
NK1R	receptor NK1
NK1R-F1	pełna (F1 — ang. <i>full</i>) forma receptora NK1
NK1R-Tr	skrócona (Tr — ang. <i>truncated</i>) forma receptora NK1
NK2R	receptor NK2
NK3R	receptor NK3
NKA	neurokinina A
NKB	neurokinina B
NMR	spektroskopia magnetycznego rezonansu jądrowego (ang. <i>Nuclear Magnetic Resonance</i>)
NOTA	kwasy 1,4,7-triazacyklononan-1,4,7-trioctowy
NPK	neuropeptyd K
NPγ	neuropeptyd γ
pAff	ujemny logarytm eksperymentalnego powinowactwa
Pal3	3-(3-pirydylo)alanina
PDB	baza danych Protein Data Bank
PEG	poli(tlenek etylenu) (glikol polietylenowy)
PET	pozytonowa tomografia emisyjna (ang. <i>positron emission tomography</i>)
pGlu	kwasy piroglutaminowy
PI3K	kinaza 3-fosfoinozytydu
PIP₂	4,5-bisfosforan fosfatydyloinozytolu
PIP₃	3,4,5-trifosforan fosfatydyloinozytolu
PKA	kinaza białkowa A (ang. <i>protein kinase A</i>)
PKC	kinaza białkowa C (ang. <i>protein kinase C</i>)
PLCβ	fosfolipaza Cβ (ang. <i>phospholipase Cβ</i>)
PPT	protachykinina

QSAR	ilościowe zależności struktura-aktywność (ang. <i>Quantitative Structure–Activity Relationships</i>)
ROC	krzywa charakterystyki operacyjnej odbiornika (ang. <i>receiver operating characteristic curve</i>)
ROCK	kinaza białkowa związana z białkami Rho (ang. <i>Rho-associated protein kinase</i>)
QSAR	ilościowe zależności struktura-aktywność (ang. <i>Quantitative Structure–Activity Relationships</i>)
SAR	badania zależności struktura chemiczna–aktywność biologiczna (ang. <i>Structure–Activity Relationship</i>)
SBVS	przesiew wirtualny oparty na strukturze (ang. <i>structure-based virtual screening</i>)
SP	substancja P
SPECT/CT	tomografia emisyjna pojedynczych fotonów/tomografia komputerowa (ang. <i>single-photon emission computed tomography/computed tomography</i>)
TFA	kwas trifluorooctowy (ang. <i>trifluoroacetic acid</i>)
TFMP	grupa 3,5- <i>bis</i> (trifluorometylo)fenylowa
Thi	tienyloalanina
TLC	chromatografia cienkowarstwowa (ang. <i>thin-layer chromatography</i>)
TM	helisa przezbłonowa (od ang. <i>transmembrane</i>)
TMG	1,1,3,3-tetrametyloguanidyna
TRT	celowana terapia radionuklidowa (ang. <i>targeted radionuclide therapy</i>)
TSSF	funkcja/funkcje oceniające specyficzne dla celu molekularnego (ang. <i>target-specific scoring functions</i>)
ZN IMDiK PAN	Zakład Neuropeptydów Instytutu Medycyny Doświadczalnej i Klinicznej im. M. Mossakowskiego Polskiej Akademii Nauk

I. Wstęp

W obliczu coraz bardziej złożonych wyzwań zdrowotnych, konieczne jest poszukiwanie nowych strategii terapeutycznych, opartych na interdyscyplinarnym podejściu i najnowszych osiągnięciach naukowych. Szczególnie ważna jest identyfikacja nowych substancji czynnych, które mogłyby stanowić skuteczne i bezpieczne terapie.

W tej perspektywie receptor tachykininowy NK1 (NK1R) oraz jego endogenny ligand, substancja P (SP), wyłaniają się jako ważne cele molekularne. Ich udział w regulacji szeregu procesów fizjologicznych i patologicznych czyni je punktami centralnymi zarówno jeśli chodzi o mechanizmy chorobowe jak i potencjalne terapie. Obecnie w medycynie stosowanych jest pięć antagonistów receptora NK1 we wskazaniu do zapobiegania nudnościom i wymiotom związanym z emetogenną chemioterapią nowotworów. Wydaje się jednak, że nie wyczerpuje to możliwości wykorzystania NK1R jako punktu uchwytu dla terapii. Badania prowadzone na przestrzeni lat sugerowały obiecujący potencjał NK1R w m.in. terapiach psychiatrycznych (depresji, schizofrenii, uzależnień), czy leczeniu bólu. W wielu przypadkach początkowe nadzieje nie spełniły się, ale dalsze, bardziej zaawansowane próby celowania w NK1R, nie ustają. W ostatnich latach, badacze zwrócili uwagę na m.in.:

- możliwość włączania farmakoforu NK1R do wielocelowych związków przeciwbólowych,
- antagonizm endosomalnego sygnałowania przez NK1R,
- selektywną cytotoksyczność antagonistów NK1R,
- wykorzystanie ligandów NK1R jako wektorów dla celowanego dostarczania elementu terapeutycznego (m.in. terapeutycznych radionuklidów).

Te specyficzne zastosowania NK1R wymagają poszukiwania nowych ligandów tego receptora lub przeprojektowania ligandów już istniejących. W tym kontekście, podjęłam się **zaprojektowania i syntezy biblioteki nowych pochodnych fenyloalaniny i tryptofanu jako potencjalnych ligandów NK1R.**

Inny aspekt moich badań dotyczył rozwoju funkcji oceniających specyficznych dla receptora NK1. We współczesnym poszukiwaniu nowych substancji aktywnych, kluczową rolę odgrywa wykorzystanie metod *in silico*. Jedną z istotnych technik stosowanych w tym

obszarze jest dokowanie molekularne, które ma na celu przewidywanie oddziaływań między potencjalnymi cząsteczkami leków a wybranymi celami molekularnymi. Dokowanie molekularne jest metodą stosunkowo szybką i tanią obliczeniowo. Dzięki temu zyskało ono bardzo dużą popularność jako narzędzie racjonalnego projektowania leków. Z drugiej jednak strony, stosowanie tej techniki wiąże się z różnymi ograniczeniami, z których być może najważniejszym jest niedokładność predykcji siły wiązania przez funkcje oceniające (ang. *scoring functions*).

Niektórzy autorzy wskazują, że przynajmniej częściowo problem ten wynika z zakładanej uniwersalności popularnych funkcji oceniających. Funkcje takie są bowiem kalibrowane na danych eksperymentalnych dotyczących wielu różnych białek receptorowych, pochodzących z różnych rodzin. Rozwiązaniem poprawiającym dokładność *scoringu* może być przygotowywanie funkcji oceniających „szytych na miarę”, specyficznych dla danego receptora. Można to osiągnąć przez dostrajanie parametrów funkcji do danych eksperymentalnych związanych z jednym konkretnym celem molekularnym (lub rodziną celów molekularnych). Takie specyficzne modele *scoringu* powinny dawać lepsze predykcje w badaniach związanych z wybranym celem molekularnym. W tym kontekście postanowiłam wykorzystać dane eksperymentalne o powinowactwie receptorowym zsyntezowanych przez mnie związków (zmierzone przez współpracowników) do **opracowania funkcji oceniających specyficznych dla receptora NK1**.

W związku z przedmiotem moich badań, w części literaturowej prezentowanej rozprawy omówiłam najpierw szczegółowo rolę receptora NK1 i SP w fizjologii i patofizjologii człowieka (Rozdział II.1). Dalej, scharakteryzowałam znanych antagonistów NK1R z różnych grup chemicznych (Rozdział II.2) i przedstawiłam opisane próby ich stosowania w farmakologii. Oddzielną sekcję poświęciłam omówieniu najnowszych prób celowania w receptor NK1 jako innowacyjnej interwencji terapeutycznej (podrozdział II.2.3). W kolejnej części skupiłam się na dokowaniu molekularnym i funkcjach oceniających stosowanych w dokowaniu molekularnym (Rozdział II.3). Tam przedstawiłam problem związany z niedokładnością predykcji uzyskiwanych w trakcie dokowania oraz opisałam tworzenie funkcji oceniających specyficznych dla celu molekularnego (Rozdział II.3.3).

II. Część literaturowa

II.1. Substancja P i receptor NK1

II.1.1. Rodzina peptydów tachykininowych

Tachykininy to grupa peptydów szeroko rozpowszechniona w obwodowym i ośrodkowym układzie nerwowym, a także w komórkach i tkankach poza układem nerwowym. Obecność tachykinin opisano w wielu organizmach, zarówno u bezkręgowców jak i u kręgowców, w szczególności u ssaków [5,6]. Przez lata tachykininy klasyfikowane były jako peptydy pochodzenia neuronalnego — neuropeptydy. Obecnie uważa się, że istnieją także inne, poza układem nerwowym, miejsca syntezy i wydzielania tych peptydów, np. poprzez limfocyty [7,8]. Nazwa „tachykininy” powiązana jest ze zdolnością tych związków do szybkiego wywoływania skurczu mięśni gładkich jelit (gr. *tachys* — szybki) [5].

Wszystkie tachykininy charakteryzują się posiadaniem podobnej sekwencji aminokwasów na C-końcu cząsteczki: Phe-X-Gly-Leu-Met-NH₂, gdzie X oznacza aminokwas aromatyczny, fenyloalaninę (Phe) lub tyrozynę (Tyr) bądź aminokwas alifatyczny, walinę (Val) lub izoleucynę (Ile) [5]. Wykazano, że obecność na C-końcu ugrupowania Met-NH₂ jest kluczowa dla aktywności tachykinin. Usunięcie tego fragmentu, utlenienie łańcucha bocznego Met do sulfotlenku, lub też usunięcie C-końcowej grupy amidowej znosi aktywność tych peptydów [5]. C-końcowy fragment sekwencji uważany jest za „sekwencję przenoszącą wiadomość” (ang. *message*), niezbędną i odpowiadającą za przekazanie sygnału, zaś fragmenty N-końcowe poszczególnych tachykinin nazywane są „sekwencjami adresowymi” (ang. *address*), odpowiadającymi za selektywność względem poszczególnych receptorów [5].

Do najbardziej znanych przedstawicieli tachykinin zalicza się substancję P (SP), neurokininę A (NKA) i neurokininę B (NKB). Substancja P, odkryta w 1931 roku, została pierwotnie wyizolowana jako ekstrakt z końskich jelit i mózgu, o silnych właściwościach pobudzających skurcze mięśni gładkich [9]. W 1934 roku ekstrakt ten nazwano „Substancją P” nawiązując do fizycznej postaci uzyskanego ekstraktu – proszku (ang. *powder*) [10]. Szeroko zakrojone prace badawcze doprowadziły do wyizolowania SP w czystej postaci oraz poznania jej właściwości i działania, co przyczyniło się do dalszego rozwoju wiedzy o tachykininach [11].

W roku 1983 odkryto kolejne tachykininy, NKA oraz NKB. NKA, znana również jako Substancja K (nazywana tak z powodu podobieństwa do kassininy, peptydu tachykininowego występującego u płazów), została początkowo wyizolowana z ekstraktów bydłęcego rdzenia kręgowego jako związek o charakterze kationowym [5,12]. NKB, znana również pod nazwą neuromedyna K, została wyizolowana ze świńskiego rdzenia kręgowego jako tachykinina o charakterze anionowym [13]. SP oraz NKA występują w głównie w obwodowym układzie nerwowym, natomiast NKB w ośrodkowym układzie nerwowym [14]. U ssaków rozpoznano również inne tachykininy, takie jak: neuropeptyd K (NPK), neuropeptyd γ (NP γ), hemokininę-1 (HK-1) i endokininy (EK) [14]. Do tachykinin występujących u zwierząt innych niż ssaki zalicza się eledoizynę, fisalaeminę, filomeduzynę, uperoleinę i kassininę [5]. Znanych jest ponad czterdzieści tachykinin [15]. W Tabeli 1 zestawiono przykładowe tachykininy i ich sekwencje.

Tabela 1. Przykładowe peptydy tachykininowe i ich sekwencje.

Nazwa	Sekwencja*
Substancja P	Arg-Pro-Lys-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH ₂
Neurokinina A	His-Lys-Thr-Asp-Ser-Phe-Val-Gly-Leu-Met-NH ₂
Neurokinina B	Asp-Met-His-Asp-Phe-Phe-Val-Gly-Leu-Met-NH ₂
Neuropeptyd γ	Asp-Ala-Gly-His-Gly-Gln-Ile-Ser-His-Lys-Arg-Hys-Lys-Thr-Asp-Ser-Phe-Val-Gly-Leu-Met-NH ₂
Eledoizyna	pGlu ^a -Pro-Ser-Lys-Asp-Ala-Phe-Ile-Gly-Leu-Met-NH ₂
Kassinina	Asp-Val-Pro-Lys-Ser-Asp-Gln-Phe-Val-Gly-Leu-Met-NH ₂

* - wytłuszczono charakterystyczny fragment C-końcowy. ^a - pGlu - kwas piroglutaminowy.

Za biosyntezę tachykinin odpowiedzialne są geny *TAC1*, *TAC3* i *TAC4*. Geny te kodują białka prekursorowe — preprotachykininy PPT-A (*TAC1*), PPT-B (*TAC3*) i PPT-C (*TAC4*). W przypadku *TAC1*, w wyniku alternatywnego splicingu, mogą powstawać różne izoformy PPT-A, z których w następstwie modyfikacji posttranslacyjnych produkowane są SP, NKA, NPK, NP γ . *TAC3* koduje prekursor, z którego powstaje NKB, z kolei PPT-C odpowiada za syntezę HK-1 i jej krótszej pochodnej hemokinin 4–11 oraz czterech endokinin: A, B, C, D (EKA, EKB, EKC, EKD) [6,11,16].

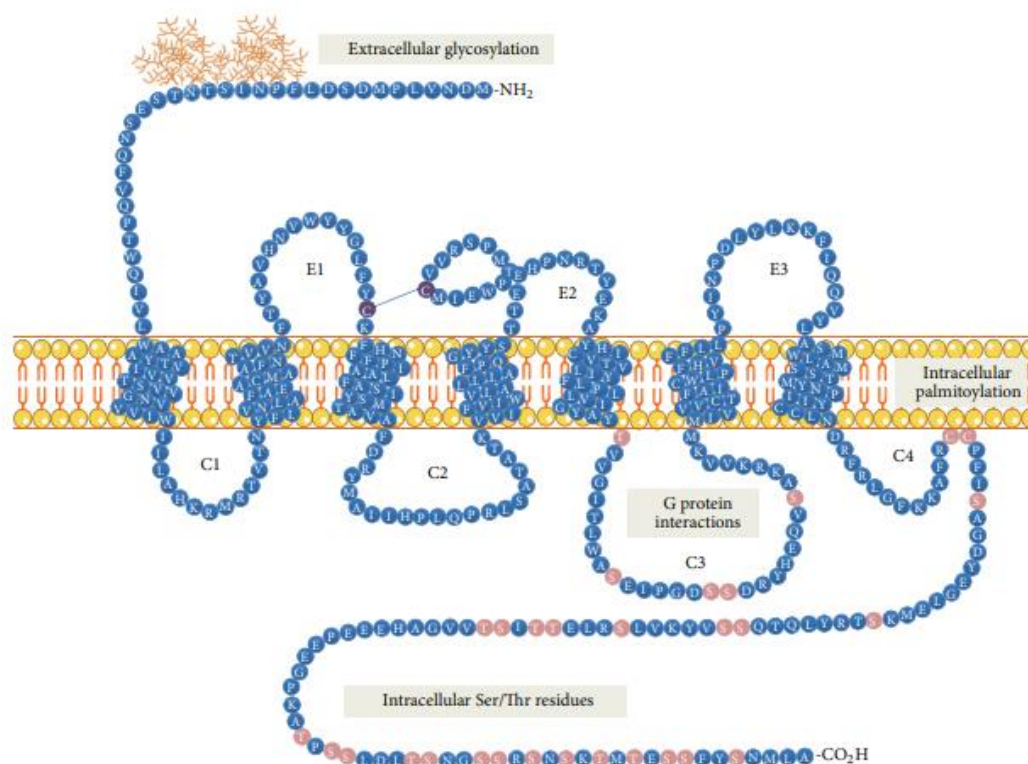
II.1.2.Receptor NK1

Działanie biologiczne tachykinin związane jest ze zdolnością tych peptydów do aktywacji receptorów tachykininowych, zwanych także receptorami neurokininowymi (NKR, ang. *neurokinin receptor*). Receptory tachykininowe należą do rodziny przezbłonowych receptorów metabotropowych sprzężonych z białkiem G (GPCR, ang. *G-protein coupled*

receptors). Wyróżnia się trzy typy receptorów tachykininowych — NK1R, NK2R, NK3R, które zbudowane są z około 300 do 500 reszt aminokwasowych, tworzących pojedynczy polipeptyd [17]. Sekwencja aminokwasowa poszczególnych receptorów jest w ponad 65% homologiczna [18]. Głównym ligandem receptora NK1 jest SP, natomiast NKA i NKB wykazują oddziaływanie odpowiednio około 100 i 500 razy słabsze niż SP. Do receptora NK2 największe powinowactwo wykazuje NKA, mniejsze NKB i SP. Najsilniej z receptorem NK3 oddziałuje NKB, słabiej NKA i SP [19]. Każda z tych tachykinin jest agonistą receptorów NK, o różnym stopniu powinowactwa do poszczególnych jego typów.

Ponieważ tematyka badań przedstawianych w rozprawie związana jest z receptorem NK1, dalsza treść wstępu ograniczona będzie zasadniczo do tego receptora i jego głównego liganda, substancji P.

Receptor NK1, kodowany przez gen *TACR1*, występuje w postaci dwóch izoform. Dominującą formą jest postać pełna receptora, składająca się z 407 aminokwasów, która ulega ekspresji głównie w ośrodkowym i obwodowym układzie nerwowym [20]. Forma skrócona, zawierająca 311 aminokwasów (pozbawiona fragmentu od strony C-końca) powstaje gdy intron znajdujący się pomiędzy czwartym a piątym egzonem nie zostanie usunięty, co prowadzi do przedwczesnej identyfikacji kodonu STOP [19]. Forma ta ulega ekspresji w różnych tkankach i komórkach, m.in. w monocytach, limfocytach T, komórkach NK (ang. *natural killers*) [20].



Rysunek 1. Schematyczny model budowy receptora NK1.

Reprodukcja z Ref. [19]. Obraz na licencji Creative Commons Attribution License. Legenda: E1-E3 – extracellular loop, pętle zewnątrzkomórkowe, C1-C3 - intracellular loop, pętle wewnątrzkomórkowe, intracellular palmitoylation – miejsce palmitoilacji cysteiny z możliwością utworzenia pętli wewnątrzkomórkowej C4, extracellular glycosylation – zewnątrzkomórkowe miejsce glikozylacji asparaginy Asn, intracellular Ser/Thr residues – miejsca fosforylacji seryny lub treoniny, G protein interactions – miejsca interakcji z białkiem G.

Receptor NK1, podobnie jak pozostałe receptory tachykininowe, zanurzony jest w błonie komórkowej, przez którą łańcuch polipeptydowy przechodzi siedmiokrotnie tworząc domenę transmembranową, na którą składa się siedem hydrofobowych α -helis przezbłonowych (TM1–TM7; TM od ang. *transmembrane*) połączonych trzema pętlami zewnątrzkomórkowymi (ECL1–ECL3; ECL od ang. *extracellular loop*) i trzema pętlami wewnątrzkomórkowymi (ICL1–ICL3; ICL od ang. *intracellular loop*). Istnieje także możliwość utworzenia czwartej pętli, ze względu na palmitoilację cysteiny, powodującą zakotwiczenie receptora w błonie plazmatycznej [20]. N-koniec łańcucha białkowego z wolną grupą aminową znajduje się po stronie zewnątrzkomórkowej i zawiera potencjalne miejsca glikozylacji asparaginy, zaś C-koniec położony jest po stronie cytozolowej błony komórkowej, gdzie znajdują się miejsca fosforylacji seryny lub treoniny przez odpowiednie kinazy białkowe GRK (ang. *G protein coupled receptor kinase*) [21]. Receptor NK1 od strony zewnątrzkomórkowej posiada miejsce wiążące ligandy. Miejsce wiązania białka G znajduje się po stronie cytozolowej [22].

Receptory NK1 różnych gatunków są wysoce homologiczne. Przykładowo, identyczność sekwencji pomiędzy szczurzym a ludzkim receptorem NK1 wynosi 95%, a pomiędzy receptorem kawii domowej a ludzkim 96.5%. Pomimo tego występują duże różnice we właściwościach farmakologicznych receptorów NK1 różnych gatunków i w powinowactwie poszczególnych ligandów do tych receptorów [18].

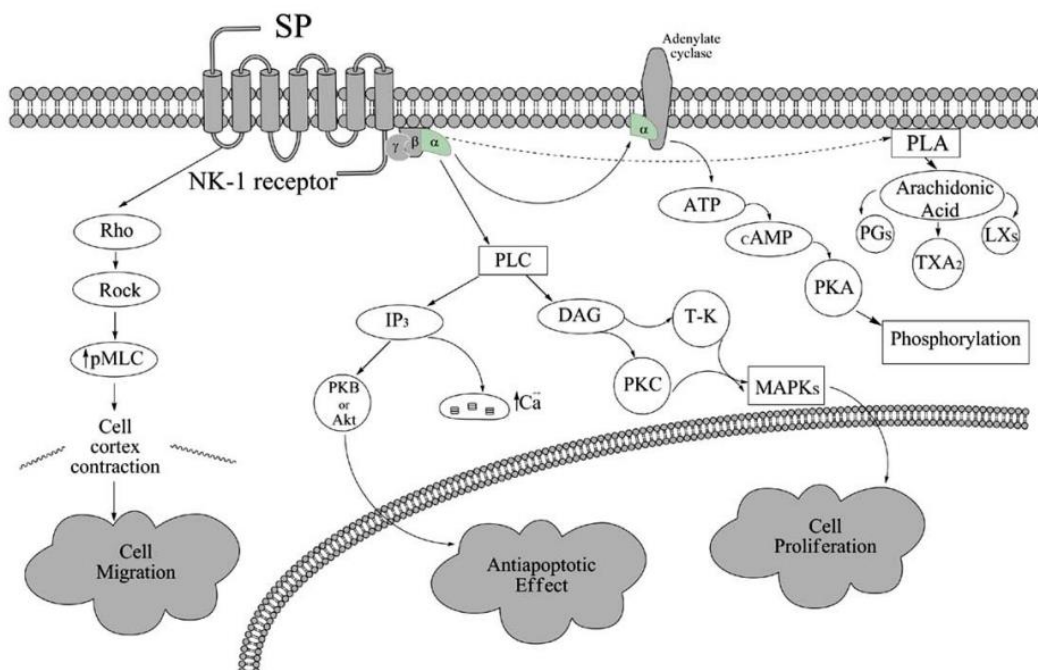
W roku 2018 rozwiązano pierwszą strukturę krystalograficzną ludzkiego receptora NK1R z antagonistą receptora w miejscu wiążącym, która dostępna jest pod kodem 6E59 w bazie danych Protein Data Bank (PDB) [23]. W następnych latach opublikowano jedenaście dalszych struktur (rozwiązanych za pomocą metod rentgenografii strukturalnej lub mikroskopii krioelektronowej) z różnymi antagonistami i agonistami [24–28]. Niedawno rozwiązano także strukturę receptora NK2 [29]. Odkrycia te przyczyniają się do lepszego zrozumienia oddziaływań receptor-ligand, selektywności na poziomie molekularnym i w przyszłości dają możliwość racjonalnego projektowania substancji aktywnych o wysokim powinowactwie do ludzkiego NK1R w oparciu o jego strukturę molekularną.

Jak już wspomniałam, endogennym agonistą receptora NK1 o wysokim powinowactwie jest Substancja P. Aktywacja receptora NK1 przez SP lub innego agonistę prowadzi do zmian konformacyjnych, dzięki którym możliwe jest związanie białka G przez receptor. Białka G są heterotrimerami i składają się podjednostek: α (która w formie nieaktywnej związana jest z guanozyno-5'-difosforanem, GDP), β oraz γ [30]. W zależności od rodzaju liganda i przyjętej konformacji aktywnej NK1R może być sprzęgany z białkami G różnego rodzaju, tj. zawierającymi w swej budowie podjednostki α różnego typu ($G\alpha_q$, $G\alpha_s$, $G\alpha_i$, $G\alpha_o$, $G\alpha_{12/13}$). Poszczególne rodzaje białek G różnią się co do tego, które wewnątrzkomórkowe ścieżki sygnalizacyjne uruchamiają [19].

Związanie białka G przez receptor powoduje jego aktywację. Następuje wymiana cząsteczki GDP na guanozyno-5'-trifosforan (GTP) i rozpad trimery na podjednostkę $G\alpha$ -GTP oraz dimer $G\beta\gamma$. Receptorowe miejsce wiązania białka G zostaje uwolnione i możliwe jest sprzęganie receptora z kolejnymi cząsteczkami białka G. Zdysocjowane podjednostki $G\alpha$ -GTP oraz $G\beta\gamma$ rozpoczynają kaskady sygnalizacyjne oddziałując na specyficzne dla siebie enzymy efektorowe i kanały jonowe, przez aktywowanie lub hamowanie ich funkcji [21]. Sygnał jest przekazywany poprzez szereg białek i cząsteczek sygnalizacyjnych od miejsca wiązania liganda do jądra komórkowego lub innych organelli. Pojedyncza cząsteczka $G\alpha$ -GTP przestaje wpływać na sygnalizację wewnątrzkomórkową w momencie hydrolizy

GTP i połączenia podjednostki $G\alpha$ z GDP, a w następstwie utworzenie nieaktywnego kompleksu połączonych trzech podjednostek (powrót do stanu wyjściowego) [21].

W wyniku oddziaływania białek G z NK1R związanym z ligandem dochodzi do aktywacji różnych szlaków przekąźnikowych.



Rysunek 2. Szlaki sygnałowe związane z receptorem NK1.

Reprodukcja z Ref. [31]. Obraz na licencji Creative Commons Attribution License.

Legenda: Aktywacja NK1R przez SP prowadzi do proliferacji komórek, działania antyapoptycznego i migracji komórek. SP – Substancja P, α , $\beta\gamma$ – podjednostki białka G, Adenylate cyclase – cyklaza adenylova, ATP – trójfosforan adenozyiny, cAMP – cykliczny monofosforan adenozyiny, PKA – kinaza białkowa A, PLC – fosfolipaza C, IP₃ – 1,4,5-trifosforan inozytoli, DAG – 1,2-diacyloglicerol, PKC – szlak kinazy białkowej C, T-K – kinaza tyrozynowa, MAPKs – kinazy aktywowane mitogenami, PKB lub Akt – kinaza białkowa B, Rho – białko Rho, Rock – kinaza białkowa związana z białkami Rho, pMLC- fosforylacja lekkich łańcuchów miozynowych, PLA – fosfolipaza A, arachidonic acid – kwas arachidonowy, PGs – prostacykliny, TXA₂ - tromboksan A₂, LXs – leukotrieny.

Białka G z podjednostką $G\alpha_q$ powodują aktywację fosfolipazy C β (PLC β), która promuje hydrolizę 4,5-bisfosforanu fosfatydyloinozytoli (PIP₂), z wytworzeniem dwóch wtórnych przekąźników: 1,2-diacyloglicerolu (DAG) oraz 1,4,5-trifosforanu inozytoli (IP₃) [32]. DAG indukuje napływ jonów wapnia z przestrzeni zewnątrzkomórkowej w następstwie otwarcia kanałów wapniowych typu L oraz aktywuje szlak kinazy białkowej C (PKC) [33]. PKC może oddziaływać na inne szlaki sygnałowe, m.in. szlak kinaz aktywowanych mitogenami (MAPK). IP₃ poprzez aktywację jonotropowych receptorów IP₃R prowadzi do wewnątrzkomórkowego wzrostu stężenia jonów wapnia Ca²⁺ w wyniku ich uwolnienia z siateczki śródplazmatycznej do cytozolu [16]. Wzrost stężenia jonów wapnia może

skutkować aktywacją różnych kinaz białkowych, m.in. PKC oraz kinazy 3-fosfoinozytydu (PI3K) [34].

Białka G z podjednostką $G\alpha_s$ powodują aktywację cykazy adenylowej (AC), która katalizuje konwersję cytoplazmatycznego ATP do cyklicznego monofosforanu adenozy (cAMP) [35]. cAMP, będący wtórnym przekaźnikiem wpływa na aktywację kinazy białkowej A (PKA) [36]. PKA prowadzi do aktywacji enzymów, odpowiedzialnych za uwalnianie energii z magazynów wewnątrzkomórkowych i regulację przemiany materii. PKA może także regulować szlak PLC oraz wpływać hamująco na działanie jądrowego czynnika transkrypcyjnego NF kappa B (NF- κ B) [36].

Białka G z podjednostkami $G\alpha_{12/13}$ powodują aktywację kinazy białkowej związanej z białkami Rho (ROCK), która wpływa na fosforylację lekkich łańcuchów miozynowych. Aktywacja tego szlaku ma znaczenie w regulacji apoptozy, procesach dezorganizacji cytoplazmy i migracji komórek. Z tym szlakiem sygnałowania ma związek powstawanie uwypukleń błony komórkowej (ang. *blebbing*), które bywa elementem procesu apoptozy. W przypadku sygnałowania wywołanego przez oddziaływanie SP/NK1R, *blebbing* jest jednak niezwiązany z apoptozą [37].

Oddziaływanie SP/NK1R wiąże się także z aktywacją fosfolipazy A2 (PLA2), która indukuje wytwarzanie kwasu arachidonowego, mogącego oddziaływać na kinazy PKC i MAPK, a także będącego prekursorem do syntezy eikozanoidów — mediatorów stanu zapalnego, m.in. prostaglandyn, prostacyklin, tromboksanu, leukotrienów [38]

W kontekście udziału osi SP/NK1R w niektórych procesach fizjologicznych i patologicznych (np. regeneracja tkanek, zapalenie, nowotworzenie) wykazano aktywację szeregu szlaków przekaźnikowych oraz enzymów, takich jak szlak JAK/STAT, szlak Wnt/ β -kateinina, szlak PI3K/Akt, niereceptorowych kinaz tyrozynowych, m.in. Src, Pyk2, biorących udział w aktywacji kinaz regulowanych sygnałem zewnątrzkomórkowym (ERK) ERK1 i ERK2 [22,39,40]. Ponadto, oddziaływanie SP z NK1R sprzyja aktywacji jądrowego czynnika transkrypcyjnego NF kappa (NF- κ B), który odgrywa znaczącą rolę w promowaniu proliferacji, hamowaniu apoptozy, transkrypcji DNA oraz produkcji cytokin prozapalnych i mediatorów stanu zapalnego, takich jak interleukina-1 (IL-1), interleukina-6 (IL-6), czynnik martwicy nowotworów TNF- α , białko zapalne makrofagów-1 β (MIP-1 β), interferon γ (IFN- γ) [20]. SP/NK1R ma również wpływ na fosforylację receptora nabłonkowego czynnika wzrostu (EGFR) i może uruchamiać szlaki sygnałowe zależne od EGFR, prowadząc do aktywacji MAPK, ERK1 i ERK2, syntezy i proliferacji DNA [22,41]. Uważa się, że oddziaływanie pomiędzy SP, receptorem NK1 i EGFR mogą wpływać na migrację i inwazję

komórek, rozwój różnych chorób, takich jak nowotwory, choroby neurologiczne i zapalne [20,41].

Sygnalizacja związana z receptorem NK1 jest regulowana w wyniku działania mechanizmów, które są odpowiedzialne za usuwanie agonistów z przestrzeni zewnątrzkomórkowej (np. degradacja Substancji P przez enzymy takie jak neprylizyna czy ECE-1) oraz zmniejszanie zdolności receptora do łączenia się z białkiem G (rozprężanie) i jego odczulanie (desensytyzacja) [39]. Odczulanie NK1R odbywa się z udziałem kinaz GRK2, GRK3 oraz GRK5, które fosforylują receptor, co sprzyja jego oddziaływaniu z β -arestynami. β -Arestyny, wiążąc się z wewnątrzkomórkowym miejscem wiążącym receptora, uniemożliwiają związanie receptora z białkiem G. Resensytyzacja sygnałowania SP odbywa się przez endocytozę receptora zależną od dynaminy i klatryny, dysocjację kompleksu SP/NK1R, rozłożenie SP przez ECE-1, dysocjację β -arestyn i recycling receptora. Możliwy jest także inny mechanizm resensytyzacji (bez endocytozy) z udziałem fosfatazy białkowej 2A, która przyłącza się do kompleksu przez oddziaływanie z β -arestyną i defosforyluje receptor NK1R [42].

Warto tutaj zwrócić uwagę, że endocytoza receptora NK1 nie wiąże się jedynie z procesami odczulania. β -Arestyny mogą uczestniczyć w tworzeniu sygnałosomów (ang. *signalosomes*) z udziałem internalizowanego NK1R oraz kinaz Src, MEKK i ERK1/2. Kompleksy takie umożliwiają sygnałowanie z udziałem receptora nawet po endocytozie. Internalizowane NK1R może także sygnałować z udziałem białka G z podjednostką $G\alpha_q$, prowadząc do aktywacji jądrowego ERK i cytozolowego PKC i cAMP [43]. Wykazano znaczenie endosomalnego sygnałowania przez NK1R np. w procesach przekazywania bólu [43] czy nowotworzenia [44].

Dalej warto nadmienić, że skutki oddziaływania SP z NK1R zależą od izoformy receptora. Pełna izoforma receptora NK1 (NK1R-F1) ma większe powinowactwo do SP (około 10 razy), jak również do innych agonistów i antagonistów [45]. Ponadto, NK1R-F1, którego C-koniec może ulegać fosforylacji poprzez działanie kinaz GRK [39], jest w stanie oddziaływać z β -arestynami [46]. Forma skrócona (NK1R-Tr), pozbawiona fragmentu C-końcowego, nie podlega fosforylacji, co zaburza internalizację receptora i powoduje, że jest on odporny na odczulanie. W odróżnieniu od pełnej izoformy, sygnałowanie przez NK1R-Tr charakteryzuje się zmniejszoną fosforylacją PKC, opóźnioną aktywacją ERK, słabszymi właściwościami aktywującymi NF- κ B oraz wewnątrzkomórkowy wyrzut jonów wapnia [47].

II.1.3.Rola SP i NK1R w procesach fizjologicznych i patologicznych

Substancja P oraz receptor NK1 występują w wielu różnych tkankach w organizmie ludzkim, zarówno w obwodowym, jak i ośrodkowym układzie nerwowym [48]. SP jest również wytwarzana poza układem nerwowym, m.in. w komórkach śródbłonna, w komórkach układu odpornościowego (makrofagach, eozynofilach, limfocytach, monocytach, neutrofilach, komórkach tucznych). SP jest znajdowana także w płynach ustrojowych, m.in. w krwi, ślinie, mleku matki, płynie mózgowo-rdzeniowym [19,49]. Obecność NK1R potwierdzono na komórkach układu odpornościowego, pokarmowego, moczowo-płciowego, sercowo-naczyniowego, krwiotwórczego, śledzionie, łożysku, skórze, komórkach mięśni gładkich, płuc i śródbłonna, a także komórkach nowotworowych [19,39]. Szerokie rozpowszechnienie SP i receptora NK1 w organizmie ludzkim sugeruje uczestnictwo układu SP/NK1R w regulacji wielu procesów fizjologicznych i patologicznych, co potwierdzone zostało przez liczne badania.

Substancja P może kontrolować procesy komórkowe za pomocą mechanizmów autokrynych, endokrynych, parakrynych oraz neuroendokrynych [49,50].

W ośrodkowym układzie nerwowym SP pełni funkcję neuroprzekaźnika i neuromodulatora. Może wpływać na uwalnianie innych neuroprzekaźników, m.in. dopaminy, acetylocholiny, kwasu γ -aminomasłowego (GABA) stymulować uwalnianie hormonów wzrostu i prolaktyny, uczestniczyć w regulacji homeostazy [5]. Istnieją dowody na udział SP w ośrodkowej regulacji oddychania i krążenia poprzez oddziaływanie na receptory NK1 w mózgowiu i naczyniach krwionośnych [51,52]. SP może być także zaangażowana w funkcjonowanie narządów wzroku, słuchu oraz węchu [52,53]. Układ SP/NK1R odgrywa istotną rolę w regulacji procesów emocjonalnych, mogąc przyczyniać się do zaburzeń emocjonalnych takich jak lęk, stres, alkoholizm, nerwica czy depresja [52,54]. Wykazano, że poziom SP jest podwyższony w przypadku depresji, schizofrenii, choroby afektywnej dwubiegunowej i fibromialgii [22,48]. Ponadto, SP zaangażowana jest w procesy metaboliczne komórek nerwowych — przeżywanie i degenerację neuronów [49,52]. Badania potwierdzają znaczenie SP w etiologii chorób neurodegeneracyjnych, m.in. Alzheimer, Parkinsona, Huntingtona [22]. Co więcej, układ SP/NK1R znajduje się w rejonach rdzenia przedłużonego (jądro pasma samotnego), biorących udział w inicjowaniu odruchu wymiotnego. Stymulacja NK1R przez SP prowadzi do pobudzenia neuronów w ośrodku wymiotnym i wywołania wymiotów [55]. Stwierdzono również znaczenie SP w migrenowych i klastrowych bólach głowy [22]. Zarówno SP jak i NK1R regulowane są

w górę w procesach zapalnych. W urazowym uszkodzeniu mózgu i udarze mózgu SP zaangażowana jest w neurogenne mechanizmy zapalne, przyczynia się do zwiększenia przepuszczalności naczyń krwionośnych, zwiększenia przepływu krwi i wytwarzania obrzęku mózgu, nasilenia stanu zapalnego i uszkodzenia tkanki mózgowej [56].

Wiadomo, że SP i receptor NK1 biorą udział w procesach odczuwaniu bólu (nocycepcji). W obrębie rdzenia kręgowego oraz części obwodowej układu nerwowego SP/NK1R uczestniczy w transmisji i modulacji bodźców bólowych. W zdrowych tkankach wydzielanie SP jest niewielkie, stymulowane tylko przez silne bodźce. W tkankach poddanych procesom chorobowym, w wyniku działania czynnika bólowego w miejscu uszkodzenia obserwowane są wzrost stężenia SP oraz zwiększona ekspresja receptorów NK1, które uczestniczą w przekazywaniu sygnału bólowego do wyższych piętér układu nerwowego, co przyczynia się do nasilenia odczuwania bólu [52,57]. Zasadniczo SP jest więc uważana za czynnik pronocyceptywny, choć w niektórych sytuacjach wykazano działanie przeciwbólowe SP lub jej metabolitów [58].

Znany jest także udział SP/NK1R w funkcjonowaniu układu immunologicznego. SP, która znajduje się w włóknach unerwiających narządy limfatyczne, działa jako mediator pomiędzy układem immunologicznym a nerwowym [49,59]. Ponadto, SP i NK1R są ekspresjonowane na różnych typach komórek odpornościowych, m.in: komórkach dendrycznych, monocytach, makrofagach, neutrofilach, limfocytach T. Podczas stanu zapalnego ich ekspresja jest zwiększona [22]. Wiele badań potwierdza rolę SP i NK1R w odpowiedziach immunologicznych i ich wpływ na nasilenie reakcji zapalnych. SP poprzez oddziaływanie z NK1R aktywuje czynnik NF- κ B, stymuluje uwalnianie cytokin prozapalnych, takich jak interleukina-1 (IL-1), interleukina-6 (IL-6) oraz czynnik martwicy nowotworów TNF- α [16,59]. Poza tym, SP może wpływać na chemotaksję i migrację leukocytów do miejsc stanu zapalnego, różnicowanie limfocytów B, proliferację limfocytów T, zwiększone uwalnianie immunoglobulin [52]. W wyniku tych działań SP i NK1R mogą modulować funkcjonowanie układu immunologicznego i wpływać na odpowiedź immunologiczną podczas infekcji i chorób zapalnych oraz autoimmunologicznych. Równocześnie, SP poprzez promowanie uwalniania cytokin prozapalnych i czynników wzrostu może być zaangażowana w procesy gojenia, naprawy i regeneracji skóry [60].

Istnieją dane sugerujące rolę SP/NK1R w przebiegu różnych zakażeń bakteryjnych i wirusowych, poprzez stymulowanie reakcji zapalnych i nasilenie odpowiedzi immunologicznej organizmu. Uważa się, że układ ten może odgrywać rolę w zakażeniach wirusowych, takich jak ludzki wirus niedoboru odporności (HIV), wirus opryszczki,

syncytialny wirus oddechowym (RSV), wirusowe zapalenie mózgu, wirusowe zapalenia mięśnia sercowego, wirusowe zapalenia wątroby typu C oraz w zakażeniach bakteryjnych, takich jak sepsa, zapalenie opon mózgowych [49,59].

Układ SP/NK1R zaangażowany jest również w patogenezę chorób zapalnych, m.in. reumatoidalnego zapalenia stawów (RZS) [9]. W tkankach stawowych pacjentów chorych na RZS obserwuje się zwiększoną ekspresję SP i NK1R. SP może stymulować uwalnianie cytokin prozapalnych i czynników wzrostu, wpływać na migrację komórek zapalnych, co przyczynia się do rozwoju zapalenia stawów.

SP/NK1R są kluczowymi mediatorami w wielu procesach fizjologicznych i patologicznych zachodzących w układzie pokarmowym. Są one zaangażowane w regulację motoryki żołądka i jelit oraz stymulację wydzielania wody i jonów [61]. W tkankach objętych stanem zapalnym zwiększają przepuszczalność naczyń krwionośnych, uczestniczą w procesie rozwoju i nasilenia zapalenia poprzez stymulowanie uwalniania cytokin prozapalnych [22]. SP i NK1R odrywają istotną rolę w patogenezie zapalnych chorób przewodu pokarmowego, takich jak zapalenie trzustki, zapalenie wątroby, wrzodziejące zapalenie jelita grubego oraz choroba Leśniewskiego-Crohna [48]. W wyniku nadmiernego lub nieprawidłowego działania SP/NK1R dochodzi do zaburzeń homeostazy i przewlekłych stanów zapalnych, co prowadzi do uszkodzeń tkanek i narządów przewodu pokarmowego.

SP i NK1R (wraz z NKA i NK2R) uczestniczą w szeregu funkcji biologicznych związanych z układem oddechowym. Pojawienie się w drogach oddechowych czynników szkodliwych i drażniących skutkuje zwiększonym wydzieleniem SP i NKA. Neuropeptydy te pośredniczą w regulowaniu napięcia mięśni gładkich dróg oddechowych (powodują silne działanie skurczowe mięśni oskrzeli) oraz hipersekcji śluzu. SP może zwiększać przepuszczalność bariery naczyniowej i stymulować uwalnianie cytokin prozapalnych i innych mediatorów zapalnych, przyczyniając się do nasilenia odpowiedzi zapalnej [19,62,63]. Wiadomo też, że nadekspresja receptorów NK1 i NK2 towarzyszy astmie oskrzelowej oraz przewlekłej obturacyjnej chorobie płuc (POChP) [62].

Ustalono również znaczenie NK1R/SP w układzie sercowo-naczyniowym i chorobach z nim związanych. System ten pełni rolę mediatora agregacji płytek krwi i tworzenia skrzeplin [64,65]. SP bierze udział w regulacji częstości akcji serca, ciśnienia krwi oraz rozszerzaniu naczyń krwionośnych. Dodatkowo, odgrywa ważną funkcję w niedokrwieniu i reperfuzji serca oraz w odpowiedzi układu sercowo-naczyniowego na stres [51].

W układzie moczowo-płciowym SP poprzez działanie na receptory NK1R pośredniczy w regulacji mięśni gładkich pęcherza moczowego i moczowodu, wpływając na

skurcz lub rozkurcz mięśni. Ponadto, SP może pośredniczyć w procesach zapalnych w układzie moczowym, prowadząc do zwiększonej przepuszczalności naczyń krwionośnych i migracji leukocytów do obszarów zapalnych [61].

SP oraz NK1R odgrywają istotną rolę w procesie powstawania nowotworów. Badania wykazały, że receptory są nadeksprimowane na powierzchni różnych typów komórek nowotworowych, w tym białaczki, glejaka, neuroblastomy, czerniaka, raka piersi, jajnika, prostaty, płuc, trzustki, żołądka, okrężnicy i tarczycy [19,66]. Uznaje się, że ilość receptorów NK1R na powierzchni komórek rakowych koreluje ze stopniem złośliwości nowotworu [67]. W komórkach nowotworowych, SP poprzez działanie na NK1R indukuje mitogenezę oraz migrację komórek nowotworowych. Ponadto, SP promuje proces tworzenia nowych naczyń krwionośnych w guzie nowotworowym (angiogenezę), umożliwiając dostarczenie do niego tlenu i składników odżywczych oraz wzmacnia szybkość glikolizy komórek nowotworowych, zwiększając metabolizm komórek nowotworowych [22]. Dodatkowo, SP/NK1R hamują proces apoptozy komórek nowotworowych, co prowadzi do zwiększenia ich przeżywalności [22].

II.2. Chemia medyczna receptora NK1

II.2.1. Receptor NK1R jako cel molekularny terapii

Wszechstronne zaangażowanie układu SP/NK1R w funkcje fizjologiczne i rozwój patologii w organizmie ludzkim stymulowało zainteresowanie receptorem NK1 jako potencjalnym celem molekularnym dla farmakoterapii. Na przestrzeni lat, począwszy od lat 80. XX wieku w wielu ośrodkach naukowych i firmach farmaceutycznych prowadzone były intensywne badania nad poszukiwaniem związków hamujących działanie Substancji P – antagonistów receptora NK1. Antagonistą receptora jest ligand, który wykazuje zdolność do blokowania lub osłabiania odpowiedzi biologicznej wywoływanej przez agonistę. Część cząsteczka antagonisty charakteryzuje się powinowactwem do receptora, jednocześnie nie wykazując aktywności wewnętrznej, czyli nie inicjuje sygnalizacji wewnątrzkomórkowej.

Do chwili obecnej do praktyki klinicznej wprowadzono pięć substancji aktywnych, będących antagonistami receptora NK1 (aprepitant, fosaprepitant – prolek aprepitantu, rolapitant, netupitant i fosnetupitant – prolek netupitantu). Amerykańska Agencja Żywności i Leków (US FDA) oraz Europejska Agencja Leków (EMA) zatwierdziły je jako leki we wskazaniu do leczenia wymiotów i nudności związanych z chemioterapią przeciwnowotworową oraz występujących w czasie pooperacyjnym [68]. Inny antagonist NK1R, maropitant, jest zatwierdzonym lekiem przeciwwymiotnym w medycynie weterynaryjnej.

Antagonisty NK1R cieszyły się także znacznym zainteresowaniem wśród badaczy jako potencjalna farmakoterapia przeciwbólowa. Tutaj jednak pomimo obiecujących wyników badań przedklinicznych, które wskazywały na działanie przeciwbólowe antagonistów NK1R, badania kliniczne przeprowadzane na ludziach wykazały, że substancje te nie są dostatecznie skuteczne w leczeniu różnego typu bólu [49,69]. Wątek celowania w NK1R w leczeniu bólu nadal pojawia się w literaturze, ostatnio przy poszukiwaniu wielocelowych związków przeciwbólowych (rozdział II.2.3.1) oraz w próbach antagonizowania endosomalnego sygnałowania receptora NK1 (rozdział II.2.3.2).

Prowadzono także szereg badań nad zastosowaniem antagonistów NK1R jako potencjalnych środków terapeutycznych w leczeniu m.in. stanów lękowych, depresji, uzależnień od alkoholu, migreny, zakażeń wirusowych HIV. Badania kliniczne wykazywały jednak, że antagonisty NK1R są nieskuteczne lub nie wykazują wyższej skuteczności niż dostępne terapie w leczeniu tych schorzeń [49]. W kilku przypadkach, wyniki były

niejednoznaczne. Przykładowo, jeśli chodzi o leczenie depresji, w badaniach klinicznych w zależności od podawanej dawki aprepitant wywierał działanie przeciwdepresyjne (dawka 300 mg na dobę, przez 45 dni) [70] lub wykazywał brak skuteczności (dawka 160 mg na dobę, przez 56 dni) [71]. Obecnie brak jest (publicznie dostępnych) doniesień o kontynuowaniu prac nad antydepresyjnym działaniem antagonistów NK1R.

Badanie kliniczne związku LY-686017 (tradipitant; inny kod VLY-686) w leczeniu uzależnień od alkoholu pokazało, że związek ten wywiera korzystny wpływ na samopoczucie i tłumi spontaniczne pragnienie alkoholu [72]. Nie wiadomo jednak czy związek ten jest dalej rozwijany pod kątem aplikacyjnym w tym wskazaniu. Tradipitant był także badany (małe badanie na ludziach) pod kątem zastosowania w leczeniu uzależnienia od opioidów, ale nie wykazano jego skuteczności [73].

W ostatnich latach otrzymano obiecujące wyniki w związku z użyciem antagonistów NK1R (m.in. serlopitantu, aprepitantu lub tradipitantu) w leczeniu przewlekłego świądu i chorób skórnych u ludzi [74–76]. Aktualnie prowadzone są badania kliniczne w celu zastosowania tych antagonistów NK1R w terapii przeciwświądowej [76–78].

Ostatnio rozwijane jest także zastosowanie antagonistów NK1R w leczeniu przewlekłego kaszlu. Niedawno opublikowano pozytywne wyniki badań II fazy dla orvepitantu [79]; inne badanie kliniczne dla tego związku trwa [80].

Ponadto, wiele związków będących antagonistami NK1R zostało opatentowanych w możliwym zastosowaniu klinicznym do leczenia, m.in. infekcji bakteryjnych, kardiomiopatii, zaburzeń motoryki żołądka, melanogenezy, czy chorób układu oddechowego [48]. Wiele badań przedklinicznych poświęcono przeciwnowotworowym właściwościom antagonistów NK1R (rozdział II.2.3.3).

Na marginesie warto dodać, że ostatnio (w 2023 roku) FDA zatwierdziła małowcząsteczkowego antagonistę NK3R, fezolinetant, do zastosowania w leczeniu uderzeń gorąca związanych z menopauzą [81]. Jest to pierwsze zatwierdzenie w obszarze modulacji układu tachykininowego niezwiązane z receptorem NK1R.

II.2.2. Chemotypy antagonistów NK1R

Z chemicznego punktu widzenia antagonisty receptora NK1 dzielą się na dwie główne grupy – antagonistów peptydowych, będących pochodnymi lub analogami Substancji P oraz antagonistów niepeptydowych, małowcząsteczkowych, syntetycznych związków o zróżnicowanej strukturze chemicznej.

II.2.2.1. Peptydowe antagonisty receptora NK1

Poszukiwanie peptydowych antagonistów NK1R cechowało wczesną fazę zainteresowania farmakologiczną modulacją tego receptora. W naturalny sposób, większość pierwszych prac nad peptydowymi antagonistami NK1R koncentrowała się na modyfikowaniu struktury endogennego agonisty, czyli Substancji P. Sekwencje niektórych omawianych dalej w tekście peptydów przedstawione są w Tabeli 2.

Tabela 2. Sekwencje wybranych peptydowych antagonistów NK1R w porównaniu do Substancji P, na podstawie [82].

	Nazwa związku	Sekwencja aminokwasowa **
	Substancja P (SP)	Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH ₂
1	[D-Pro ² , D-Trp ^{7,9}]SP	Arg- D-Pro -Lys-Pro-Gln-Gln- D-Trp -Phe- D-Trp -Leu-Met-NH ₂
2	Spantyd I	D-Arg -Pro-Lys-Pro-Gln-Gln- D-Trp -Phe- D-Trp -Leu-Leu-NH ₂
3	Spantyd II	D-Lys(Nic) -Pro- Pal(3) -Pro- D-Phe(Cl₂) - Asn-D-Trp -Phe- D-Trp -Leu-Nle-NH ₂
4	[D-Pro ⁹ , <i>trans</i> -β-benzyl-Pro ¹⁰ , Trp ¹¹]SP	Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe- Pro-(<i>trans</i>-β-benzyl)Pro-Trp -NH ₂
5	[D-Pro ⁴ , D-Trp ^{7,9,10} , Phe ¹¹]SP(4-11)	D-Pro -Gln-Gln- D-Trp -Phe- D-Trp-D-Trp -Phe-NH ₂
6	[Orn ⁶ , AspOBzl ¹¹]SP(6-11)-OBzl	Orn -Phe-Phe-Gly-Leu- Asp(OBzl)-OBzl
7	Sendyd ([Tyr ⁶ , D-Phe ⁷ , D-His ⁹]SP(6-11))	Tyr-D-Phe-Phe-D-His -Leu-Met-NH ₂
8	[D-Trp ⁷]Sendyd	Tyr-D-Trp -Phe- D-His -Leu-Met-NH ₂
9	[Gln ⁶]Sendyd	Gln- D-Phe-Phe-D-His -Leu-Met-NH ₂

** wytuszczono zmienione reszty aminokwasowe względem Substancji P; D-Lys(Nic) – N^ε-nikotynoilo-D-lizyna, Pal(3) – 3-(3-pirydylo)alanina, D-Phe(Cl₂) – 3,4-dichloro-D-fenylalanina, Orn – ornityna, Bzl – grupa benzylowa

Pierwsza zastosowana strategia polegała na wprowadzaniu reszt D-aminokwasowych. Badania dotyczące analogów SP wykazały, że zamiana stereochemii aminokwasów z izomerów L na D w określonych pozycjach sekwencji SP powoduje zmianę działania na antagonistyczne [5]. Jednym z pierwszych uzyskanych w ten sposób peptydów o działaniu antagonistycznym był [D-Pro², D-Trp^{7,9}]SP (Tabela 2, poz. 2). Związek ten wykazywał powinowactwo do receptorów NK1 szczura, królika, myszy i kawii domowej, ale jednocześnie powodował znaczną neurotoksyczność [83]. W toku dalszych prac odkryto Spantyd I (Tabela 2, poz. 3), różniący się od poprzedniego analogu zmianą chiralności Arg w pozycji pierwszej (z L na D), Pro w pozycji drugiej (z D na L) oraz wprowadzeniu na

C-końcu Leu w miejsce Met [84]. Spantyd I był pełnym antagonistą, ale nie wykazywał wysokiego powinowactwa receptorowego. Ponadto, obserwowano jego znaczącą neurotoksyczność oraz powodowanie degranulacji komórek tucznych [85]. W rezultacie dalszych badań zależności struktura chemiczna–aktywność biologiczna (ang. *Structure-Activity Relationship*, SAR) uzyskano silnego antagonistę NK1R, charakteryzującego się niską neurotoksycznością – Spantyd II [86] (Tabela 2, poz. 3). Sekwencja aminokwasowa Spantynu II względem Substancji P zmieniona jest w siedmiu pozycjach. W odróżnieniu od Spantynu I, D-Arg w pozycji pierwszej oraz Lys w pozycji trzeciej (prawdopodobnie odpowiedzialne za działanie neurotoksyczne) zostały zastąpione odpowiednio N^c-nikotynoilo-D-lizyną (D-Lys(Nic)) oraz 3-(3-pirydylo)alaniną (Pal3). Wprowadzono także 3,4-dichloro-D-feniloalaninę (D-Phe(Cl₂)) w pozycji piątej, Asn w pozycji szóstej oraz norleucynę (Nle) w pozycji jedynastej, zachowując D-Trp w pozycji siódmej i dziewiątej. Uważa się, że zmiana Gln na Asn wpływa na zwiększenie powinowactwa [82]. Wprowadzone modyfikacje pozwoliły na uzyskanie silnego antagonisty o 10-krotnie wyższym powinowactwie do NK1R niż Spantyd I [87] oraz niskiej neurotoksyczności [86].

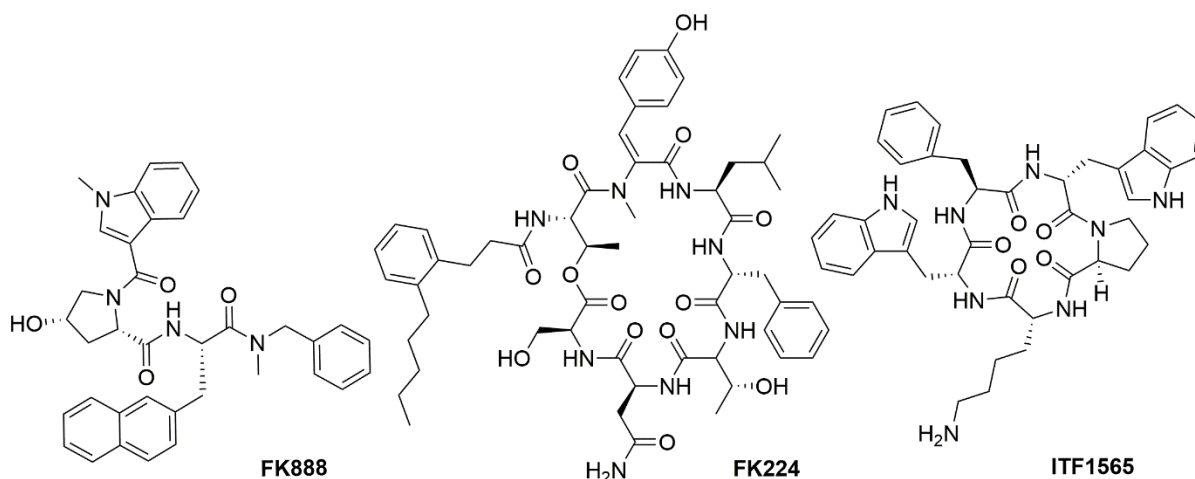
W innym podejściu podjęto próbę modyfikacji trzech ostatnich reszt aminokwasowych w SP (pozycje 9-11). W wyniku tych prac otrzymano serię antagonistów, z których najsilniejszym jest związek [D-Pro⁹, *trans*-β-benzylo-Pro¹⁰, Trp¹¹]SP (Tabela 2, poz. 4). Związek ten zawiera Pro podstawioną grupą benzylową w pozycji β pierścienia (*trans*-β-benzylo-Pro¹⁰). Modyfikacja ta umożliwiła uzyskanie silnego antagonisty o powinowactwie w zakresie nanomolowym [88].

Inna strategia poszukiwania antagonistów NK1R zakładała modyfikowanie łańcucha peptydowego SP poprzez jednoczesne skracanie go oraz dokonywanie zmian w sekwencji aminokwasowej. Zastosowanie tej strategii pozwoliło m.in. na odkrycie selektywnego i silnego antagonisty NK1R oktapeptydu [D-Pro⁴, D-Trp^{7,9,10}, Phe¹¹]SP(4-11) (Tabela 2, poz. 5), który posiada łańcuch skrócony od N-końca, D-Pro w pozycji czwartej, D-Trp w pozycji siódmej, dziewiątej i dziesiątej oraz Phe w pozycji jedenastej, która prawdopodobnie jest kluczowa dla wysokiego powinowactwa do receptora [82]. Silny antagonizm NK1R wykazywał także heksapeptyd [Orn⁶, Asp(OBzl)¹¹]SP(6-11)-OBzl (Tabela 2, poz. 6), który w pozycji szóstej zawiera zasadową ornitynę (Orn) oraz podwójnie benzylowany kwas asparaginowy (Asp(OBzl)-OBzl) na C-końcu łańcucha peptydowego. Innymi heksapeptydowymi antagonistami NK1R, są Sendyd ([Tyr⁶, D-Phe⁷, D-His⁹]SP(6-11)) (Tabela 2, poz. 7) oraz jego analogi m.in. [D-Trp⁷]Sendyd (Tabela 3, poz. 8) i [Gln⁶]Sendyd (Tabela 2, poz. 9) [82].

Przeprowadzono również badania mające na celu identyfikację minimalnego fragmentu o właściwościach antagonistycznych. W tym celu zsyntezowano serie związków, powstałych w wyniku delecji kolejnych reszt aminokwasowych od końca aminowego i/lub karboksylowego. Wysiłki te doprowadziły do odkrycia, że zabezpieczony tripeptyd Boc-Gln-D-Trp(CHO)-Phe-OBzl, który posiada grupy zabezpieczające na N-końcu (grupę *tert*-butoksykarbonylową, Boc) i C-końcu (grupę benzylową, Bzl) oraz ugrupowanie formylowe na indolowym atomie azotu tryptofanu charakteryzuje się wysokim powinowactwem do NK1R. Zsugerowano, że związek ten zawiera kluczową sekwencję aminokwasów Gln-D-Trp-Phe, zapewniającą silne wiązanie się z receptorem [85]. W wyniku dalszych modyfikacji uzyskano tripeptyd FR113680 [Ac-Thr-D-Trp(CHO)-PheNMeBzl] zawierający na końcu aminowym *N*-acetylo-L-treoninę, na karboksylowym zaś *N*-benzylo-*N*-metyloamid. Związek ten jest selektywnym antagonistą i wykazuje silnie powinowactwo do NK1R, a także jest aktywny *in vivo* po podaniu dożylnym [85,89]. W odpowiedzi na niekorzystne właściwości fizykochemiczne i niską biodostępność, prowadzono dalsze prace, które doprowadziły do odkrycia dipeptydu FK888 (Rysunek 3), będącego selektywnym antagonistą NK1R o wysokim powinowactwie. FK888 zawiera 4-hydroksyprolinę, w której grupa aminowa zacylowana jest kwasem *N*-metylo-indolilo-3-karboksylowym oraz naftyloalaninę, występującą w formie *N*-benzylo-*N*-metyloamidu [90]. Przypuszczano, że obecność reszty hydroksyproliny wpływa na przyjęcie odpowiedniej orientacji przestrzennej pierścieni aromatycznych w cząsteczce, niezbędnej do wysokiego powinowactwa do receptora NK1R. Związek ten jest aktywny zarówno w testach *in vitro* i *in vivo*, po podaniu doustnym i dożylnym [90–92]. Poszukiwanie peptydowych antagonistów NK1R obejmowało także wykorzystanie peptydów cyklicznych. Interesującym przedstawicielem tej grupy jest FK224, cykloheptapeptyd (Rysunek 3), który zawiera w swojej strukturze mostek laktonowy między grupą hydroksylową reszty treoniny w pozycji pierwszej a grupą karboksylową seryny w pozycji siódmej, co przyczynia się do stabilizacji struktury oraz utrzymania odpowiedniej geometrii cząsteczki. FK224 wykazuje jednocześnie wysokie powinowactwie do NK1R i NK2R (nie jest selektywny), co może być użyteczne w niektórych procesach terapeutycznych [85,93]. Selektywnym i silnym antagonistą NK1R jest analog SP, GR71251 [D-Pro⁹(spiro- γ -laktam)Leu¹⁰, Trp¹¹]SP [94]. Związek ten posiada spirocykliczny fragment pomiędzy D-Pro w pozycji dziewiątej i Leu w pozycji dziesiątej, powodujący ograniczenia konformacyjne. Znany przedstawicielem tej grupy jest także cykliczny pentapeptyd ITF1565 c[D-Trp-Pro-D-Lys-D-Trp-Phe] (Rysunek 3), wykazujący działanie antagonistyczne wobec receptora NK1 [85].

Pomimo szerokiej skali prac oraz postępów w badaniach SAR, antagoniści peptydowi wykazywali właściwości ograniczające ich zastosowanie w badaniach klinicznych. Podobnie jak większość związków peptydowych, syntetyczni antagoniści cechowali się niedostateczną biodostępnością po podaniu doustnym, niską stabilnością metaboliczną, niezdolnością do przekraczania bariery krew-mózg (BBB, ang. *blood-brain barrier*) oraz często ograniczoną rozpuszczalnością i siłą działania. W związku z tym dalsze prace skupiały się na identyfikacji małowcząsteczkowych antagonistów NK1R.

Warto jednak zauważyć, że w ostatnich latach niektóre peptydowe antagoniści NK1R i peptydowe analogi SP ponownie znajdują zastosowanie, zwłaszcza w badaniach nad celowanym dostarczaniem ładunku terapeutycznego (ang. *therapeutic cargo*), o czym mowa będzie w rozdziale II.3.4 i II.3.5.



Rysunek 3. Wybrane peptydowe/peptydomimetyczne antagoniści NK1R.

II.2.2.2. Małowcząsteczkowe antagoniści NK1R

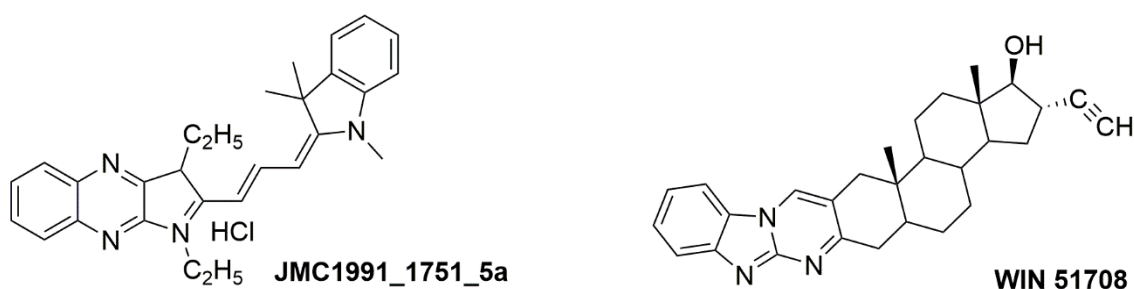
Podczas gdy zespoły akademickie skupiały się głównie na badaniach nad peptydowymi antagonistami receptora NK1, to wraz z identyfikacją pierwszych mniejszych peptydowych antagonistów i małowcząsteczkowych antagonistów NK1R, wysiłki firm farmaceutycznych skoncentrowały się głównie wokół tych ostatnich. Niniejszy rozdział przedstawia przegląd znanych małowcząsteczkowych antagonistów NK1R z różnych grup chemicznych oraz próby ich zastosowania w farmakologii.

II.2.2.2.1 Imidazochinoksaliny i heterosteroidy

Wśród pierwszych niepeptydowych antagonistów NK1R były imidazochinoksaliny odkryte i opisane przez badaczy z firmy Eastman Kodak [95]. Jednym z lepiej

scharakteryzowanych związków w tej grupie jest analog JMC1991_1751_5a (Rysunek 4), o umiarkowanym powinowactwie do szczurzego receptora NK1 ($IC_{50} = 510$ nM). Pomimo przeprowadzonych badań SAR nie udało się opracować związku o znacząco lepszym powinowactwie do receptora NK1. Co więcej, związki z tej serii okazały się mieć działanie toksyczne, powodując w badaniach *in vivo* spadek napięcia mięśniowego, trudności w oddychaniu i sinicę (cyjanozę) [85].

W wyniku dalszych poszukiwań odkryto serię heterosteroidów i związek WIN 51078 (Rysunek 4), o stokrotnie lepszym powinowactwie do szczurzego receptora NK1 ($IC_{50} = 50$ nM) [96]. Mimo, że zsyntezowano wiele pochodnych tej struktury to nie otrzymano związku o lepszej aktywności biologicznej.



Rysunek 4. Wybrane antagonisty NK1R oparte na strukturze imidazochinoksaliny i heterosteroidów.

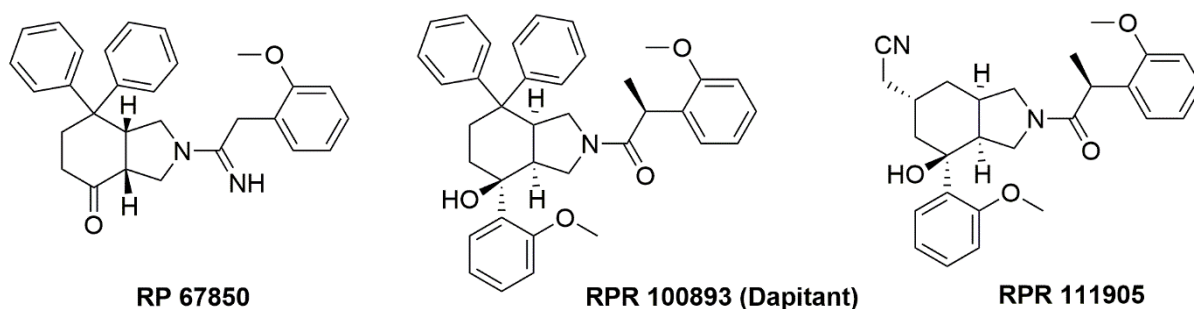
II.2.2.2 Perhydroizoindolony

W firmie Rhône-Poulenc opracowano antagonisty NK1R wykorzystujące strukturę perhydroizoindolonu jako szkielet cząsteczkowy. Wczesnym przedstawicielem tej grupy jest pochodna 2-aryloacetylo-7,7-difenyloperhydroizoindolonu, związek RP 67580 (Rysunek 5) [97]. Analiza SAR wykazała, że kluczowe znaczenie dla wysokiego powinowactwa do NK1R ma konfiguracja *R* w 3 i 7 pozycji pierścienia perhydroizoindolonu (3*aR*, 7*aR*), ponieważ izomer 3*aS*, 7*aS* okazał się pozbawiony aktywności. RP 67580 został scharakteryzowany jako selektywny antagonist NK1R, o wysokim powinowactwie do szczurzego i mysiego receptora NK1, zarówno *in vitro* i *in vivo*. W dalszych badaniach wykazano, że związek ten wiąże się znacznie słabiej do ludzkich receptorów NK1 (około 25-razy słabiej, $IC_{50} = 49$ nM) [9,97].

W wyniku dalszych prac odkryto związek należący do grupy 7,7,4-triaryloperhydroizoindol-4-oli – RPR 100893, znany również jako dapitant (Rysunek 5) [98]. Związek ten posiada dodatkowy czwarty pierścień arylowy, zmodyfikowaną grupę karbonylową, zastąpione ugrupowanie amidynowe przez amid oraz dodatkową grupę

metylową w porównaniu do RP 67580. RP 100893 charakteryzuje się wysokim powinowactwem do ludzkiego receptora NK1 ($IC_{50} = 13 \text{ nM}$) i około 50-razy niższym do szczurzego receptora NK1 [98]. W przeciwieństwie do RP 67580, enancjomer 3a*S*, 7a*S* jest bardziej aktywny. Związek ten wykazywał aktywność *in vivo* w modelach bólu i zapalenia oraz charakteryzował się dobrą biodostępnością po podaniu doustnym. RPR 100893 był testowany w badaniach klinicznych do zastosowania w leczeniu migrenowych bólów głowy, ale okazał się nieskuteczny [99].

Kolejne prace zaowocowały powstaniem związku RPR 111905 (Rysunek 5), który nie posiada ugrupowania difenylowego oraz ma wprowadzony polarny podstawnik w pozycji piątej pierścienia perhydroizoindolonu [100]. Związek ten charakteryzuje się wysokim powinowactwem ($IC_{50} = 0.3 \text{ nM}$), dobrą aktywnością po podaniu doustnym, jednak wykazuje ograniczoną penetrację do ośrodkowego układu nerwowego [101].



Rysunek 5. Wybrane antagonisty NK1R oparte na strukturze perhydroizoindolonu.

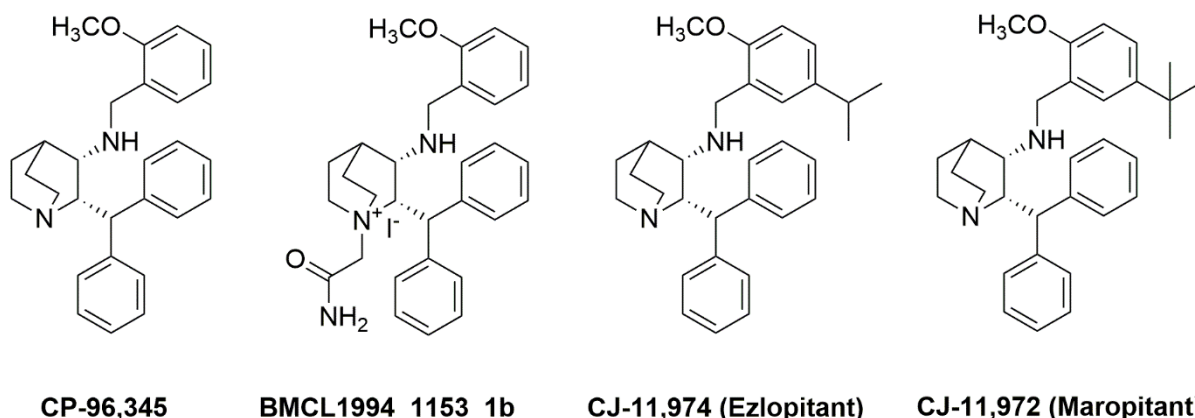
II.2.2.2.3 3-(*N*-benzyloamino)chinuklidyny

Naukowcy z firmy Pfizer przedstawili antagonistę NK1R opartego na strukturze 3-(*N*-benzyloamino)chinuklidyny o nazwie CP-96,345 (Rysunek 6) [102]. Związek ten składa się ze sztywnego rusztowania chinuklidynowego, które zawiera zasadowy atom azotu, ugrupowania benzhydrylowego i grupy *N*-(*o*-metoksybenzylo)aminowej. CP-96,345 ma wysokie powinowactwo do ludzkich NK1R ($IC_{50} = 0.77 \text{ nM}$) i bydłecych ($IC_{50} = 3.4 \text{ nM}$) oraz około 70 razy niższe do szczurzych i mysich NK1R. Badania SAR w ramach tej serii chinuklidyn wykazały kluczowe znaczenie stereochemii 2*S*, 3*S* oraz obecności grupy *ortho*-metoksy w łańcuchu bocznym *N*-benzyloaminy [103]. Usunięcie grupy metoksy skutkowało 20-krotnym spadkiem powinowactwa. Wykazano, że CP-96,345 jest aktywny *in vivo* po podaniu doustnym w modelach wynaczynienia białek osocza wywołanych przez SP i kapsaicynę oraz w modelu wydzielania śliny wywołanym przez SP [103]. W wyniku dalszych badaniach ustalono, że CP-96,345 w znacznym stopniu oddziałuje z kanałami

wapniowymi, co prowadzi do wystąpienia różnorodnych, niepożądanych działań ogólnoustrojowych niezwiązanych z blokowaniem NK1R, m.in. bradykardię oraz hipotensję [85,104]. Tę niepożądaną aktywność przypisano obecności silnie zasadowego azotu chinuklidyny.

Opracowano także serię *N*-alkilowanych analogów (z czwartorzędowym azotem) [105]. Dobre powinowactwo tych analogów pozwoliło odrzucić wcześniejszą hipotezę o istotnej roli ewentualnego oddziaływania jonowego pomiędzy chinuklidynowymi antagonistami a receptorem. W tamtych pracach otrzymano związek zawierający ugrupowanie acetamidowe (BMCL1994_1153_1b, Rysunek 6), który utrzymywał wysokie powinowactwo do ludzkiego NK1R ($IC_{50} = 1.3 \text{ nM}$), ale nie podano aktywności tych związków w kanałach wapniowych.

Z omawianym wcześniej CP-96,345 blisko strukturalnie związane są jeszcze związki CJ-11,974 (ezlopitant, Rysunek 6) oraz CJ-11,972 (maropitant, Rysunek 6), w których pierścień metoksybenzylowy CP-96,345 podstawiony jest grupą izopropylową (CJ-11,974) lub grupą *tert*-butylową (CJ-11,972) [9]. Ezlopitant przeszedł testy w II fazie badań klinicznych w leczeniu wymiotów spowodowanych chemioterapią. Wykazał dobrą tolerancję i skutecznie kontrolował wymioty, jednak stwierdzono mniejszą aktywność w łagodzeniu nudności, co spowodowało przerwanie badań. Ponadto, przeprowadzono ocenę skuteczności ezlopitantu u pacjentów z zespołem jelita drażliwego (IBS). Pomimo obiecujących wyników, badania w tym wskazaniu również nie zostały kontynuowane [9,106]. Maropitant jest lekiem w medycynie weterynaryjnej stosowanym do leczenia wymiotów pooperacyjnym u psów i kotów [107].



Rysunek 6. Wybrane antagonisty NK1R oparte na strukturze 3-(*N*-benzyloamino)chinuklidyny.

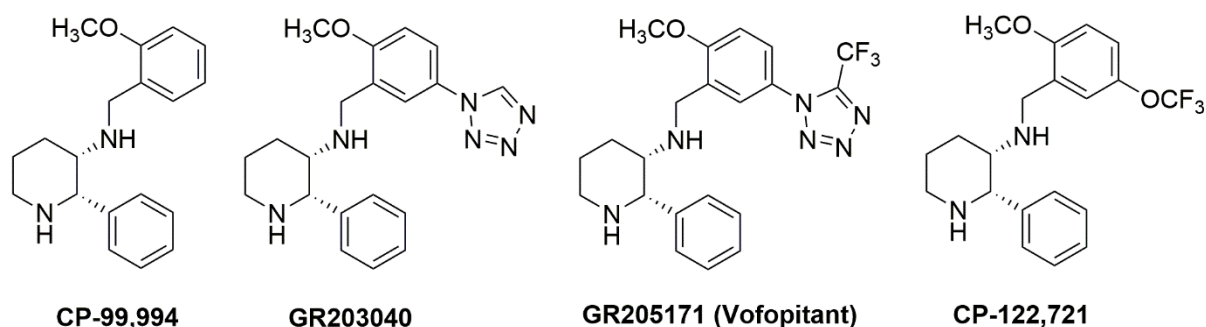
II.2.2.2.4 3-(N-benzyloamino)piperydyny

W wyniku dalszych prac nad uproszczeniem struktury CP-96,345 opracowano serię związków 3-(N-benzyloamino)piperydyn. Pierwszy związek z tej grupy, CP-99,994 (Rysunek 7), został uzyskany poprzez zamianę pierścienia chinuklidynowego na pierścień piperydynowy oraz ugrupowania benzhydrylowego na grupę fenyłową [108]. CP-99,994 zachowuje wysokie powinowactwo do ludzkiego NK1R ($IC_{50} = 0.17$ nM), jest selektywny *in vitro* oraz ma znacznie obniżoną aktywność wobec kanałów wapniowych w porównaniu do CP-96,345 [109]. CP-99,994 był poddawany badaniom klinicznym w fazie II w celu oceny jego skuteczności w leczeniu bólu zębów u pacjentów po zabiegach chirurgicznych [110]. Jednak badania zostały przerwane z powodu słabej biodostępności tego związku. CP-99,994 był przedmiotem licznych badań SAR, mających na celu określenie strukturalnych wymagań niezbędnych do uzyskania związków o wysokim powinowactwie do receptora NK1.

Badania nad poprawą właściwości farmakokinetycznych CP-99,994 oraz zwiększenia dostępności po podaniu doustnym w firmie Glaxo doprowadziły do opracowania serii analogów CP-99,994 podstawionych w pozycji piątej pierścienia benzyłowego (pozycja *para* względem grupy metoksy) [101]. Związek podstawiony pierścieniem tetrazolowym, GR203040 (Rysunek 7), charakteryzował się najlepszym powinowactwem do ludzkiego NK1R na poziomie subnanomolowym ($IC_{50} = 0.06$ nM) [111]. Wprowadzenie pierścienia tetrazolowego spowodowało polepszenie właściwości farmakokinetycznych, zwiększoną stabilność metaboliczną oraz lepszą dostępnością po podaniu doustnym. Ponadto, GR203040 wykazywał aktywność w hamowaniu wymiotów u fretek wywołanych szerokim zakresem czynników emetogennych. Dalsze modyfikacje doprowadziły do opracowania również silnego antagonisty, związku GR205171 (Rysunek 7), zwanego także vofopitantem. Związek ten zawiera pierścień tetrazolowy podstawiony grupą trifluorometylową. GR205171 wykazywał działanie przeciwwymiotne w trzy razy niższych dawkach niż GR203040 [112]. GR205171 był testowany w fazie II badań klinicznych w leczeniu nudności i wymiotów, migreny, choroby lokomocyjnej. Pomimo obiecujących wyników w leczeniu nudności i wymiotów badania nie były kontynuowane [113].

Jednocześnie badania nad poprawą właściwości fizykochemicznych CP-99,994 prowadzono w firmie Pfizer. W wyniku tych prac otrzymano związek CP-122,721 (Rysunek 7), posiadający grupę trifluorometoksyłową w pozycji piątej pierścienia benzyłowego [114]. CP-122,721 wykazywał porównywalne powinowactwo do NK1R, ale około 50-krotnie większą aktywność *in vivo* w porównaniu do CP-99,994 [115]. CP-122,721 był testowany w

leczeniu pooperacyjnych wymiotów i nudności, depresji, chorób zapalnych, w tym astmy i zespołu jelita drażliwego (IBS), ale badania te zostały przerwane [113].



Rysunek 7. Wybrane antagonisty NK1R oparte na strukturze 3-(*N*-benzyloamino)piperydyny.

II.2.2.2.5 Etery benzytowe chinuklidyny, piperydyny oraz morfolin

Firma Merck, Sharpe and Dohme Ltd (obecnie Merck) poszukiwała nowych antagonistów NK1R, wychodząc ze struktur CP-96,345 oraz CP-99,9994. W wyniku prowadzonych badań stwierdzono, że zmiana ugrupowania aminowego na eterowe w CP-96,345 pozwala na zachowanie wysokiego powinowactwa do NK1R [116]. Kolejne modyfikacje dotyczyły podstawienia pierścienia benzyłowego w różnych pozycjach grupami o różnej wielkości i polarności. Otrzymano związek posiadający dwie grupy metylowe w pozycjach *meta* o wysokim powinowactwie do ludzkiego NK1R ($IC_{50} = 1$ nM). Dalsze optymalizacje struktury skutkowały uzyskaniem analogu podstawionego dwoma grupami trifluorometylowymi w pozycjach 3 i 5 pierścienia benzyłowego, związku L-709,210 (Rysunek 8). Wykazano, że tak jak w przypadku CP-96,345 niezbędna jest konfiguracja *S* atomu węgla w pozycji 3 pierścienia, ale w przypadku pozycji 2 zarówno konfiguracja *S* jak i *R* pozwala zachować wysokie powinowactwo do ludzkiego NK1R ($IC_{50} = 1.3$ nM dla diastereoizomeru 2*S*, 3*S* oraz $IC_{50} = 0.7$ nM dla diastereoizomeru 2*R*, 3*S*) [117].

Badania SAR oparte na strukturze CP-99,994 doprowadziły do uzyskania związku L-733,060 (Rysunek 8), wykazującego wysokie powinowactwo do NK1R ($IC_{50} = 0.87$ nM) [118]. L-733,060 w miejscu ugrupowania *N*-(*o*-metoksybenzylo)aminowego CP-99,994 posiada ugrupowanie eteru 3,5-*bis*(trifluorometylo)benzyłowego. Mimo wysokiego powinowactwa do NK1R i doskonałej selektywności L-733,060 wykazywał również powinowactwo do kanałów wapniowych typu L ($IC_{50} = 760$ nM), powodując występowanie działań niepożądanych [9].

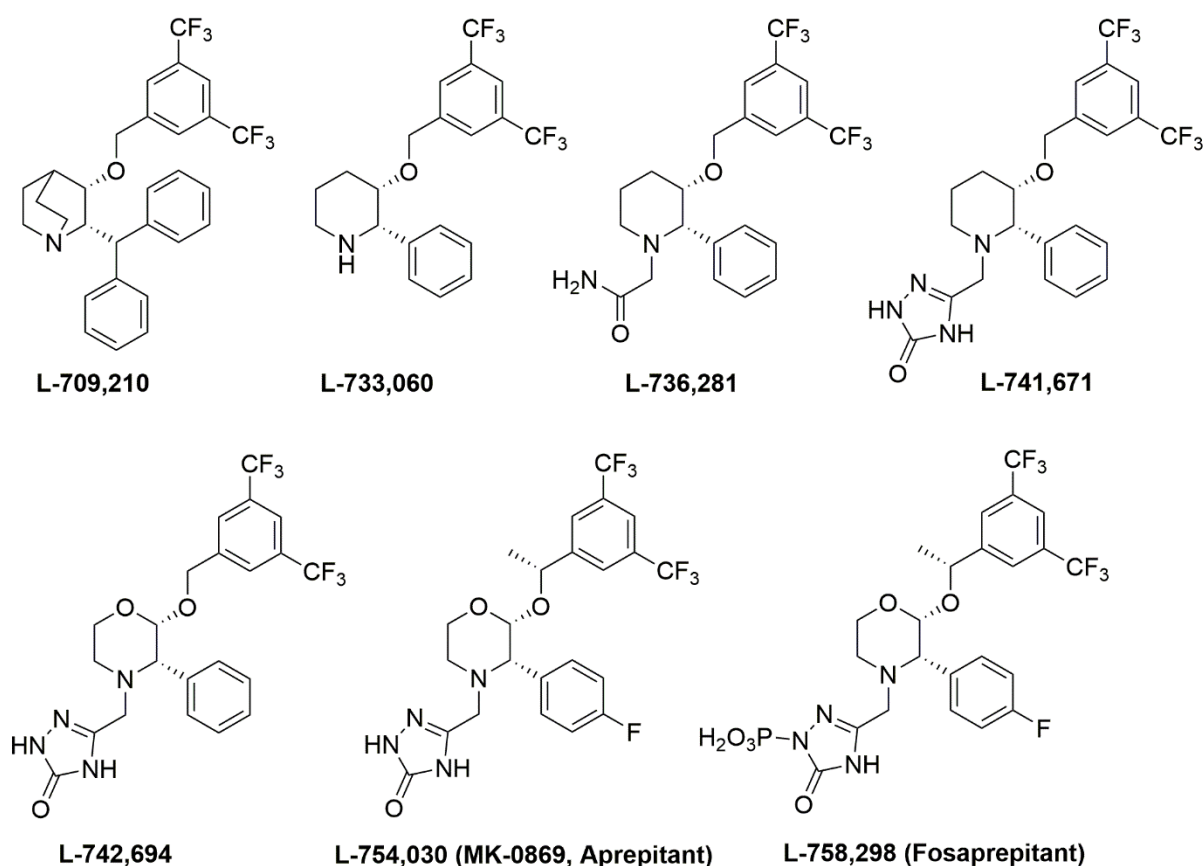
W celu zmniejszenia powinowactwa do kanałów wapniowych podjęto próby podstawienia azotu pierścienia piperydynowego grupami wyciągającymi elektrony [119]. Wśród opracowanych analogów, związek L-736,281 (Rysunek 8), zawierający grupę

acetamidową, zachowywał powinowactwo do NK1R na poziomie nanomolowym ($IC_{50} = 1.3 \text{ nM}$) oraz wykazywał znacznie zmniejszoną aktywność wobec kanałów wapniowych, co przypisano obniżeniu pK_a azotu w pierścieniu piperidynowym.

W toku dalszych badań opracowano serię związków, w których azot pierścienia piperidynowego był podstawiony pięcioczłonowym pierścieniem heterocyklicznym. Prace te doprowadziły do uzyskania związku L-741,671 (Rysunek 8), podstawionego ugrupowaniem 3-okso-1,2,4-triazol-5-ilometylowym [120]. Wprowadzenie tej modyfikacji pozwoliło na uzyskanie silnego, selektywnego antagonisty NK1R ($IC_{50} = 0.05 \text{ nM}$) o lepszej biodostępności po podaniu doustnym.

W innej pracy przedstawiono związek L-742,694 (Rysunek 8), zawierający pierścień morfoliny w miejscu pierścienia piperidyny, którego azot był podstawiony także ugrupowaniem 3-okso-1,2,4-triazol-5-ilometylowym [121]. L-742,694 był aktywny po podaniu doustnym i wykazywał silne powinowactwo do NK1R ($IC_{50} = 0.09 \text{ nM}$). Dalsze modyfikacje, obejmujące wprowadzenie grupy metylowej na węglu alfa łańcucha benzyloвого oraz wprowadzenie podstawnika fluorowego w pozycji *para* do pierścienia fenyloвого pozwoliły na uzyskanie związku L-754,030 (MK-0869, nazwa międzynarodowa: aprepitant, Rysunek 8) [122]. Jest to silny, selektywny antagonist ($IC_{50} = 0.09 \text{ nM}$) o wysokiej aktywności po podaniu doustnym. Zachowanie pierścienia morfoliny pozwoliło na utrzymanie wysokiego powinowactwa, a wprowadzone modyfikacje przyczyniły się do zminimalizowania metabolizmu oksydacyjnego cząsteczki. Związek ten jest odporny na hydrolizę i epimeryzację oraz nie wykazuje żadnej aktywności wobec kanałów wapniowych [123]. Stwierdzono, że wprowadzenie podstawienia pierścienia benzyloвого w pozycjach 3 i 5 w serii związków eterowych, zamiast podstawienia grupą metoksylową w pozycji 2, stosowanego we wcześniejszych strukturach benzyloaminowych (takich jak CP-99,994 i CP-96,345), prowadzi do zwiększenia aktywności biologicznej związku. Ponadto, wydaje się, że grupa 3,5-*bis*(trifluorometylo)fenylowa (TFMP) odgrywa istotną rolę we wzmacnianiu aktywności *in vivo* oraz poprawie metabolizmu cząsteczki [123]. Aprepitant był szeroko testowany w badaniach klinicznych i w 2003 roku substancja ta została zatwierdzona jako lek we wskazaniu do leczenia wymiotów i nudności związanych z chemioterapią przeciwnowotworową oraz występujących w czasie pooperacyjnym [113]. Prowadzono również badania w kierunku zastosowania aprepitantu jako leku przeciwbólowego [124] oraz przeciwdepresyjnego [71,125], ale związek ten nie wykazywał w tych próbach wystarczającej skuteczności. Inne prace przedkliniczne pokazały, że aprepitant ma także działanie przeciwnowotworowe [126,127], przeciwwzapalne [128] i przeciwwirusowe [129].

Aprepitant charakteryzuje się dobrą biodostępnością po podaniu doustnym i penetruje do ośrodkowego układu nerwowego. Jednocześnie cechuje się niską rozpuszczalnością w wodzie, dlatego w toku dalszych prac opracowano rozpuszczalny w wodzie ufosforylowany prolek aprepitantu, fosaprepitant (L-758,298, Rysunek 8), który jest odpowiedni do podawania zarówno dożylnego i doustnego [130,131]. Fosaprepitant został zatwierdzony w 2008 roku do użytku klinicznego w tym samym wskazaniu co aprepitant [132].



Rysunek 8. Wybrane antagonisty NK1R oparte na strukturze eterów benzytowych chinuklidyny, piperydyny oraz morfoliny.

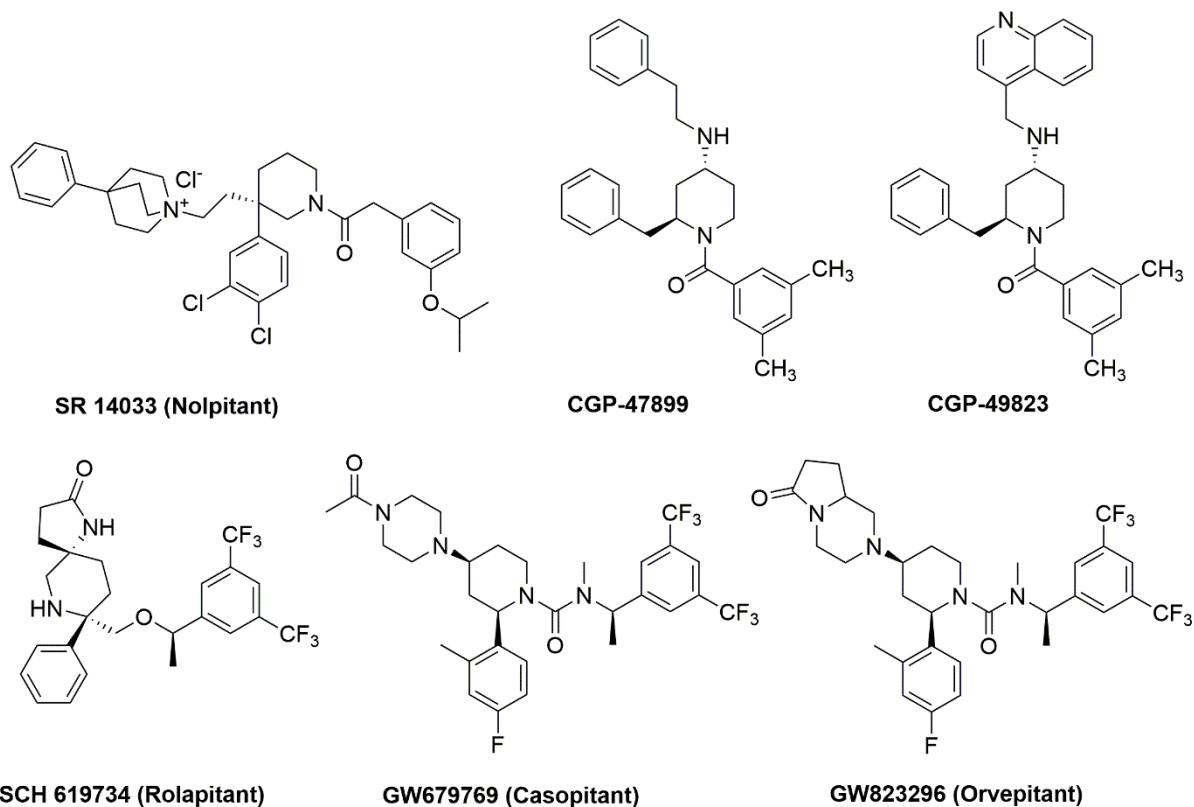
II.2.2.2.6 Antagonisty oparte na strukturze piperydyny

Przeprowadzono szereg badań nad modyfikacjami związków zawierających strukturę piperydyny w celu identyfikacji nowych antagonistów NK1R. W wyniku tych prac naukowcy z firmy Sanofi-Aventis odkryli SR 140333 (nolpitant, Rysunek 9), będący *N*-acylowaną-3-dipodstawioną piperydyną [133]. SR 140333 wykazuje doskonałą selektywność i wysokie powinowactwo zarówno do szczurzych ($IC_{50} = 0.02$ nM) i ludzkich ($IC_{50} = 0.01$ nM) NK1R. Obecność trzech ugrupowań aromatycznych uznano za niezbędne do zachowania wysokiego powinowactwo i selektywności.

W firmie Ciba Geigy w wyniku selektywnego przesiewu firmowej kolekcji substancji chemicznych odkryto związek wiodący, który umożliwił opracowanie serii selektywnych antagonistów należących strukturalnie do 2-benzylo-4-aminopiperydyn [134]. Skryning i optymalizację wykonano przy założeniu, że niezbędna do wysokiego powinowactwa jest obecność co najmniej dwóch grup aromatycznych, oddalonych od siebie od pięciu do dziewięciu atomów. Otrzymano związek o wysokim powinowactwie do NK1R, CGP-47899 (Rysunek 9), w którym ugrupowanie piperydyny jest podstawione trzema pierścieniami aromatycznymi. W wyniku dalszych modyfikacji obejmujących zamianę ugrupowania *N*-2-fenetylowego na *N*-chinolino-4-ylometylove uzyskano związek CGP-49823 (Rysunek 9) [134]. CGP-49823 charakteryzował się dobrym powinowactwem do NK1R ($IC_{50} = 12$ nM) i biodostępnością po podaniu doustnym. Związek ten był testowany w fazie I badań klinicznych do potencjalnego zastosowania w leczeniu zaburzeń lękowych, jednakże badania te nie były kontynuowane [113].

Innym ważnym antagonistą NK1R zawierającym fragment piperydyny jest opracowany w firmie Schering-Plough związek SCH 619734 (rolapitant, Rysunek 9) [135]. Związek ten jest pochodną 1,7-diazaspiro[4,5]dekanon-2-onu podstawioną w pozycji 8 pierścieniem fenylowym oraz łańcuchem z grupą TFMP połączoną wiązaniem eterowym. Rolapitant jest silnym ($IC_{50} = 0.66$ nM) i selektywnym antagonistą NK1R, aktywnym po podaniu doustnym i zdolnym do penetracji do ośrodkowego układu nerwowego. Rolapitant został zatwierdzony do stosowania w praktyce klinicznej jako składnik terapii przeciwwymiotnej w leczeniu wymiotów i nudności związanych z chemioterapią [136].

Innym przykładem związków z tej grupy są opracowane w firmie GlaxoSmithKline GW679769 (Casopitant, Rysunek 9) oraz GW823296 (Orvepitant, Rysunek 9), w których piperydyna podstawiona jest *N*-podstawioną piperazyną, dipodstawionym pierścieniem fenylowym oraz łańcuchem zawierającym grupę TFMP połączoną przez wiązanie amidowe. Casopitant był badany pod kątem leczenia nudności i wymiotów wywołanych chemioterapią, ale badania te zostały przerwane ze względu na konieczność dalszej oceny bezpieczeństwa tego związku [9]. Obecnie prowadzone są badania nad zastosowaniem casopitantu jako potencjalnego leku przeciwdepresyjnego [137]. Testowany klinicznie w leczeniu depresji był także orvepitant, ale okazał się nie być wystarczająco skuteczny [138]. Niedawno opublikowano pozytywne wyniki badań II fazy dla orvepitantu w terapii kaszlu przewlekłego opornego na leczenie (ang. *chronic refractory cough*) [79].



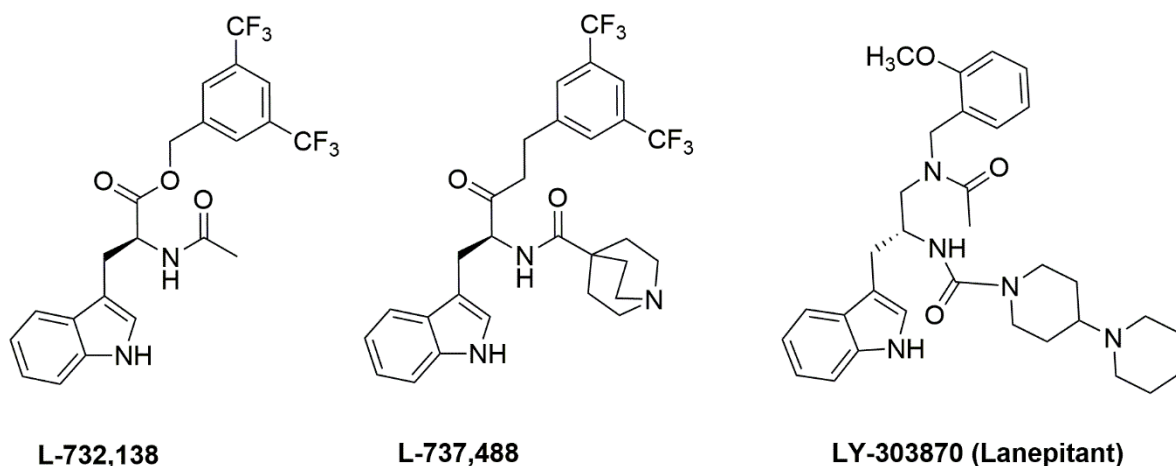
Rysunek 9. Wybrane antagonisty NK1R oparte na strukturze piperydyny.

II.2.2.2.7 Pochodne tryptofanu

Badania przesiewowe kolekcji bibliotek chemicznych przeprowadzone w firmie Merck doprowadziły do zidentyfikowania innej serii antagonistów NK1R zawierających ugrupowanie tryptofanu. Podczas tych prac odkryto, że ester benzyłowy *N*-acetylo-*L*-tryptofanu jest słabym inhibitorem SP i charakteryzuje się mikromolowym powinowactwem do ludzkiego NK1R ($IC_{50} = 3.8 \mu M$) [139]. W wyniku modyfikacji tej cząsteczki zsyntezowano i przebadano wiele analogów, spośród których związek zawierający grupę TFMP, ester 3,5-*bis*(trifluorometylo)benzyłowy *N*-acetylo-*L*-tryptofanu (L-732,138, Rysunek 10) wykazywał silne i selektywne działanie wobec ludzkiego NK1R ($IC_{50} = 3.8 nM$) oraz wysoką aktywność *in vivo* [140,141]. L-732,138 charakteryzował się niższą bioaktywnością po podaniu doustnym, co przypisano obecności wiązania estrowego, podatnego na szybką degradację metaboliczną. W toku dalszych prac przebadano analogi, w których wiązanie estrowe było zastąpione ugrupowaniem eterowym, ketonowym, a także drugorzędowymi i trzeciorzędowymi amidami. Stwierdzono, że związki zawierające grupę ketonową zachowują wysokie powinowactwo. Dalsze optymalizacje mające na celu poprawę biodostępności i rozpuszczalności obejmowały modyfikacje podstawnika acyloaminowego, co doprowadziło do uzyskania związku L-737,488 (Rysunek 10) zawierającego zasadową chinuklidynę oraz

ugrupowanie ketonowe. Związek ten wykazywał wysokie powinowactwo do NK1R ($IC_{50} = 0.9$ nM) [141] oraz wysoką aktywność *in vivo*, dobrą rozpuszczalność w wodzie i słabe powinowactwo do kanałów wapniowych.

Poprzez zastosowanie analogicznej strategii przesiewu kolekcji związków chemicznych, opartej na istotności kluczowych reszt aromatycznych, zespół badawczy z Eli Lilly odkrył szereg amidów oraz estrów *N*-acylowanego tryptofanu [142]. Prace te doprowadziły do otrzymania związku LY-303870 (lanepitant, Rysunek 10). LY-303870 można uznać za *N*-acylowany zredukowany amid D-tryptofanu. LY-303870 zawiera dwa pierścienie piperydyny, które wprowadzono dla zwiększenia powinowactwa receptorowego oraz rozpuszczalności w wodzie. LY-303870 jest selektywnym i silnym antagonistą NK1R ($IC_{50} = 0.15$ nM), o wysokiej aktywności *in vivo*. Związek ten był testowany w II fazie badań klinicznych do zastosowania w leczeniu bólu związanego z chorobą zwyrodnieniową stawów, ale okazał się nie być wystarczająco skuteczny [143].



Rysunek 10. Wybrane antagonisty NK1R oparte na strukturze tryptofanu.

II.2.3. Nowsze koncepcje w chemii medycznej receptora NK1

Intensywne programy poszukiwania antagonistów NK1R prowadzone przez duże firmy farmaceutyczne zaowocowały syntezą setek związków, bardzo dokładnym przebadaniu zależności SAR, i co najważniejsze, wprowadzeniem na rynek aprepitantu i kilku innych antagonistów NK1R. Sukces ten częściowo wyhamował zainteresowanie projektowaniem i syntezą nowych antagonistów receptora ze strony dużego przemysłu farmaceutycznego (sądząc po braku ogłaszanych patentów). Nie ustają jednak próby wykorzystania celowania w NK1R jako interwencji terapeutycznej. W tych nowszych badaniach często wykorzystuje się znane już antagonisty NK1R (w tym zatwierdzone do innego zastosowania), niemniej jednak niektóre specyficzne zastosowania wymagają poszukiwania nowych ligandów NK1R lub

kreatywnego przeprojektowania znanych już związków (np. w przypadku związków hybrydowych lub celowanych biokoniugatów). Próby takie zostaną omówione poniżej.

II.2.3.1. Związki hybrydowe działające przez receptory opioidowe i NK1R

Szerokie zaangażowanie układu SP/NK1R w procesie nocyciepcji skłoniło badaczy do testowania ligandów receptorów NK1R jako potencjalnych związków przeciwbólowych. Jednak pomimo obiecujących wyników uzyskanych w modelach przedklinicznych, w badaniach klinicznych zastosowanie antagonistów NK1R nie przyniosło oczekiwanych efektów [69]. Dlatego też podjęto próby tworzenia związków wielocelowych jednocześnie działających przez NK1R i receptory opioidowe, zaangażowane w modulowaniu bólu. Podejście to opiera się na danych sugerujących wzajemną zależność między działaniem receptorów opioidowych i NK1R [144]. Na ich podstawie wywiedziono potencjał synergistycznego działania ligandów NK1R z agonistami opioidowymi w leczeniu bólu. Przykładowo, uważa się, że podawanie antagonistów NK1R z agonistami opioidowymi może przyczyniać się do lepszej odpowiedzi przeciwbólowej i jednocześnie zmniejszać rozwój tolerancji na działanie analgetyczne oraz minimalizować rozwój działań niepożądanych (np. zaparcia opioidowe). Sugerowano także, że taka dualna farmakoterapia mogłaby być także skuteczna w leczeniu bólu neuropatycznego. W odpowiedzi na te obserwacje opracowano i przebadano wiele związków bifunkcyjnych (inaczej określane w literaturze jako „ligandy bifunkcyjne”, „ligandy biwalentne” lub „hybrydy”), które w jednej cząsteczce łączą farmakofor agonisty opioidowego i farmakofor celowany w receptor NK1, wykazujący właściwości bądź to agonistyczne (uważa się, że SP i jej metabolity w pewnych warunkach mogą wywoływać działanie przeciwbólowe) bądź też właściwości antagonistyczne (hamowanie działania SP, która jest zasadniczo uważana za czynnik pronocycieptywny ma potęgować działanie opioidów i hamować rozwój opioidowych działań niepożądanych wynikających z up-regulacji SP i NK1R).

Związki zbudowane z agonisty opioidowego i antagonisty NK1R po raz pierwszy zostały opisane przez grupę Lipkowskiego [145]. Wśród wczesnych przykładów można wymienić związek AWL-60 (Tyr-Pro-D-Phe-Phe-D-Phe-D-Trp-Met-NH₂) składający się z fragmentu kazomorfiny (opioidowy peptyd pochodzący z białka mleka) i fragmentu peptydowego antagonisty NK1R. W wyniku tego połączenia uzyskano hybrydę o silnym i długotrwałym działaniu analgetycznym *in vivo*, pomimo umiarkowanego powinowactwa do receptora μ -opioidowego (MOR) (IC₅₀ = 210 nM) i δ -opioidowego (DOR) (IC₅₀ = 270 nM) [146].

Dalsze badania doprowadziły do uzyskania związku AA501 (Tyr-D-Ala-Gly-Phe-NH-NH<-Trp(Z)), powstałego w wyniku połączenia fragmentu silnego agonisty opioidowego, [D-Ala²]-enkefaliny (Tyr-D-Ala-Gly-Phe-) z fragmentem antagonisty NK1R, *N*-benzyloksykarbonylotryptofanem (Z-Trp) za pomocą mostka hydrazydowego (-NH-NH-) [147]. Hybryda charakteryzowała się wysokim powinowactwem zarówno do MOR i NK1R oraz zależnym od dawki działaniem przeciwbólowym *in vivo* w modelach bólu ostrego, zapalnego oraz neuropatycznego. Podobne korzystne działanie przeciwbólne oraz dodatkowe obniżenie rozwoju tolerancji analgetycznej i zaparcia opioidowego uzyskano dla analogu, w którym farmakoforem celowanym w NK1R był izomeryczny Z-D-Trp (związek AA3266) [148,149]. Inne hybrydy skonstruowane z różnorodnych peptydowych fragmentów opioidowych i fragmentów organicznych zawierających grupę TFMP (celowane w NK1R) przedstawiła Dyniewicz i wsp. [150,151].

Wtorek i wsp. przedstawił serię peptydowych hybryd zbudowanych z cyklicznego farmakoforu opioidowego (Tyr-c[D-Lys-Phe-Phe-Asp]) oraz fragmentu substancji P (agonista NK1R) lub spantylu II (antagonista NK1R) [152]. Co ciekawe, zarówno hybrydy typu agonista opioidowy/antagonista NK1R jak i agonista opioidowy/agonista NK1R, wykazały *in vivo* działanie przeciwbólne, nie powodując rozwoju tolerancji analgetycznej i zaparcie.

Warto wzmiankować także związek AA3052 (Tyr-D-Arg-Phe-Lys-D-Phe-Phe-D-Phe-Leu-Leu-NH₂) [153], który został zaprojektowany jako połączenie fragmentu peptydowego agonisty MOR (DALDA) i modyfikowanego fragmentu SP(7-11). Związek ten, pomimo słabego powinowactwa do MOR ($K_i = 729.8$ nM) wykazywał działanie przeciwbólne w badaniach *in vivo* oraz nie wywoływał rozwoju tolerancji analgetycznej.

Prace nad ligandami bifunkcyjnymi agonista opioidowy – antagonistą NK1R prowadzone były także w zespole Hruby'ego. Yamamoto i wsp. przedstawili serię hybryd, w których jako sekwencję sygnałową do receptorów opioidowych wybrano fragment [D-Ala²]-enkefaliny, a jako sekwencję sygnałową receptora NK1 użyto fragmenty Phe-Pro-Leu-Trp-O-[3',5'-Bzl(CF₃)₂] lub D-Phe-Pro-Leu-Trp-O-[3',5'-Bzl(CF₃)₂], zawierające kluczową dla wysokiego powinowactwa do NK1R grupę TFMP [154]. Związek TY005 (Tyr-D-Ala-Gly-Phe-Met-Pro-Leu-Trp-O-[3',5'-Bzl(CF₃)₂]) wykazywał aktywność przeciwbólową w badaniach *in vivo* w modelach bólu ostrego i neuropatycznego, hamując lub osłabiając rozwój tolerancji wywołanej opioidami [155]. Dalsze modyfikacje obejmowały zmianę podatnego na hydrolizę enzymatyczną wiązania estrowego na wiązanie amidowe, w wyniku czego otrzymano związek TY027 (Tyr-D-Ala-Gly-Phe-Met-Pro-Leu-Trp-NH-[3',5'-

Bzl(CF₃)₂]). Hybryda ta wykazywała większą stabilność enzymatyczną, 4-krotnie większe powinowactwo do DOR i 2-krotnie większe do MOR, zachowując właściwości przeciwbólowe. Usunięcie dwóch grup trifluorometylowych w pierścieniu aromatycznym (związek TY025, Tyr-D-Ala-Gly-Phe-Met-Pro-Leu-Trp-NH-Bzl) spowodowało także zwiększenie powinowactwa do MOR i DOR, ale jednocześnie znaczny spadek do NK1R.

W innym podejściu, Vardanyan i wsp. przedstawili pracę, w której jako farmakofor celowany w MOR zastosowano fentanyl (małocząsteczkowy agonista MOR, zatwierdzony lek), zaś farmakofor NK1R stanowił antagonistę NK1R, L-732,138 [156]. Przetestowano hybrydy, w których farmakofory połączone były wiązaniem jonowym (sole) lub kowalencyjnym. Powinowactwo do NK1R w obydwu grupach było dobre ($K_i = 6.8\text{--}44\text{ nM}$). Związki z farmakoforami połączonymi kowalencyjnie charakteryzowały się umiarkowanym powinowactwem do MOR ($K_i = 120\text{--}400\text{ nM}$). Natomiast hybrydy-pary jonowe wiązały się do MOR na znacznie niższym poziomie ($K_i > 1\text{ }\mu\text{M}$).

Badania nad ligandami bifunkcyjnymi opioid/antagonista NK1R prowadzone były także w grupie Balleta. Zespół ten przedstawił m.in. bifunkcyjny peptydomimetyk o sekwencji Dmt-D-Arg-Aba-Gly-NMe-[3',5'-Bzl(CF₃)₂] (Dmt – 2',6'-dimetylotyrozyna) [157]. Jako farmakofor opioidowy zastosowano fragment oparty na sekwencji peptydu Dmt¹-DALDA (Dmt-D-Arg-Phe-Lys-NH₂), zaś farmakofor NK1R stanowił nowo odkryty antagonistę NK1R o strukturze Ac-Aba-Gly-NMe-[3',5'-Bzl(CF₃)₂] (Ac – acetylo; Aba – 4-amino-1,2,4,5-tetrahydro-2-benzazepin-3-on). Zaobserwowano, że ugrupowanie Aba jest korzystne dla wiązania związku do receptorów MOR i DOR. Otrzymana hybryda charakteryzowała się wysokim powinowactwem do MOR ($K_i = 0.42\text{ nM}$), DOR ($K_i = 10.4\text{ nM}$) i NK1R ($K_i = 0.5\text{ nM}$). Połączenie farmakoforów w związek bifunkcyjny spowodowało wzrost powinowactwa do NK1R względem komponenty NK1R (Ac-Aba-Gly-NMe-3',5'-(CF₃)₂-Bn) i jednocześnie obniżenie powinowactwa do MOR względem komponenty opioidowej (Dmt-D-Arg-Aba-Gly). Hybryda wykazywała silne działanie analgetyczne oraz zdolność do przenikania BBB, jednak siła działania farmakofora opioidowego podanego samodzielnie była wyższa niż po podaniu hybrydy. Co więcej stosowanie tej hybrydy nie pozwoliło na uniknięcie efektów ubocznych wywołanych stosowaniem opioidów [158].

Jak widać poszczególne próby tworzenia hybryd opioidowo-(anty)tachykininowych przynosiły zróżnicowane rezultaty, jeśli chodzi o spełnienie założenia o poprawie działania przeciwbólowego przy jednoczesnym zredukowaniu działań niepożądanych. Jak do tej pory, żadna z eksperymentalnych hybryd nie została przeprowadzona do badań klinicznych, co pozwoliłoby na całościową ocenę tej strategii projektowania innowacyjnych analgetyków.

Dodatkowo, omówione przykłady pokazują, iż połączenie dwóch komponent farmakoforowych nie zawsze gwarantuje zachowanie wysokiego powinowactwa receptorowego do zamierzonych celów molekularnych. Warto przy okazji nadmienić, że niedawno Gadais i wsp. przedstawili peptydomimetyczne hybrydy celowane w receptory opioidowe MOR/DOR oraz receptory NK2R i NK3R jako potencjalny nowy kierunek poszukiwania wielocelowych analgetyków [159].

II.2.3.2. Antagonizm endosomalnego sygnałowania przez NK1R

Inna nowa koncepcja leczenia bólu przez celowanie w receptor NK1 wiąże się z antagonizmem endosomalnego sygnałowania przez ten receptor.

Powszechnie receptory GPCR uznaje się za receptory sygnalizujące z powierzchni komórek poprzez oddziaływanie z białkami G. Jednak, według nowszych ustaleń, wiele GPCR, m.in. NK1R, może także sygnalizować z przedziałów wewnątrzkomórkowych, w tym z endosomów [160]. Sygnały wewnątrzkomórkowe prowadzą do reakcji fizjologicznych, różniących się od tych, które pochodzą z sygnalizacji w powierzchniowej błonie komórkowej.

Receptor NK1 po związaniu się z SP (lub innym agonistą), jest fosforylowany przez kinazy GRK, co sprzyja jego oddziaływaniu z β -arestynami, które rozprzegają NK1R od białka G i odczulają (desensytyzują) sygnalizację prowadzoną na poziomie błony komórkowej. β -arestyny umożliwiają również połączenie NK1R z klatryną i białkiem adaptorowym-2 (AP2), które pośredniczą w endocytozie receptora. Dawniej, endocytoza uznawana była za mechanizm kończący sygnalizację NK1R. Obecnie uważa się jednak, że NK1R może dalej sygnalizować po endocytozie, tworząc w obrębie endosomów kompleksy sygnałowe, które generują wtórne przekaźniki i aktywują kinazy w przedziałach subkomórkowych, przy czym sygnalizacja z endosomów jest trwalsza niż sygnalizacja z poziomu powierzchni komórki [160].

Uważa się, że sygnalizacja endosomalna NK1R jest zaangażowana w nocycepcję, co wykazano w modelach bólu ostrego i neuropatycznego [43]. Bódźce bólowe prowadzą do uwolnienia SP z pierwotnych neuronów czuciowych. Następnie SP w rogu grzbietowym rdzenia kręgowego aktywuje NK1R w neuronach drugiego rzędu, co prowadzi do jego endocytozy. Wówczas sygnalizacja endosomalna NK1R pośredniczy w przedłużonej aktywacji neuronów rdzeniowych i przekazywaniu sygnałów bólowych [160]. Wobec tego, zaproponowano, że hamowanie tej endosomalnej sygnalizacji NK1R, może przynosić działanie przeciwbólowe i być użyteczną strategią terapeutyczną.

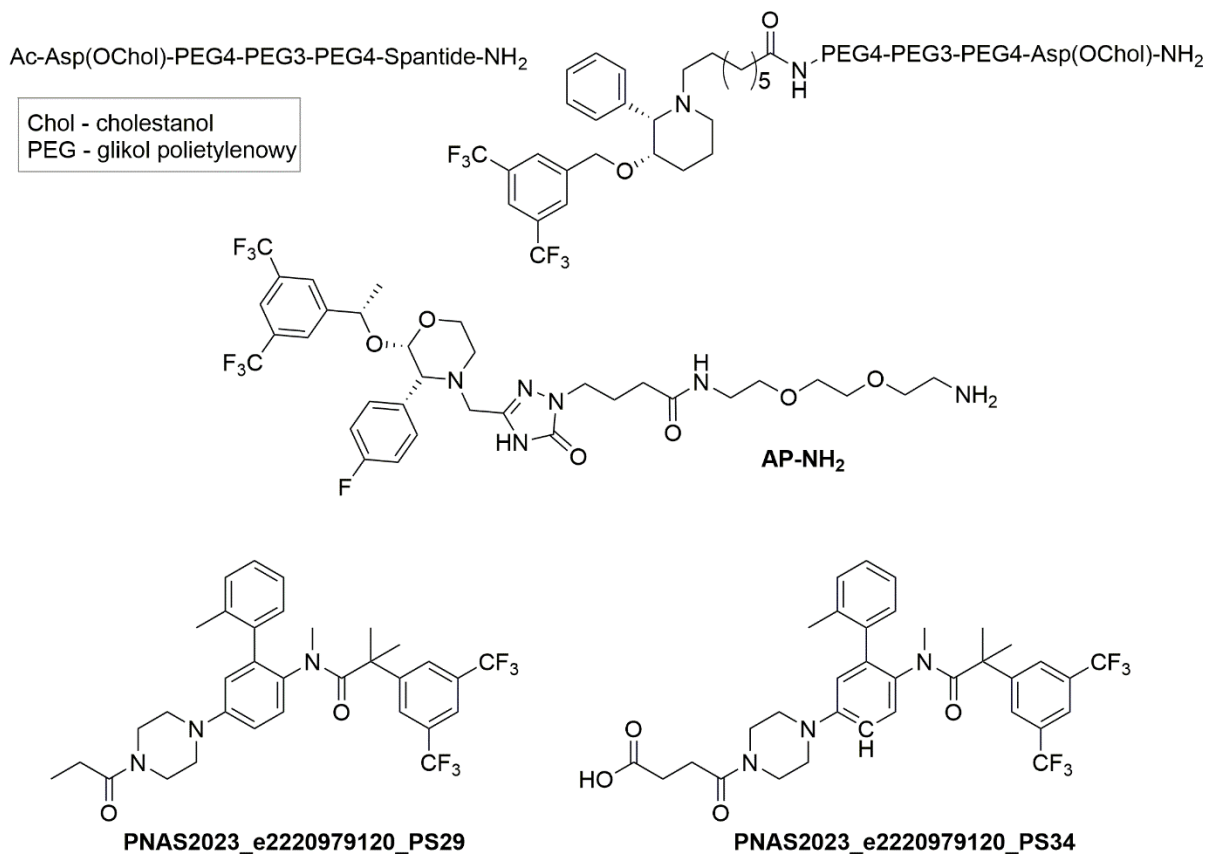
Jensen i wsp. przedstawili antagonistę NK1R Span-Chol (Rysunek 11), zbudowanego z sekwencji spantydu I (Span; nieprzechodzący przez błony komórkowe peptydowy antagonistą NK1R), łącznika PEG oraz cholestanolu (Chol), który miał umożliwić zakotwiczenie koniugatu w błonie komórkowej, dostarczenie i zatrzymanie antagonistów w endosomach zawierających NK1R [43]. Rozbudowa spantydu na N-końcu nie wpłynęła negatywnie na aktywność antagonistyczną (IC_{50} w zakresie pojedynczych nanomoli). Span-Chol gromadził się w endosomach, powodując trwałe (ok. 6 godzin) i selektywny antagonizm endosomalny i przedłużone działanie przeciwbólne hamując mechaniczną przeczulicę bólową, podczas gdy działanie antynocyceptywne dla Span było nieznaczne i krótkie. W dalszych badaniach wykazano, że zakotwiczone w lipidach antagonisty NK1R (Span-Chol) trwale hamują sygnalizację z endosomów, ale także wpływają na sygnalizację NK1R zlokalizowanego w błonie komórkowej, hamując rekrutację NK1R- β -arestyny na powierzchni komórki i zmniejszając endocytozę receptora [161]. W przedklinicznym modelu bólu Span-Chol wywoływał przedłużoną antynocycepcję powyżej 9h. Analogiczny koniugat L-733,060-Chol (Rysunek 11), zbudowany z małowcząsteczkowego antagonisty, linkera i cholestanolu, wykazywał podobne przedłużone działanie przeciwbólne w porównaniu do nieskoniugowanego związku wyjściowego [43].

Ramírez-García i wsp. opisali wrażliwe na pH nanocząstki polimerowe zawierające aprepitant, które mogą precyzyjnie hamować sygnalizację endosomalną, odpowiedzialną za przewlekły ból [162]. Nanocząstki zaprojektowano tak, aby nie oddziaływały z receptorem na powierzchni komórek, a bezpośrednio trafiały do endosomów. Gdy nanocząstki trafiają do endosomów z NK1R, na skutek niższego pH panującego w tym otoczeniu, uwalniają aprepitant w sposób długotrwały, co prowadzi do zahamowania aktywacji neuronów rdzenia kręgowego indukowanej przez SP i w efekcie zapobiega przekazywaniu sygnałów bólowych. W modelach przedklicznych wykazano, że użycie tych nanocząstek wzmocniło antynocyceptywne działania aprepitantu w modelach bólu zapalnego i neuropatycznego. Wykorzystanie nanocząstek wrażliwych na pH, pozwoliło na zastosowanie minimalnej dawki aprepitantu wykazującej aktywność przeciwbólową, co może przyczynić się do uniknięcia działań niepożądanych.

Ponieważ nanocząstki wrażliwe na pH są względnie niestabilne i podatne na efekty rozcieńczenia, stosunkowo szybko rozkładają się w endosomach i uwalniają antagonistę, co może ograniczyć czas działania antynocyceptywnego, Latorre i wsp. podjęli próbę opracowania stabilnych nanogwiazd polimerowych, zawierających zmodyfikowaną cząsteczkę aprepitantu [163]. Struktura aprepitantu została rozbudowana na pierścieniu

triazolinonowym o łącznik butyrylo-PEG₂ zakończony grupą aminową (AP-NH₂, Rysunek 11), pozwalającą na połączenie ze strukturą polimeru. Nanogwiazdy charakteryzują się odpornością na rozcieńczanie i możliwością zaprojektowania do kontrolowanego i długotrwałego uwalniania ładunku, co pozwala na przedłużone działanie przeciwbólne. Skuteczność działania nanogwiazd zawierających AP-NH₂ oceniano w przedklinicznych modelach bólu neuropatycznego i zapalnego. W endosomach nanogwiazdy uwalniały antagonistę przez 24 godziny, zakłócały sygnalizację endosomalną NK1R oraz zapewniały trwałą i skuteczną analgezję (ponad 10 godzin). Wykorzystanie nanogwiazd umożliwiła jednoczesną inkorporację wielu antagonistów i pozwala na celowanie ich do neuronów odpowiedzialnych za przekazywanie sygnałów bólowych, co może być skutkować długotrwałym działaniem przeciwbólowym.

Inne podejście zaproponowali Hegron i wsp. opisując analogi netupitantu, które trwale hamowały sygnalizację endosomalną [164]. W pracy porównano działanie przeciwbólne aprepitantu, netupitantu oraz specjalnie zaprojektowanych analogów netupitantu o różnej lipofilowości i kwasowości. Opracowano serię pochodnych, w których ugrupowanie pirydynowe w netupitancie zastąpiono pierścieniem benzenowym, a w miejsce azotu piperazyny podstawiano różne grupy (obojętne, kwaśne, słabo zasadowe, zasadowe i silnie zasadowe). Lipofilowe i kwasowe analogi (PNAS2023_e2220979120_PS29 oraz PNAS2023_e2220979120_PS34, Rysunek 11) najskuteczniej antagonizowały endosomalną sygnalizację NK1R, wykazywały trwalsze i silniejsze działanie przeciwbólne *in vivo* niż aprepitant i netupitant.



Rysunek 11. Wybrane cząsteczki oparte o antagonistów NK1R w endosomalnym sygnalowaniu.

Doniesienia o hamowaniu endosomalnego sygnalowania NK1R pojawiły się w literaturze w kilku ostatnich latach. Zaproponowany mechanizm działania przeciwbólowego wydaje się obiecujący i potwierdzony danymi przedklinicznymi. Jak uczą jednak doświadczenia z wcześniejszych badań nad „zwykłymi” antagonistami NK1R w leczeniu bólu, nie jest pewne, czy ewentualne badania kliniczne potwierdzą przydatność tego mechanizmu w leczeniu bólu u ludzi.

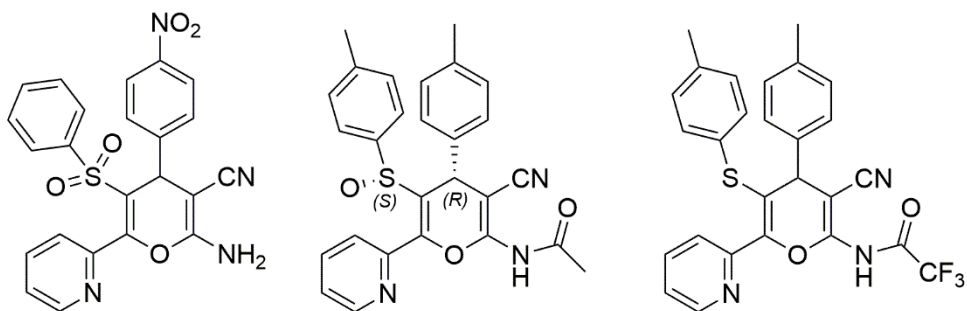
II.2.3.3. Antagonisty NK1R o działaniu przeciwnowotworowym

Liczne badania wykazały, że układ SP/NK1R odgrywa też znaczącą rolę w procesach nowotworowych. Po związaniu z NK1R SP indukuje proliferację i migrację komórek nowotworowych, hamuje procesy apoptotyczne, a także nasila angiogenezę. Skutkuje to rozrostem tkanki nowotworowej i powstawaniem przerzutów [50,66,67]. Ponadto stwierdzono, że receptor NK1 jest nadeksprimowany w licznych liniach komórkowych nowotworów, w tym białaczce, glejaku, neuroblastomie, czerniaku, raku piersi, jajnika, prostaty, płuc, trzustki i tarczycy [19,31].

Obserwacje te doprowadziły do zainteresowania badaczy zastosowaniem znanych już antagonistów NK1R jako potencjalnych substancji o działaniu przeciwnowotworowym. Wykazano, że antagoniści NK1R takie jak L-732,138, L-733,060, aprepitant czy fosaprepitant charakteryzują się szerokim działaniem przeciwnowotworowym [31,165]. Selektywnie hamują one wzrost komórek nowotworowych, angiogenezę oraz migrację komórek nowotworowych, wykazują działanie proapoptyczne na komórki nowotworowe, zmniejszają objętość guza oraz stan zapalny [31,50,66,67,126,165–167]. Aktywność przeciwnowotworową obserwowano *in vitro* na wielu liniach komórkowych z nadekspresją NK1R, między innymi czerniaka, neuroblastomy, glejaka, siatkówczaka, raka trzustki, krtani czy przewodu pokarmowego. Działanie to wykazano także *in vivo* na mysich modelach opartych na heteroprzeszczepie ludzkich nowotworów [126,127,166,168–171].

Pomimo tych pozytywnych danych przedklinicznych, oraz znanego profilu bezpieczeństwa w przypadku zatwierdzonych antagonistów NK1R, kliniczne dowody na skuteczność antagonistów NK1R w leczeniu nowotworów są skromne i ograniczają się do kilku opisów przypadku [172,173]. Dopiero niedawno, w 2021 roku rozpoczęto badanie kliniczne II fazy mające na celu ocenę skuteczności i bezpieczeństwa stosowania dużych dawek oraz długotrwałego leczenia aprepitantem u pacjentów z zaawansowanym rakiem płuc [174].

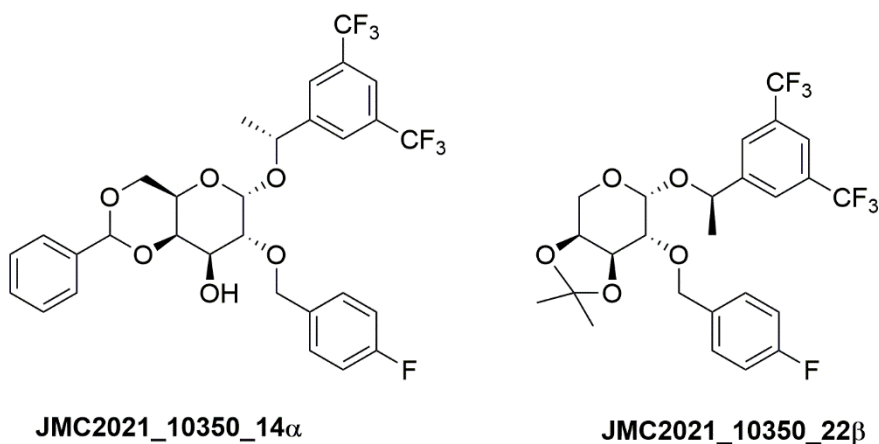
W ciągu ostatnich lat podjęto także próby opracowania nowych antagonistów NK1R o zamierzonym działaniu przeciwnowotworowym. Recio i wsp. przedstawili serię 5-arylotio-2-amino-4*H*-piranów jako nowych ligandów NK1R (Rysunek 12) [175]. Najbardziej aktywnym antagonistą w tej grupie jest związek EJMC2017_644_*rac*-17 (Rysunek 12). Określono wpływ działania wybranych pochodnych na liniach komórkowych ludzkiego raka płuc (A549) oraz prawidłowych płuc (MRC-5). Najwyższą aktywność przeciwnowotworową uzyskano dla pochodnych amidowych EJMC2017_644_18(*S_s*,*R_C*) i EJMC2017_644_*rac*-19, przy czym związek EJMC2017_644_*rac*-19 wykazywał większą selektywność względem komórek nowotworowych. Co ciekawe, wśród badanych analogów znajdowały się także związki o działaniu agonistycznym na NK1R. Jest to o tyle warte odnotowania, że znanych jest niewiele małowcząsteczkowych agonistów tego receptora [176].



EJMC2017_644_rac-17 **EJMC2017_644_18(S_S,R_C)** **EJMC2017_644_rac-19**

Rysunek 12. Wybrane przykłady z serii pochodnych 5-arylotio-2-amino-4H-piranołów.

W innej pracy Recio i wsp. opisali serię nowych antagonistów NK1R opartych na strukturze węglowodanów (D-galaktozy lub L-arabinozy) [177]. Wszystkie otrzymane związki wykazywały wysokie powinowactwo i aktywność antagonistyczną wobec NK1R. Dla poszczególnych pochodnych oceniono cytotoksyczność wobec ludzkich komórek raka płuc (A549) oraz komórek prawidłowych płuc (MRC-5). Wszystkie przebadane związki wykazywały pewną selektywność względem komórek nowotworowych. Najbardziej selektywnymi związkami były JMC2021_10350_14 α (Rysunek 13) i JMC2021_10350_22 β (Rysunek 13), przy czym pochodne D-galaktozy (np. JMC2021_10350_14 α) były bardziej cytotoksyczne niż pochodne L-arabinozy (np. JMC2021_10350_22 β). Związek JMC2021_10350_14 α został dodatkowo przebadany na dwunastu różnych typach linii komórkowych ludzkich nowotworów, wykazując znaczącą aktywność cytotoksyczną oraz selektywność. Kontynuując te badania, zespół zaprojektował i syntezował również pochodne oparte o szkielet D-glukozy, wśród których znalazły się związki o znacznej aktywności cytotoksycznej i selektywności [178].



JMC2021_10350_14 α

JMC2021_10350_22 β

Rysunek 13. Nowe antagonisty NK1R oparte na strukturze węglowodanów.

II.2.3.4. Receptor NK1R w radiofarmacji

Nadekspresja receptora NK1 na licznych typach komórek nowotworowych, przy jednocześnie względnie niskiej ekspresji w tkankach prawidłowych, stwarza możliwość zastosowania ligandów NK1R jako wektorów radionuklidów w celowanej terapii radionuklidowej (TRT, ang. *targeted radionuclide therapy*) i diagnostyce radionuklidowej. Idea TRT (lub analogicznego podejścia diagnostycznego) opiera się na selektywnym dostarczaniu do komórek nowotworowych radiofarmaceutyku/radiokoniugatu, który zawiera radionuklid o odpowiednich właściwościach rozpadu promieniotwórczego. Selektywność uzyskuje się dzięki selektywnemu wektorowi, który wiąże się z pewnym celem molekularnym nadekspresywowanym na powierzchni komórek nowotworowych. Zazwyczaj do celów diagnostycznych stosuje się radionuklidy o stosunkowo krótkim czasie połowicznego rozpadu, emitujące promieniowanie γ lub β^+ , a do celów terapeutycznych radionuklidy o dłuższym czasie połowicznego rozpadu, emitujące promieniowanie β^- , α lub elektrony Augera [167]. Radiofarmaceutyki składają się z wektora (cząsteczka o wysokim i selektywnym powinowactwie do określonego celu molekularnego), łącznika (umożliwia przyłączenie radionuklidu lub ugrupowania chelatującego z radionuklidem do struktury wektora bez zmiany jego właściwości), chelatora (wielokleszczowy ligand kompleksujący radionuklid) oraz radionuklidu.

Istnieje wiele badań *in vivo* i *in vitro* opisujących zastosowanie w TRT i diagnostyce radionuklidowej ligandów NK1R, głównie SP i jej analogów [167]. Niektóre z tych preparatów zostały poddane analizie klinicznej, wykazując obiecujące wyniki terapeutyczne. Wczesne eksperymenty kliniczne obejmowały zastosowanie [$^{111}\text{In-DTPA-Arg}^1$]-SP (DTPA – kwas dietyleno-triamino-pentaoctowy) do scyntygrafii miejsc zapalnych w różnorodnych chorobach immunologicznych oraz obrazowania zmian w grasicy [179]. Mimo, że radiokoniugat ten po podaniu dożylnym ulegał szybkiej degradacji oraz charakteryzował się krótkim czasem połowicznego rozpadu, może być on użyteczny w obrazowaniu. Inne próby kliniczne dotyczyły zastosowania radiofarmaceutyków opartych na strukturze SP w terapii glejaka [180]. Radiokoniugaty te składały się z SP połączonej makrocyklicznym ligandem chelatującym DOTAGA (kwas 1,4,7,10-tetraazacyklododekan-4,7,10-trioctowo-1-[2-glutarowy]), wyznakowane ^{90}Y , ^{177}Lu lub ^{213}Bi . We wstępnych eksperymentach klinicznych wykazano, że zastosowanie tych radiofarmaceutyków (podanie dokanałowe) powodowało martwicę komórek nowotworowych i umożliwiało skuteczniejszą resekcję chirurgiczną guza, ale ze względu na niską stabilność wektorów w surowicy stosowanie tych związków jest

ograniczone m.in. w podaniu dożylnym. W toku dalszych prac otrzymano radiofarmaceutyki oparte na pochodnej SP, [Thi⁸, Met(O₂)¹¹]SP (Thi – tienyloalanina), o lepszej stabilności. Wektor ten został wykorzystany w radiofarmaceutykach, poddanych szeregu badań *in vitro*, *in vivo* oraz testom klinicznym do diagnostyki i leczenia glejaka. Aktualnie w badaniach klinicznych testowane są [²¹³Bi]Bi-DOTA-[Thi⁸, Met(O₂)¹¹]SP (DOTA – kwas 1,4,7,10-tetraazacyklo-dodekan-1,4,7,10-tetraoctowy) lub [²²⁵Ac]Ac-DOTA-[Thi⁸, Met(O₂)¹¹]SP w połączeniu z [⁶⁸Ga]Ga-DOTA-[Thi⁸, Met(O₂)¹¹]SP do obrazowania i terapii glejaka wielopostaciowego, które we wstępnych badaniach okazały się być dobrze tolerowane, bezpieczne dla pacjenta, i skuteczniejsze w porównaniu do standardowych metod leczenia [181–184].

Jednak ze względu na stosunkowo niską stabilność metaboliczną oraz niską lipofilowość (a zatem słabą dyfuzję) SP i jej analogów, stosowanie peptydowych wektorów w radiofarmaceutykach jest ograniczone. Rozwiązaniem tego problemu może być zastosowanie małowiązanych antagonistów NK1R, które są stabilne i nie ulegają degradacji enzymatycznej *in vivo* oraz charakteryzują się wyższą lipofilowością. Zhang i wsp. opracowali radiofarmaceutyk NK1R-NOTA (Rysunek 14), wykorzystujący jako wektor znanego antagonistę L-733,060, który za pomocą łącznika PEG został połączony z cyklicznym ligandem chelatującym NOTA (kwas 1,4,7-triazacyklononan-1,4,7-trioctowy) i wyznakowany radionuklidami ⁶⁴Cu lub ⁶⁷Ga (do badań *in vitro*) [185]. NK1R-NOTA charakteryzował się wysokim powinowactwem do NK1R oraz pełną stabilnością w krwi mysiej po 30 minutach. W badaniu *in vivo* (podanie dożylne) za pomocą pozytonowej tomografii emisyjnej (PET, ang. *positron emission tomography*) wykazano, że [⁶⁴Cu]NK1R-NOTA może być użyteczny w obrazowaniu zmian nowotworowych wykazujących ekspresję NK1R, aczkolwiek dogodna obserwacja możliwa było dopiero po 20 godzinach od wstrzyknięcia. W innej pracy Kanduluru i wsp. opisali kolejny radiofarmaceutyk bazujący na antagoniście L-733,060 (MIB2020_377_3, Rysunek 14), który za pomocą łącznika PEG i fragmentu Arg-Asp-Lys został połączony z chelatującym ugrupowaniem peptydowym i wyznakowany radionuklidem ^{99m}Tc do potencjalnego obrazowania i oceny stopnia zaawansowania wielu guzów neuroendokrynych [186]. Fragment Arg-Asp-Lys wprowadzono w celu zmniejszenia niespecyficznego wiązania do błon. Przedstawiony radiofarmaceutyk wykazywał wysokie powinowactwo i selektywność do NK1R. W obrazowaniu za pomocą tomografii emisyjnej pojedynczych fotonów/ tomografii komputerowej (SPECT/CT, ang. *single-photon emission computed tomography/computed*

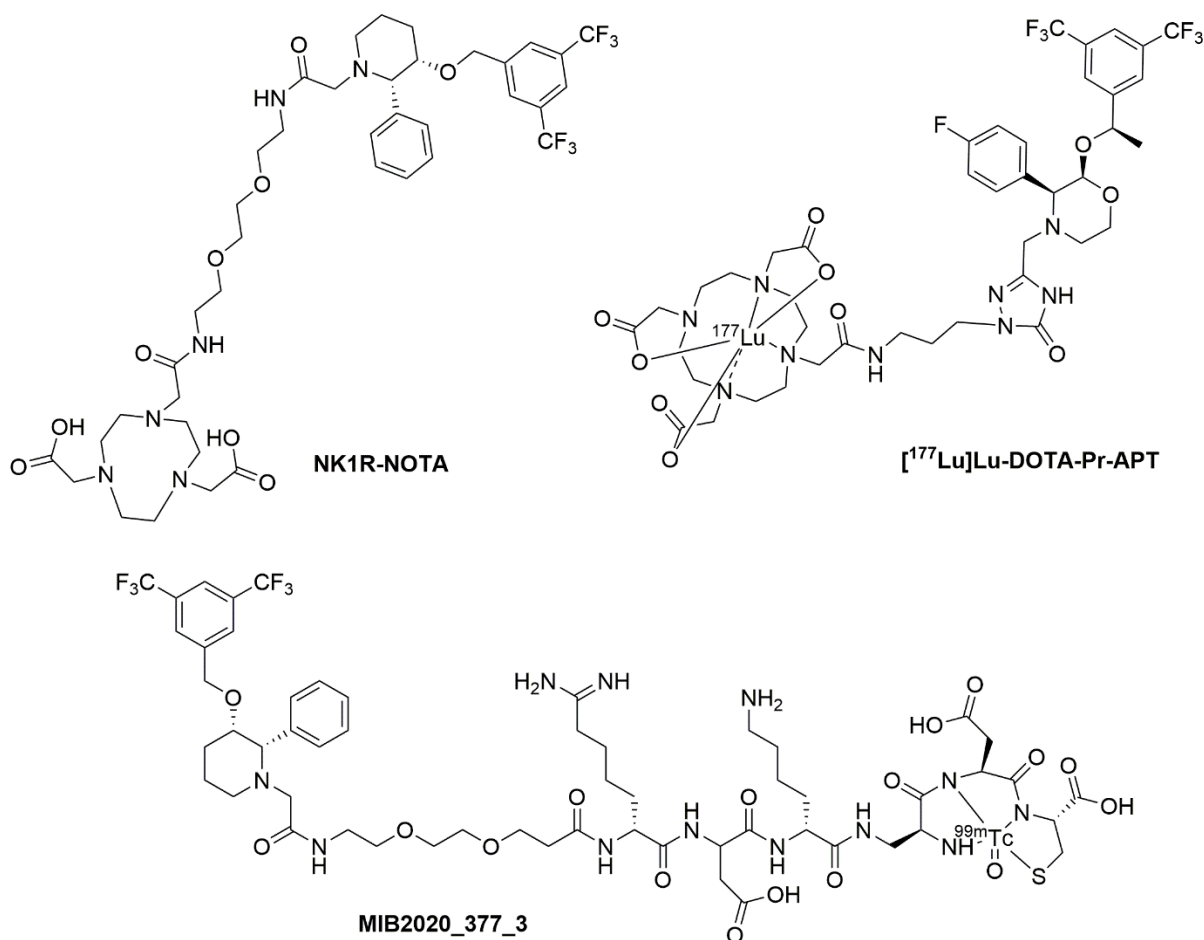
tomography) radiokoniugat po podaniu dożylnym był w wysokim stopniu oraz selektywnie wychwytywany do guza eksprymującego NK1R już 2 godziny po wstrzyknięciu.

Halik i wsp. opracowali i przebadali *in vitro* nowe radiokoniugaty bazujące na strukturze aprepitantu i jego analogach [187,188]. Otrzymano serię pochodnych aprepitantu poprzez dołączenie do pierścienia triazolinonu łączników alifatycznych o różnej długości oraz łączników acetamidowych zakończonych pierwszorzędową grupą aminową. Po wstępnych badaniach stabilności wszystkie pochodne połączone z chelatorem DOTA wyznakowano ^{68}Ga i ^{177}Lu . Otrzymane radiokoniugaty poddano ocenie stabilności, lipofilowości oraz wiązalności do ludzkiego NK1R. Większość przedstawionych radiokoniugatów charakteryzowała się pełną stabilnością w surowicy, wysoką lipofilowością oraz wysokim powinowactwem do NK1R. Najbardziej interesującym radiokoniugatem w tej grupie wydaje się być związek [^{177}Lu]Lu-DOTA-Pr-APT (Rysunek 14), w którym cząsteczka aprepitantu połączona jest z kompleksem radionuklidu za pomocą łącznika alifatycznego zawierającego trzy grupy CH_2 , niezbędne jednak jest przeprowadzenie dalszych badań *in vivo*.

Stosowanie radiofarmaceutyków ukierunkowanych na receptor NK1R stanowi obiecującą strategię w dziedzinie medycyny nuklearnej. Wykorzystanie takich radiokoniugatów pozwala na precyzyjne identyfikowanie i lokalizowanie obszarów nowotworowych nadeksprymujących NK1R, umożliwiając dokładniejszą diagnostykę, monitorowanie postępów terapii oraz leczenie przy zastosowaniu celowanej terapii radionuklidowej. Konieczny jest jednak dalszy rozwój i poszukiwanie nowych ligandów NK1R o lepszych właściwościach fizykochemicznych.

Antagonisty NK1R znakowane ^{11}C , ^{14}C , ^{18}F stosowane są także w PET do obrazowania stopnia wysycenia receptorów. [^{11}C]GR205171 został wykorzystany do mapowania NK1R w mózgu w różnych zaburzeniach lękowych, stresowych i padaczkowych [189–191]. [^{18}F]SPA-RQ (pochodna vofopitantu) wykorzystywany był w ilościowym oznaczaniu NK1R w mózgach pacjentów z lękiem napadowym, czy skutecznej ocenie zmian raka trzustki [167,192]. Mapowanie wysycenia NK1R z wykorzystaniem PET za pomocą radioznacznika [^{11}C]GR205171 obejmowało także ocenę korelacji stężenia casopitantu w osoczu z zajęciem receptorów oraz oszacowanie dawki casopitantu do kolejnych badań klinicznych [193]. Podobne badania z użyciem [^{11}C]GR205171 były stosowane do określenia wysycenia receptorów NK1R, bezpieczeństwa, tolerancji i właściwości farmakologicznych rolapitantu [194]. W trakcie prac rozwojowych określenie biodostępności (właściwości

wchłaniania, dystrybucji, metabolizmu) możliwe jest dzięki zastosowaniu antagonistów znakowanych ^{14}C , przykładowo [^{14}C]aprepitant, [^{14}C]casopitant, [^{14}C]netupitant [167].



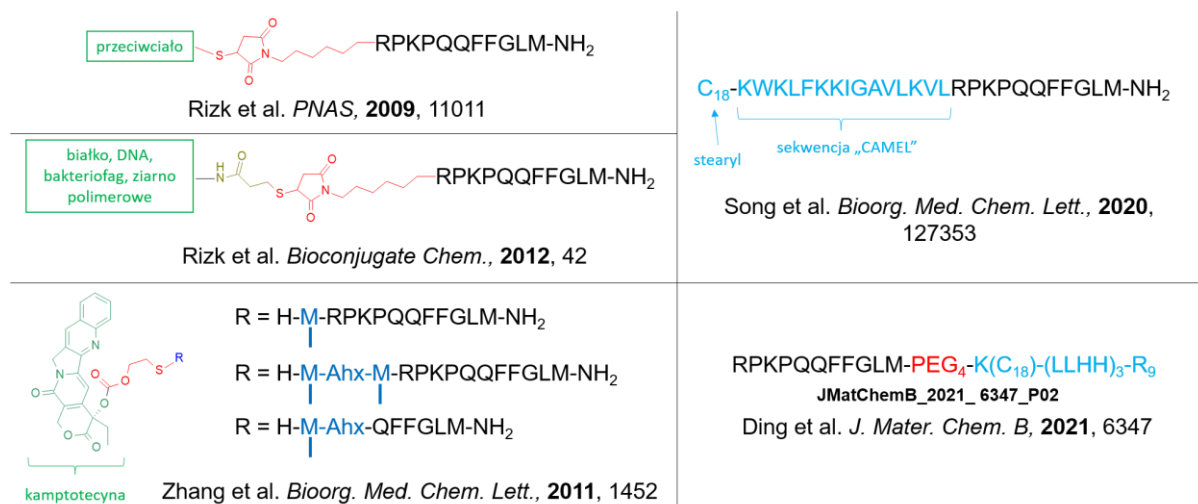
Rysunek 14. Wybrane radiofarmaceutyki oparte na strukturze antagonisty NK1R.

II.2.3.5. Receptor NK1R w celowanym dostarczaniu elementu (cargo) terapeutycznego (innego niż radionuklid)

Poza omówionym powyżej celowanym dostarczaniem radionuklidów diagnostycznych/terapeutycznych, receptor NK1R jest także punktem uchwytu dla dostarczania innego rodzaju elementów terapeutycznych (ang. *therapeutic cargos*). Podobnie jak w przypadku zastosowań radiofarmaceutycznych, chodzi tutaj o opracowanie skutecznych i bezpiecznych terapii nowotworów. Podejście to opiera się na 1) zjawisku internalizacji ligandów NK1R, a także na 2) wielokrotnie obserwowanej nadekspresji NK1R na powierzchni komórek nowotworowych.

Rizk i wsp. przedstawili wykorzystanie modyfikowanej Substancji P do dostarczania do tkanki nowotworowej syntetycznych fragmentów przeciwciała do aktyny [195]. Fragmenty te (wybrane za pomocą eksperymentów „phage display”) miały wiązać się z różnymi konformacjami aktyny i wpływać na organizację cytoszkieletu. Strukturalną

modyfikacją umożliwiającą doczepienie ładunku było wprowadzenie N-końcowego ugrupowanie maleimidowego na sześciowęglowym łączniku (przez kwas 6-maleimidoheksanowy, Rysunek 15). Ugrupowanie takie umożliwia reakcję z wolnymi łańcuchami bocznymi cysteiny na powierzchni białka i wytworzenie połączenia tioeterowego. Lokalizacja linkera i ładunku na N-końcu podyktowana była znanym faktem, iż wysokie powinowactwo SP do NK1R zależy głównie od obecności w sekwencji kilku C-końcowych reszt aminokwasowych. Prezentowane koniugaty SP-białko wykazywały (*in vitro*) internalizację w komórkach nowotworowych glejaka U87, zdolność do „ucieczki” z endosomu, kolokalizację z cytoszkieletem aktynowym oraz znacząco obniżały przeżywalność komórek U87. Jednocześnie koniugaty nie były internalizowane przez normalne astrocyty (nieeksprymujące znacznych ilości NK1R).



Rysunek 15. Struktury niektórych koniugatów lub nośników opartych na strukturze SP.

Kontynuując omówione badania, Rizk i wsp. udowodnili, że wektory oparte na SP mogą służyć do dostarczania dużego białka do cytozolu lub do jądrowego dostarczania fragmentu DNA (kodowane przez nie białko było eksprymowane przez komórki), a nawet cząsteczek tak dużych jak np. cały bakteriofag albo polistyrenowe ziarno o średnicy 1 μM [196]. Dostarczenie DNA do jądra umożliwiające zostało przez wprowadzenie do sekwencji SP na N-końcu sygnału lokalizacji jądrowej (NLS, ang. *nuclear localization signal*, Rysunek 15). Dodatkowo, dołączanie elementów terapeutycznych pozbawionych powierzchniowo dostępnych łańcuchów bocznych cysteiny, możliwe były dzięki użyciu heterobifunkcyjnego łącznika krzyżowego SPDP (estru *N*-sukcynimidowego kwasu 3-(2-pyridylotio)propionowego), który z jednej strony przyłącza się do powierzchniowych reszt lizynowych białka, z drugiej zaś umożliwia przyłączenie peptydu przez maleimid.

Zhang i wsp. wykorzystali SP i skrócony peptyd SP(6-11) do dostarczania do tkanki nowotworowej kamptotecyny (CPT; eksperymentalny lek przeciwnowotworowy; inhibitor topoisomerazy, Rysunek 15) [197]. Molekuła CPT (jedna lub dwie) była przyłączana do peptydu na N-końcu przez łącznik zawierający ugrupowania węglanowe i disiarczkowe, które miałyby ulegać rozpadowi w obecności dużych stężeń glutationu po internalizacji do komórki. Przedstawione koniugaty wykazywały cytotoksyczność na komórki nowotworowe (eksprymujące NK1R) porównywalną lub wyższą niż wolna CPT. Jednocześnie były one znacznie bardziej selektywne w teście cytotoksyczności, gdzie porównywano komórki CHO-NK1 i CHO bez ekspresji NK1R.

Co ciekawe, badacze zaobserwowali pewien stopień internalizacji omawianych koniugatów nawet w przypadku komórek CHO nietransfekowanych NK1R, co sugerowałoby współdziałanie jakiegoś niereceptorowego (niselektywnego) mechanizmu internalizacji. W związku z tym, CPT była analogicznie łączona z peptydowym antagonistą NK1R, [D-Arg¹, D-Trp^{5,7,9}, Leu¹¹]SP [198], który miałby odgrywać rolę peptydu penetrującego komórki. Koniugat ten wykazał (*in vitro*) dobrą zdolność dostarczania CPT do komórek CHO nietransfekowanych NK1R. Był także cytotoksyczny wobec komórek nowotworowych (nie badano selektywności). Podobne badania przeprowadzono z *N*-stearylowaną [D-Arg¹, D-Trp^{5,7,9}, Leu¹¹]SP pod kątem dokomórkowego dostarczania plazmidu.

W innej pracy przedstawiono hybrydowy peptyd złożony z *N*-stearylowanej sekwencji „CAMEL” i Substancji P (Rysunek 15) do dostarczenia plazmidu p53 do komórek wykazujących ekspresję NK1R [199]. Kompleksy takie były (*in vitro*) selektywnie cytotoksyczne wobec komórek transfekowanych NK1R.

Kolejna próba inspirowana nadekspresją NK1R na komórkach glejaka została zaprezentowana przez Ding i wsp. [200]. Zsyntezowali oni peptydowe wektory (przykładowo związek JMatChemB_2021_6347_P02, Rysunek 15) złożone z *N*-acetylowanej SP, linkera polietylenoglikolowego (PEG), *N*^ε-stearylolizyny (K(C₁₈)), segmentu odpowiedzialnego za ucieczkę z endosomu ((LLHH)₃) oraz segmentu penetrującego komórkę (R₉). Wektory te były kompleksowane z plazmidem pGL3, a następnie oceniono (*in vitro*) wydajność internalizacji oraz transfekcji. Najlepsze wyniki uzyskano dla wektora P02. Dla tego wektora, potwierdzono zdolność do przechodzenia przez BBB *in vitro* (w modelu z komórek szczurzych) oraz *in vivo* na Danio pręgowanym. Choć w obecności nadmiaru wolnej SP wydajność transfekcji oraz internalizacji spadała, co sugerowałoby udział receptora NK1R w procesie internalizacji omawianych koniugatów, struktura tych koniugatów (rozbudowa SP

na C-końcu sekwencji) powoduje, że w świetle znanych zależności SAR oraz struktury krystalograficznej NK1R z SP [24,25] bezpośrednie oddziaływanie wektora z miejscem wiążącym receptora NK1 jest, moim zdaniem, wątpliwe.

Należy zauważyć, że opisane powyżej próby wykorzystania celowania w NK1R do specyficznego dostarczania ładunku terapeutycznego do komórek, miały charakter wstępny i ograniczały się jedynie do badań *in vitro* (na pojedynczych parach linii komórkowych nowotwór *vs* komórki prawidłowe). Tymczasem jedynie badania *in vivo* (na modelu ssaczym) pozwoliłyby na realistyczne oszacowanie farmakokinetyki tego typu koniugatów. Szczególnie wydaje się, że niestabilność (wobec peptydaz) peptydowego nośnika opartego na strukturze SP może uniemożliwiać uzyskanie okresu półtrwania dostatecznie długiego dla dystrybucji, wniknięcia do komórek i uzyskania efektu leczniczego. Innym problemem może być polarność peptydowego nośnika, która może negatywnie wpływać na dystrybucję. W świetle tego, korzystne mogłoby być wykorzystanie nośników opartych na strukturach małowładczkowych ligandów NK1R, podobnie jak miało to miejsce w przypadku zastosowań radiofarmaceutycznych. Dodatkowo, warto zauważyć, że w omawianych pracach brakowało sprawdzenia powinowactwa wektorów i/lub koniugatów do NK1R. Jest to o tyle ważne, że doświadczenia z badań nad hybrydowymi ligandami opioidowo-NK1R lub z zakresu radiofarmacji pokazują, że hybrydyzacja fragmentów może wiązać się z utratą powinowactwa receptorowego. Nie przebadano także mechanizmu internalizacji omawianych koniugatów pod kątem tego, czy na pewno jest on związany z NK1R.

II.3. Funkcje oceniające

II.3.1. Dokowanie molekularne

II.3.1.1. Rola dokowania we współczesnym poszukiwaniu leków

Kluczową rolę we współczesnym poszukiwaniu nowych substancji biologicznie aktywnych odgrywają metody modelowania molekularnego (metody *in silico*). Jedną z technik takiego modelowania jest dokowanie molekularne. Dokowanie molekularne przewiduje czy i w jaki sposób dana cząsteczka (ligand) będzie oddziaływać z określonym celem molekularnym (np. receptorem białkowym). Dokowanie dostarcza informacji o prawdopodobnych sposobach ułożenia liganda w miejscu wiążącym celu molekularnego (pozycja, konformacja, rodzaj oddziaływań) oraz o szacowanej sile wiązania kompleksu ligand-cel molekularny [201].

Uzyskane w wyniku dokowania molekularnego predykcje, dotyczące struktury kompleksu ligand-cel molekularny oraz występujących w takim kompleksie oddziaływań, stanowią ważne informacje w projektowaniu nowych substancji aktywnych. Na podstawie danych otrzymanych z dokowania można np. projektować potencjalne modyfikacje strukturalne znanych związków aktywnych, np. w fazie optymalizacji związku wiodącego (ang. *lead optimization*), które mogą mieć korzystny wpływ na oddziaływanie z celem molekularnym, co daje możliwość projektowania związków o lepszych właściwościach farmakologicznych.

Dokowanie molekularne znajduje także zastosowanie w przesiewie wirtualnym opartym na strukturze (ang. *structure-based virtual screening*, SBVS) [202]. Celem takiej procedury jest identyfikacja nowych związków wiodących (ang. *lead compounds*). W SBVS, dokuje się do receptora bardzo duże biblioteki różnorodnych związków chemicznych, identyfikując te, które mają największe szanse na wykazanie wiązania do receptora. Dzięki dokowaniu w SBVS możliwe jest względnie szybkie zawężenie grupy związków, które warto przebadać eksperymentalnie, co skraca czas i koszty związane z poszukiwaniem nowych związków wiodących.

Dokowanie molekularne jest również wykorzystywane w projektowaniu cząsteczek *de novo* [203]. W takiej procedurze, program komputerowy generuje propozycje nowych (wirtualnych) związków, które nie mają analogii strukturalnej do znanych ligandów danego receptora. Dzięki symulacjom dokowania, oceniany jest potencjał wiązania takich

wirtualnych ligandów z wybranym celem molekularnym. Projektowanie *de novo* pozwala na odkrywanie całkowicie nowych struktur substancji czynnych.

II.3.1.2. Proces dokowania

Istotą procesu dokowania molekularnego jest wygenerowanie struktury kompleksu receptor-ligand na podstawie struktury receptora i liganda w stanie niezwiązany. Dokowanie przebiega w dwóch etapach: 1) generowanie wielu możliwych geometrii kompleksu ligand-cel molekularny za pomocą algorytmów przeszukiwania (ang. *search algorithms*) oraz 2) ocena tych konfiguracji za pomocą funkcji oceniających (ang. *scoring functions*), która polega na szacowaniu siły wiązania danych konformacji kompleksu ligand-cel molekularny i znalezieniu konfiguracji najkorzystniejszej energetycznie (o najniższej energii wiązania) [204]. Istnieje szereg programów do dokowania molekularnego, spośród których do najpopularniejszych należą AutoDock [205], AutoDock Vina [206], DOCK [207], FlexX [208], GOLD [209], Glide [210]. Wszystkie działają na podobnej zasadzie, lecz różnią się stosowanymi algorytmami przeszukiwania i oceniania.

Proces dokowania wymaga odpowiedniego przygotowania struktur liganda oraz celu molekularnego, którym najczęściej jest receptor białkowy. Konieczna jest znajomość trójwymiarowej struktury obu molekuł. Przygotowanie liganda do dokowania obejmuje wygenerowanie jego struktury oraz optymalizację geometrii. Podczas tego procesu należy uwzględnić kwestie protonowania, stereochemii, ewentualnej tautomerii oraz konformacji liganda.

Jeśli chodzi o strukturę receptora, w dokowaniu najczęściej stosuje się struktury rozwiązane eksperymentalnie, z wykorzystaniem metod rentgenografii strukturalnej, spektroskopii magnetycznego rezonansu jądrowego lub mikroskopii krioelektronowej. Te empiryczne struktury białek można znaleźć w bazach danych, z których najpopularniejszą jest PDB [211]. W przypadku braku informacji o pełnoatomowej strukturze białka możliwe jest przeprowadzenie dokowania do struktur otrzymanych w wyniku modelowania homologicznego, w którym tworzy się model docelowego białka na podstawie empirycznej struktury innego białka o zbliżonej sekwencji (homologii). Struktury receptora muszą zostać przygotowane do dokowania przez np. ustalenie stanów protonacji jonizowalnych grup funkcyjnych. Przed przeprowadzeniem symulacji dokowania kluczowe jest również określenie miejsca wiążącego, w którym ligand będzie się wiązał z białkiem. Istotnym aspektem jest także ewentualne uwzględnienie obecności cząsteczek wody w obszarze wiążącym, które mogą wpływać na geometrię i stabilność kompleksu ligand-receptor [212].

W większości standardowych protokołów dokowania, obecność cząsteczek wody nie jest brana pod uwagę, co obniża koszt obliczeniowy symulacji, lecz może prowadzić do uproszczonych lub błędnych predykcji sposobu wiązania.

Jeśli chodzi o przeszukiwanie przestrzeni rozwiązań, dokowanie molekularne, w zależności od sposobu uwzględniania zmienności konformacyjnej liganda i receptora, może być realizowane na trzy różne sposoby [213]:

1) *dokowanie sztywne*, w którym uznaje się, że ligand i receptor są strukturami sztywnymi, konformacje obu molekuł nie ulegają zmianie w czasie dokowania; jedna z cząsteczek podlega rotacjom i translacjom,

2) *dokowanie półgiętkie*, w którym uwzględnia się zmienność konformacyjną ligandów, a konformacja receptora pozostaje stała; cząsteczki liganda poddawane są rotacjom i translacjom;

3) *dokowanie giętkie*, w którym konformacje obu molekuł podlegają zmianom, a dodatkowo cząsteczki liganda poddawane są rotacjom i translacjom. W praktyce, uwzględnienie pełnej giętkości receptora jest zbyt kosztowne obliczeniowo i jest domeną raczej symulacji dynamiki molekularnej. W typowym dokowaniu giętkim, program uwzględnia zmienność konformacyjną pewnych obszarów receptora, np. wybranych łańcuchów bocznych aminokwasów w miejscu wiążącym, co pozwala na bardziej realistyczne modelowanie oddziaływań ligand-receptor w obszarach, gdzie zmienność konformacyjna jest kluczowa.

Współcześnie, większość programów do dokowania wykorzystuje algorytmy przeszukiwania uwzględniające giętkość ligandów. Ogólnie, algorytmy przeszukiwania można podzielić (klasyfikacja przedstawiona m.in. przez Brooijmansa i Kuntza [214]) na algorytmy systematyczne, stochastyczne oraz deterministyczne.

Algorytmy systematyczne opierają się na systematycznym przeszukiwaniu przestrzeni konfiguracyjnej kompleksu ligand-receptor [213]. W celu znalezienia minimum energetycznego, algorytm systematycznie modyfikuje wartości zmiennych opisujących formalne stopnie swobody kompleksu (translacje, rotacje, zmiany konformacyjne), przez zmienianie je w określonych interwałach. W miarę wzrostu liczby stopni swobody liganda, ilość możliwych kombinacji rośnie, co prowadzi do drastycznego wzrostu złożoności obliczeniowej. Dlatego często stosuje się kryteria zakończenia, aby przerwać proces wyszukiwania przed kompletnym przetestowaniem wszystkich możliwych kombinacji. Jednym z typów algorytmów systematycznych jest przeszukiwanie wyczerpujące, które polega na badaniu różnych ułożeń liganda poprzez systematyczne obracanie wszystkich

możliwych wiązań rotacyjnych w ligandzie. Dodatkowo, metoda uwzględnia systematyczne przemieszczanie liganda w przestrzeni translacyjnej. Ze względu na generowanie dużej liczby kombinacji, zwłaszcza przy większej liczbie wiązań rotujących, metoda ta zazwyczaj znajduje zastosowanie w przypadku małych, niezbyt giętkich ligandów [212]. Innym rodzajem algorytmu systematycznego są metody oparte na fragmentacji ligandów [213]. W ramach tego podejścia ligand jest początkowo podzielony na mniejsze fragmenty, spośród których jeden fragment (zazwyczaj największy lub posiadający kluczowe dla wiązania elementy), uznawany jest za fragment bazowy i dokowany jest do miejsca aktywnego jako pierwszy. Kolejne fragmenty przyłączane są stopniowo, z jednoczesnym generowaniem różnych orientacji. Umożliwia to uwzględnienie giętkości liganda i systematyczne przeszukiwanie położenia poszczególnych części liganda. Podobnie jak w przypadku metody przeszukiwania wyczerpującego, metoda fragmentacji jest ograniczona do mniejszych oraz średnich ligandów, w których ilość fragmentów nie byłaby zbyt duża. Alternatywnym podejściem jest dokowanie zespołów konformacyjnych (ang. *ensemble docking*), w którym wstępnie generuje się potencjalne konformacje liganda, a następnie każda konformacja dokowana jest jako sztywna cząsteczka do miejsca wiążącego receptora. Zastosowanie tego podejścia pozwala na redukcję nakładów obliczeniowych związanych z badaniem przestrzeni konformacyjnej liganda, jednakże wymaga użycia dodatkowych narzędzi do generowania jak najbardziej kompletnego zestawu konformacji liganda [212].

Algorytmy stochastyczne, w odróżnieniu od algorytmów systematycznych, przeszukują przestrzeń konformacyjną poprzez losową, wielokrotną modyfikację stopni swobody liganda (konformacyjnych, translacyjnych i rotacyjnych) [212]. Zmiany są następnie akceptowane lub odrzucane zgodnie z przyjętym kryterium probabilistycznym. Zaletą stosowania tego podejścia jest możliwość generowania dużych, różnorodnych zbiorów rozwiązań, co zwiększa prawdopodobieństwo znalezienia globalnego minimum energetycznego w krótszym czasie. Metody wykorzystujące algorytmy genetyczne oraz metoda Monte Carlo (MC) to dwa klasyczne przykłady metod należących do algorytmów stochastycznych. W algorytmach genetycznych informacje o stopniach swobody liganda zapisywane są w postaci ciągów binarnych nazywanych „genami”, które formują „chromosomy”, reprezentujące położenia liganda w przestrzeni. Chromosomy ulegają mutacjom, polegającym na losowych pojedynczych zmianach w genach oraz krzyżowaniu, obejmującym wymianę genów pomiędzy dwoma chromosomami. W wyniku tych działań powstaje nowa konfiguracja liganda w miejscu wiążącym receptora, która jest następnie oceniana przez funkcję oceniającą. Nowe, najbardziej korzystne struktury, które uzyskują

scoring powyżej odpowiedniego progu, biorą udział w tworzeniu kolejnych „pokoleń” obliczeń. Proces ten powtarza się przez wiele „pokoleń”, umożliwiając skuteczne przeszukiwanie przestrzeni konformacyjnej. W metodzie Monte Carlo konfiguracje liganda w miejscu wiążącym generowane są poprzez losowe zmiany we wszystkich stopniach swobody. Otrzymana konfiguracja jest następnie poddawana analizie za pomocą kryterium selekcji opartego na energii. W przypadku spełnienia określonych warunków, jest ona zapisywana i podlega dalszym modyfikacjom w celu wygenerowania kolejnej konformacji. Proces ten powtarza się iteracyjnie aż do osiągnięcia określonego kryterium zbieżności.

W algorytmach deterministycznych aktualny stan systemu określa (determinuje) zmiany, które należy wykonać na strukturze kompleksu, aby mógł zostać wygenerowany kolejny stan na „ścieżce” przeszukiwania możliwych rozwiązań, a który zazwyczaj musi być równy lub niższy pod względem energii w porównaniu do stanu wyjściowego [214]. Algorytmy deterministyczne opierają się na matematycznych modelach, wykorzystujących zasady mechaniki molekularnej do modelowania ruchów i oddziaływań między molekułami na poziomie atomowym. Do algorytmów deterministycznych zaliczyć można te, które wykorzystują różne metody minimalizacji energii lub dynamikę molekularną [213]. Metoda minimalizacji energii ma na celu zbadanie krajobrazu energetycznego, czyli różnych możliwych konformacji molekularnych, i znalezienie najkorzystniejszego energetycznie układu. W trakcie tego procesu system jest naprowadzany w kierunku najniższego lokalnego minimum potencjału energetycznego. Metody dynamiki molekularnej umożliwiają symulowanie ruchów systemu w czasie, przy uwzględnianiu różnych zmiennych termodynamicznych, takich jak temperatura i ciśnienie. Obie te metody wiążą się z pewnymi ograniczeniami, w tym z niemożnością przewyciężenia barier energetycznych i utknięciem w minimach lokalnych oraz wysokimi kosztami obliczeniowymi. Metody oparte na algorytmach deterministycznych stosowane są często w połączeniu z innymi algorytmami przeszukiwania.

II.3.2. Funkcje oceniające

W trakcie dokowania, wygenerowane przez algorytmy przeszukujące możliwe konfiguracje (geometrie) kompleksu ligand–cel molekularny są oceniane przez funkcje oceniające (*scoring*). Celem *scoringu* jest szybkie oszacowanie (według pewnego założonego modelu) energii (entalpii swobodnej) wiązania danego hipotetycznego kompleksu. W ten sposób *scoring* mierzy „jakość” poszczególnych wygenerowanych konfiguracji i wybiera „najlepszą” (o najbardziej ujemnej energii), która ma odpowiadać rzeczywistej strukturze

kompleksu. Idealnie, funkcja oceniająca powinna dobrze szacować energię wiązania kompleksu (powinowactwo/wiązalność), identyfikować najkorzystniejszą energetycznie pozę liganda w miejscu wiążącym (ang. *binding pose*) oraz poprawnie szeregować różne ligandy według ich siły działania.

Ze względu na złożoność i nieaddytywność zjawisk wpływających na siłę wiązania w kompleksach ligand–receptor, szybkie szacowanie energii wiązania na podstawie obliczeń uwzględniających wszystkie ważne z punktu widzenia teorii czynniki nie jest możliwe. Z konieczności zatem funkcje oceniające opierają się na różnych założeniach (modelach) i uproszczeniach, aby w jak najkrótszym czasie uzyskać jak najbardziej zbliżone do rzeczywistych energie wiązania. Istnieje wiele funkcji oceniających, które można podzielić na trzy główne kategorie: 1) oparte na polach siłowych, 2) empiryczne, 3) oparte na wiedzy [215].

Funkcje oceniające oparte na polach siłowych przybliżają entalpię swobodną wiązania przez obliczenie energii kompleksu ligand-receptor z wykorzystaniem wyrażeń znanych z typowych pól siłowych stosowanych w mechanice molekularnej. Konceptyjnie, wyrażenia te mają jasny sens fizyczny, dlatego funkcje takie są określane także jako „oparte na fizyce” (ang. *physics-based*) [216]. Energia wiązania kompleksu obliczana jest poprzez sumowanie wkładów opisujących oddziaływania „niezwiązane” (ang. *non-bonded*) między ligandem a receptorem (oddziaływania van der Waalsa oraz elektrostatyczne) oraz oddziaływania „związane” (ang. *bonded*) wewnątrz liganda. Te drugie dotyczą naprężenia sterycznego w cząsteczce liganda w konformacji przyjętej w miejscu wiążącym receptora. Oddziaływania van der Waalsa opisywane są np. za pomocą potencjału Lennarda-Jonesa, a oddziaływania elektrostatyczne są reprezentowane np. przez potencjał Coulomba. W nowszych, mieszanych funkcjach oceniających, dodawane są także wyrazy opisujące wkład energetyczny wiązań wodorowych jako dodatkowy składnik. Nowsze funkcje mogą uwzględniać również wkłady związane z entropią i solwatacją. Parametry funkcji opartych na polach siłowych zwykle są wyprowadzane zarówno z danych eksperymentalnych, jak i obliczeń *ab initio* mechaniki kwantowej. Ogólny wzór funkcji opartych na polach siłowych można przedstawić za pomocą poniższego równania (Równanie 1) [212]:

$$E_{bnd} = E_{bond} + E_{non-bond} + E_{H-bond} + E_{solv} \quad (1)$$

gdzie E_{bind} – energia wiązania, E_{bond} – wkład energetyczny oddziaływań „związanych” (wewnątrz liganda), $E_{\text{non-bond}}$ – wkład energetyczny oddziaływań „niezwiązanych” (pomiędzy ligandem a receptorem oraz ewentualnie wewnątrz liganda), $E_{\text{H-bond}}$ – wkład energetyczny wiązań wodorowych, E_{solv} – wkład energetyczny związany z solwatacją.

Klasycznym przykładem funkcji opartej na polach siłowych jest funkcja ze starszych wersji programu DOCK (Równanie 2) [217],

$$E = \sum_i \sum_j \left(\left(\frac{A_{i,j}}{r_{i,j}^{12}} \right) - \left(\frac{B_{i,j}}{r_{i,j}^6} \right) + \left(332,0 \frac{q_i^* q_j}{D * r_{i,j}} \right) \right) \quad (2)$$

gdzie: r_{ij} – odległość pomiędzy atomem receptora i a atomem liganda j , $A_{i,j}$ oraz $B_{i,j}$ – parametry sił van der Waalsa, q_i oraz q_j – ładunki atomowe, D – stała dielektryczna. Poszczególne składniki w tym równaniu odpowiadają wyrazom energii „niezwiązanej” z równania energii pola siłowego AMBER.

W empirycznych funkcjach oceniających energia wiązania kompleksu ligand-białko jest przybliżana za pomocą sumy różnych deskryptorów opisujących poszczególne rodzaje oddziaływań występujących w danym kompleksie [204]. Deskryptory te mogą dotyczyć m.in. np. oddziaływań hydrofobowych, dopasowania sterycznego, oddziaływań elektrostatycznych, wiązań wodorowych oraz entropii i desolvatacji. Dla każdego z tych czynników przypisuje się odpowiedni współczynnik wagowy, który jest ustalany na podstawie analizy statystycznej, w tym wielorakiej regresji (ang. *multiple regression*), na danych eksperymentalnych znanych kompleksów ligand-białko, o eksperymentalnie zmierzonej energii wiązania i ustalonej strukturze. Przykładem funkcji empirycznej jest funkcja programu AutoDock Vina (Równania 3–6) [206,218], w której energia wiązania (E) jest sumą energii oddziaływań wszystkich możliwych par atomów liganda i miejsca wiążącego receptora (e_{pair}) podzieloną przez nieliniową „karę” entropową (mianownik Równania 3) związaną z ilością wiązań rotujących w ligandzie (N_{rot}). Energia ta jest zależna od odległości "powierzchniowej" (ang. *surface distance*) między atomami, obliczanej zgodnie z Równaniem 4, gdzie r oznacza odległość międzyatomową, a R_i i R_j to promienie atomów w parze. Energia oddziaływania par (e_{pair}) jest sumą ważoną składników, które uwzględniają, wkład dopasowania sterycznego (pierwsze trzy składniki Równania 5), „oddziaływania” hydrofobowego (czwarty składnik Równania 5) oraz wiązania wodorowego (piąty składnik Równania 5). Czynniki w_1 – w_6 są wagami wkładu poszczególnych składników oddziaływania w modelu. Ich wartości dla oryginalnej funkcji AutoDock Vina są

przedstawione w Tabeli 3. Wielkości tych parametrów zostały dobrane przez kalibrację względem danych eksperymentalnych z bazy PDDBind [219].

$$E = \frac{\sum e_{pair}(d)}{1+w_6*N_{rot}} \quad (3)$$

$$d = r - R_i - R_j \quad (4)$$

$$e_{pair}(d) = \begin{cases} w_1 * Gauss_1(d) + \\ w_2 * Gauss_2(d) + \\ w_3 * Repulsion(d) + \\ w_4 * Hydrophobic(d) + \\ w_5 * HBond(d) \end{cases} \quad (5)$$

Tabela 3. Wagi poszczególnych wkładów energetycznych do siły oddziaływania w modelu domyślnej funkcji oceniającej programu AutoDock Vina.

Czynnik wagowy	w ₁	w ₂	w ₃	w ₄	w ₅	w ₆
Wartość	-0.0356	-0.00516	0.840	-0.0351	-0.587	0.0585

Funkcje oparte na wiedzy (ang. *knowledge-based scoring functions*) do oceny siły wiązania kompleksu ligand-białko wykorzystują dane uzyskane z analiz statystycznych znanych kompleksów białko-ligand, które posiadają eksperymentalnie rozwiązane struktury [204]. W kompleksach tych badane są kontakty pomiędzy poszczególnymi typami atomów. W szczególności, analizowane są częstości występowania kontaktów międzyatomowych i/lub odległości między atomami. Korzystając z prawa Boltzmana, przekształca się te dane w pseudopotencjał nazywany potencjałem średniej siły (ang. *potential of mean force*), który jest wykorzystywany do modelowania preferowanych geometrii par atomów różnych typów w kompleksie ligand-białko (Równanie 6):

$$A_{ij}(r) = -k_B T \ln p_{ij}(r) \quad (6)$$

gdzie $A_{ij}(r)$ – wartość potencjału średniej siły oddziaływania grup i, j , k_B – stała Boltzmana, T – temperatura, $p_{ij}(r)$ – prawdopodobieństwo znalezienia grup i, j w odległości r [212].

Funkcje oparte na wiedzy, zakładają, że oddziaływania, które występują częściej w eksperymentalnie rozwiązanych kompleksach białko-ligand, są bardziej korzystne energetycznie dla wiązania. Dlatego też, ostateczny wynik jest obliczany poprzez faworyzowanie preferowanych (statystycznie częstszych) kontaktów i „karanie” oddziaływań

odpychających pomiędzy atomami liganda i białka. *Scoring* danego kompleksu jest sumą wartości potencjału po wszystkich (w obrębie pewnej odległości od liganda) parach atomów i j w kompleksie (Równanie 7):

$$score = \sum_i \sum_j A_{ij}(r) \quad (7)$$

Przykładami funkcji opartych na wiedzy są np. funkcja DSX [220] albo funkcja PMF [221].

Pomimo dostępności wielu funkcji oceniających, żadna z nich nie zapewnia dokładnej oceny siły wiązania kompleksu ligand-białko. Funkcje oceniające opierają się na różnych założeniach i uproszczeniach dotyczących energetyki wiązania (m.in. założenie o addytywność oddziaływań międzycząsteczkowych, często nieprawidłowy, uproszczony opis tych oddziaływań i ich termodynamiki), co wpływa też na jakość predykcji. W celu zredukowania ograniczeń pojedynczych funkcji oceniających, możliwe jest zastosowanie podejścia konsensusowego (ang. *consensus scoring*) [212]. Polega ono na łączeniu wyników z różnych funkcji oceniających dla uzyskania bardziej wiarygodnych przewidywań siły wiązania [204].

Warto przy okazji dodać, że z racji wspomnianej powyżej niedokładności funkcji oceniających, *scoring* jest w praktyce rzadko jedynym kryterium oceny wyników dokowania i podejmowania decyzji „co dalej?” (np. co do wyboru związków do zakupu lub syntezy). Powszechną praktyką jest wizualna inspekcja pól z dokowania [222], porównywanie z dostępnymi strukturami krystalograficznymi dla analogicznych kompleksów, sprawdzenie spełnienia określonych więzów eksperymentalnych [223] (np. oddziaływania z konkretnymi aminokwasami wywiedzione z eksperymentów kierowanej mutagenyzy), porównywanie z danymi SAR dla podobnych związków, prowadzenie dodatkowych obliczeń (np. dynamiki molekularnej [224]), stosowanie filtrów opartych o *fingerprinting* oddziaływań [225,226] *etc.*

Nowszą alternatywą dla klasycznych funkcji oceniających są funkcje oparte na uczeniu maszynowym (ang. *machine learning scoring functions*, MLSF), które wykorzystują zaawansowane algorytmy trenowania (m.in. bez predefiniowania postaci modelu) i nieliniowe modele. Funkcje te mają na celu ujęcie nieliniowych zjawisk występujących w oddziaływaniach ligand-białko, które są trudne do modelowania za pomocą tradycyjnych metod [212]. Dodatkowo, MLSF – w przeciwieństwie do klasycznych liniowych modeli – są w stanie „zaabsorbować” informacje zawarte w dużych zestawach danych [227] i w ten sposób poprawiać skuteczność wraz ze wzrostem wielkości zestawu danych treningowych. W wielu testach, MLSF dały znacznie lepsze wyniki jeśli chodzi o moc przesiewania czy korelację z powinowactwem niż klasyczne funkcje oceniające [228,229].

MLSF są jednak przedmiotem krytyki ze strony niektórych autorów, którzy zarzucają im m.in. przeuczenie (ang. *overfitting*), brak sensu fizycznego, brak prawdziwej mocy predykcyjnej itp. [230–235]. Według Shena i wsp. większość ogólnych funkcji oceniających opartych na uczeniu maszynowym daje dużo gorsze lub gorsze wyniki w przesiewie niż klasyczne funkcje oceniające [236]. Proponenci MLSF odpierają jednak te zarzuty [229]. Pojawiają się też perspektywne zastosowania MLSF [237].

II.3.3. Funkcje oceniające specyficzne dla celu molekularnego

Typowe, powszechnie używane funkcje oceniające mają być – w zamyśle ich twórców – stosowalne uniwersalnie. Oznacza to, że powinny dawać dobre oszacowania entalpii swobodnej wiązania dla różnorodnych ligandów i różnorodnych celów molekularnych. W związku z tym, w przypadku funkcji empirycznych i opartych na wiedzy, dobór współczynników równania funkcji i konstrukcja potencjałów odbywa się z wykorzystaniem danych eksperymentalnych (powinowactwa i/lub struktur) uzyskanych dla wielu bardzo różnych białek pochodzących z różnych rodzin. W przypadku funkcji opartych na polach siłowych („na fizyce”), uniwersalność funkcji ma pochodzić z uniwersalności praw fizyki ujętych liczbowo w równaniu pola.

W praktyce zaobserwowano jednak, że funkcje oceniające prezentują różną skuteczność zależnie od konkretnego celu molekularnego [238–241]. Dalej, Ross i wsp., wykorzystując teorię uczenia statystycznego i teorię informacji, wykazali formalnie, że empiryczne modele powinowactwa „ogólne” muszą być bardziej niedokładne niż modele trenowane specyficznie dla poszczególnych białek, a wynika to stąd, iż zbiory danych strukturalnych i powinowactwa dla różnych par białko-ligand są rządzone różnymi rozkładami prawdopodobieństwa [242].

W związku z tymi obserwacjami zaproponowano opracowywanie funkcji oceniających specyficznych (dostrojonych) dla konkretnego celu molekularnego lub rodziny celów molekularnych (ang. *target-specific scoring function*, TSSF) [243]. W literaturze opisano wiele podejść do tego problemu, zróżnicowanych co do

1) sposobu „dostrajania”, np.

- *przez rozszerzanie* klasycznej funkcji oceniającej,
- *przez rekaliibrację* współczynników (wag) klasycznej funkcji oceniającej,
- *przez podejścia mieszane* i nietypowe,

- *przez uczenie maszynowe* tworzące nowe funkcje (bez predefiniowanej postaci);
- 2) **celu optymalizacji**, np.
- poprawa zdolności wyboru pozycji natywnej spośród innych możliwych (ang. *docking power*),
 - poprawa korelacji predykcji entalpii swobodnej z eksperymentem (ang. *scoring power*),
 - poprawa mocy rankingowania (ang. *ranking power*),
 - poprawa mocy przesiewania. (ang. *screening power*);
- 3) **natury zbiorów danych używanych do trenowania funkcji i walidacji**, np.
- struktury krystalograficzne vs struktury z dokowania albo mieszanina obu typów,
 - dane tylko strukturalne lub tylko dane o powinowactwie,
 - dane strukturalne połączone z danymi o powinowactwie;
- 4) **źródła danych o powinowactwie lub kategorii (aktywny/nieaktywny)**, np.
- wszystkie dane o powinowactwie/kategorii z eksperymentu,
 - dane związków aktywnych z eksperymentu, a zbiór związków nieaktywnych („przynęt”, „wabików”, ang. *decoys*) wygenerowany komputerowo jako zestaw związków podobnych fizykochemicznie (ang. *property-matched*) do zestawu aktywnego, lecz o innej topologii cząsteczek (stąd zresztą przypisanie im kategorii nieaktywnych jest uzasadnionym, ale jednak tylko założeniem);
- 5) **zastosowania dostrojonych funkcji**, np.
- ocena pól na etapie przeszukiwania przestrzeni rozwiązań oraz szacowanie energii,
 - *rescoring* pól wybranych przez klasyczne funkcje oceniające,
 - *rescoring* pól wybranych przez klasyczne funkcje oceniające w ramach oceny konsensusowej.

Spośród prób tworzenia TSSF **przez rozszerzanie**, można wymienić podejście AutoShim przedstawione przez Martin i Sullivana [244]. Polega ono na rozszerzeniu dowolnej ogólnej funkcji oceniającej o wyrazy (*shimy*) związane ze spełnieniem pewnych (specyficznych dla receptora) wymagań farmakoforowych. Zarówno wybór punktów

farmakoforowych, jak i wagi poszczególnych *shimów* jest przedmiotem procedury optymalizacyjnej, która dopasowuje model do danych eksperymentalnych. Metodologia ta została użyta do dostrojenia funkcji oceniającej specyficznej dla kinazy PDK1, która umożliwiła odkrycie 5 submikromolowych inhibitorów spośród biblioteki 100000 zadokowanych związków (23 przebadane eksperymentalnie).

Santos-Martins i wsp. rozszerzyli pole siłowe (funkcję oceniającą) programu AutoDock o dodatkowe wyrażenia opisujące oddziaływanie kationu cynku z grupami wiążącymi cynk [245]. Parametry równania zostały skalibrowane wobec zbioru eksperymentalnych struktur białek zawierających cynk w miejscu wiążącym (katalitycznym). Opracowana funkcja, nazwana AutoDock4_{Zn}, jest specjalnie przeznaczona do dokowania małych cząsteczek do metaloenzymów cynkowych. W tym zastosowaniu AutoDock4_{Zn} wykazuje znaczną poprawę, jeśli chodzi o przewidywane geometrie kompleksów (szczególnie koordynację kationu cynku) oraz korelację *scoringu* z powinowactwem.

Ważną grupą TSSF są funkcje oceniające uzyskane **przez rekaliibrację** współczynników klasycznych funkcji empirycznych. Wśród wczesnych prób tego typu jest praca Logeana i wsp., którzy zoptymalizowali współczynniki równania funkcji Fresno do przewidywania wiązania peptydów do białka HLA-B2705 z grupy białek MHC [246], uzyskując poprawę korelacji predykcji z eksperymentalną entalpią swobodną (zbiór 26 oligopeptydów).

Niektóre programy dokujące zawierają elementy ułatwiające rekaliibrację domyślnego *scoringu* względem dostępnych danych eksperymentalnych. Program Surflex-Dock został wyposażony w specjalny moduł umożliwiający użytkownikowi optymalizację współczynników funkcji oceniającej pod kątem dopasowania do eksperymentu [247]. Pham i Jain wykazali na kilku zbiorach walidacyjnych, że tego typu dostrojenie prowadzi do poprawy zdolności przesiewania z użyciem tego programu [247].

Twórcy programu eHiTs również zaprezentowali metodykę dostrajania funkcji oceniającej (suma ważona 20 zmiennych; funkcja jest mieszaną funkcją empiryczną/opartą na wiedzy) do konkretnych celów molekularnych [248], która została wykorzystana podczas opracowywania tego programu. Opisano warianty *tuningu* mające na celu poprawę skuteczności wyboru pozy, skuteczności przesiewania, albo korelacji z powinowactwem. Program eHiTs zawiera domyślnie wiele schematów wag *scoringu*, które są wybierane automatycznie podczas dokowania zależnie od charakterystyki miejsca wiążącego. Dodatkowo umożliwia on samodzielne dopasowanie funkcji oceniającej przez użytkownika. Autorzy wykazali, że dopasowane funkcje oceniające poprawiają jakość wyników dokowania

w porównaniu do wyników uzyskanych z użyciem „ogólnej” funkcji oceniającej programu. Dodatkowo, dla funkcji dostrojonych eHiTs wykazano większą skuteczność przesiewania niż w przypadku niektórych ogólnych funkcji oceniających związanych z innymi programami dokującymi.

Funkcja programu eHiTS była punktem wyjścia do opracowania funkcji oceniającej specyficznej dla reduktazy InhA przez Kinnings i wsp. [249]. Autorzy wykorzystali metodę maszyny wektorów nośnych (ang. *support vector machine*) do utworzenia nowych funkcji o poprawionej korelacji z powinowactwem i zdolności przesiewowej.

Antes i wsp. zaproponowali metodę POEM (**P**arameter **O**ptimization using **E**nsemble **M**ethods) [250], która łączy procedury typowe dla projektowania eksperymentów (ang. *design-of-experiments*) z różnymi metodami regresji. Autorzy wykorzystali to podejście do uzyskania TSSF (do wyboru prawidłowej pozy) dla kinaz i ATPaz przez dostrojenie współczynników funkcji FlexX and Screenscore. Trenowanie i testowanie odbyło się na zestawach kilku/kilkunastu struktur białek z tych rodzin z bazy PDB. Modyfikowane funkcje dawały pozy wiążące bardziej zbliżone do danych eksperymentalnych niż funkcje oryginalne.

Seifert zaproponował podejście TOP (**T**arget-specific **O**ptimization) jako uniwersalną metodę tworzenia funkcji oceniających specyficznych dla konkretnych celów molekularnych lub rodzin białek [251,252]. W podejściu tym iteratywnie zmienia się współczynniki funkcji oceniającej (za pomocą algorytmu DIRECT będącego algorytmem typu przeszukiwanie tabu) dążąc do maksymalizacji rozdzielania rozkładów *scoringu* ligandów prawdziwych i molekuł-wabików. Rozdzielenie rozkładów mierzone jest parametrami zaczerpniętymi z analizy wariancji ANOVA. W omawianych pracach zoptymalizowano empiryczne funkcje oceniające LUDI oraz ProPose wobec kilku celów molekularnych, uzyskując znaczącą poprawę mocy przesiewania na dużych zbiorach walidacyjnych ligandów i molekuł-wabików.

Wang i wsp. opracowali grupę funkcji (do *rescoringu*) TS-Chemscore, specyficznych dla 14 różnych celów molekularnych [253]. TS-Chemscore powstały przez dostrajanie parametrów empirycznej funkcji Chemscore z programu GOLD. Zbiory danych eksperymentalnych zawierały od 100 do ponad 700 związków aktywnych (podzielone 3:1 na zbiory treningowe i walidacyjne). Dodatkowo funkcje były walidowane na zbiorach zawierających molekuły-wabiki (ang. *decoys*). TS-Chemscore był także zwalidowany prospektywnie w przesiewie biblioteki związków wobec kinazy JAK3 i fosfatazy YopH, w którym to umożliwił odkrycie kilku nowych związków aktywnych.

Shamsara i Shamsara opracowali funkcje *rescorujące* dla sześciu celów molekularnych na bazie funkcji programu AutoDock Vina [254]. Wagi oryginalnej funkcji

zostały dostrojone za pomocą optymalizatora opartego na algorytmie ewolucyjnym (*natural evolution strategy*). Trenowanie zostało przeprowadzone na danych z bazy DUD-E [255], a walidacja dotyczyła skuteczności przesiewu. Efektywność nowych funkcji była zróżnicowana zależnie od celu molekularnego. Co ciekawe, niektóre zmodyfikowane wagi były drastycznie różne od oryginalnych.

Rzęsikowska i wsp. opracowały funkcje oceniające (do *rescoringu*) specyficzne dla 24 białek z klasy A rodziny receptorów GPCR [256]. W tym celu współczynniki (wagi) funkcji oceniającej programu AutoDock 4.2 zostały skalibrowane za pomocą algorytmu genetycznego do danych eksperymentalnych o powinowactwie z bazy ChEMBL [257,258] (warto zauważyć, że zarówno trenowanie jak i podwójna walidacja były przeprowadzone na małych liczbowo zbiorach po 36 cząsteczek). Nowe funkcje wykazały lepszą zdolność do klasyfikacji ligandów (w kilka kategorii powinowactwa) niż oryginalna funkcja AutoDock.

Spiegel i Senderowitz utworzyli TSSF przez dostrajanie współczynników empirycznej funkcji ChemPLP (program GOLD) [259]. W ich podejściu kalibracja równania wobec danych eksperymentalnych jest wykonywana za pomocą algorytmu Enrichment Optimization Algorithm nakierowanego na maksymalizację wzbogacenia (ang. *enrichment*). Autorzy wykonali taką optymalizację dla 6 celów molekularnych, wykorzystując zbiory ligandów i molekuł-wabików z bazy DUD-E [255]. Nowe funkcje były używane do *rescoringu* póź wybranych oryginalną funkcją, a także do dokowania. Skuteczność przesiewania nowych funkcji była zróżnicowana zależnie od celu molekularnego.

W pewnym sensie do grupy TSSF tworzonych przez **rekalibrację** domyślnych funkcji można zaliczyć jeszcze „małe” modyfikacje funkcji czynione niesystematycznie przez różnych autorów bez prowadzenia szerszej walidacji. Przykładowo, Buzko i wsp. zmieniał sposób obliczania wkładu wiązania wodorowego w funkcji AutoDock’a w celu poprawy dokowania ligandów do kinaz [260].

Również **funkcje oceniające oparte na wiedzy** doczekały się wersji specyficznych, **reaklibrowanych** dla konkretnych celów molekularnych.

Xue i wsp. przedstawili specyficzny dla kinaz potencjał Kinase-PMF [261]. Funkcja ta jest analogiczna w swojej budowie do „ogólnego” potencjału PMF04 [221], ale została wytrenowana na zestawie eksperymentalnych struktur ponad 800 kompleksów inhibitor-kinaza (bez brania pod uwagę eksperymentalnej wiązalności). Na zbiorze testowym, funkcja wykazywała lepszą (w porównaniu do „ogólnego” PMF04) zdolność do wyboru natywnej konformacji z zestawu konformacji wygenerowanych przez program AutoDock. W dodatkowym teście, autorzy użyli funkcji do przesiewu małej biblioteki wobec kinazy CDK2

i stwierdzili znaczną poprawę zdolności przesiewania w porównaniu do kilku innych popularnych funkcji oceniających.

SAM-score, kolejną TSSF będącą modyfikacją funkcji opartych na wiedzy przedstawili Wang i wsp. [262]. Funkcja ta jest specyficzna dla metylotransferaz zależnych od (S)-adenozyl-L-metioniny. Autorzy rekalibrowali potencjał iPMF [263] (modyfikacja potencjału PMF) na danych dla 12 enzymów z tej rodziny (562 ligandy o znanym powinowactwie). Zbiór danych strukturalnych do trenowania pochodził z dokowania programem AutoDock Vina albo ze struktur krystalograficznych. Rekalibrowany potencjał posłużył do konstrukcji SAM-score z użyciem metody maszyny wektorów nośnych. Na zbiorze testowym, nowa funkcja (jako element *rescoringu*) wykazywała lepszą korelację z eksperymentalnym powinowactwem niż funkcje programów AutoDockVina czy Glide. SAM-score został użyty do przesiewu wirtualnego biblioteki SPECS (około 200 tysięcy związków) wobec metylotransferazy DOT1L. Wyselekcjonowane w ten sposób 171 związków zostało przebadanych eksperymentalnie, z czego 21 wykazywało mierzalną inhibicję przy 50 μM . Analiza dodatkowych analogów jednego z *hitów*, pozwoliła na znalezienie inhibitorów o IC_{50} rzędu pojedynczych mikromoli.

Pfeffer i Gohlke przedstawili funkcję DrugScore^{RNA} [264] dostrojoną do oceny kompleksów RNA-ligand. Statystyczna analiza kontaktów międzycząsteczkowych w krystalograficznych strukturach kompleksów RNA-ligand i RNA-białko pozwoliła na opracowanie potencjałów międzyatomowych specyficznych dla oddziaływań z RNA, które zostały włączone do funkcji DrugScore^{RNA}. Nowa funkcja cechowała się poprawą skuteczności wyboru prawidłowej geometrii wiązania i poprawą korelacji z powinowactwem eksperymentalnym w porównaniu do „ogólnych” funkcji DrugScore czy DrugScore^{CSD}.

Inne oparte na wiedzy funkcje oceniające specyficzne dla oddziaływań ligandów z RNA to np. AnnapuRNA zaproponowana przez Stefaniaka i Bujnickiego (oparta o potencjały pseudoatomowe) [265] albo funkcja SPA-LN przedstawiona przez Yan i wsp. [266]. Inne podejście do modelowania oddziaływań z kwasami rybonukleinowymi przedstawili Chen i wsp., którzy uzyskali funkcje oceniające iMDLScore1 i iMDLScore2 przez optymalizację parametrów równania funkcji AutoDock 4.1 względem danych eksperymentalnych o powinowactwie [267].

Ciekawym przykładem **nietykowego** TSSF jest PocketScore, funkcja specyficzna dla poszukiwania inhibitorów białka HSP90 [268]. PocketScore obliczany jest programem AutoDock Vina na uproszczonej strukturze kompleksu ligand-białko, gdzie ze strony receptora uwzględnione są tylko dwa aminokwasy miejsca wiążącego oraz jedna cząsteczka

wody obecna w miejscu wiążącym. Reszta białka jest ignorowana. Do obliczenia wykorzystuje się domyślną funkcję Viny (patrz Równania 3-5) z zastosowaniem domyślnych wartości dla wag w_1-w_5 (Tabela 3), natomiast bez uwzględnienia wkładu giętkości cząsteczki liganda (waga $w_6 = 0$). Funkcja PocketScore wykazywała dobrą skuteczność zarówno jako element kierujący wyborem pozy (przy przeszukiwaniu) jak i przy *rescoringu* póź wybranych przez inne funkcje oceniające. Mimo iż PocketScore był optymalizowany na wybór prawidłowej geometrii wiązania bez uwzględniania powinowactwa, testy funkcji wykazały również poprawę skuteczności przesiewania.

Nowszym trendem są próby opracowywania **TSSF opartych na uczeniu maszynowym** (ML-TSSF, ang. *machine-learned target specific scoring functions*), gdzie nie zadaje się predefiniowanej postaci modelu.

Yan i wsp. przedstawili metodę PLEIC-SVM (**Protein-Ligand Empirical Interaction Components Support Vector Machine**) [269]. W podejściu tym, geometria kompleksu (wygenerowana przez dokowanie lub ze struktury eksperymentalnej) służy do stworzenia macierzy oddziaływań ligandów z poszczególnymi resztami miejsca wiążącego. Dla każdego oddziaływania obliczone są fizycznie zrozumiałe wkłady do energii opisujące oddziaływania van der Waalsa, kontakty hydrofobowe, czy wiązania wodorowe. Macierz jest daną wejściową do stworzenia modelu klasyfikatora aktywny/niekatywny za pomocą metody maszyny wektorów nośnych. Autorzy przedstawili trenowanie klasyfikatora dla 36 celów molekularnych na danych z bazy DUD-E (prawdziwe związki aktywne vs wygenerowane molekuly-wabiki). Modele PLEIC-SVM wykazywały lepszą zdolność przesiewania niż funkcja GlideScore na zbiorach testowych.

Fourches i wsp. opracowali dla sztucznego białka wiążącego digoksygeninę specjalny klasyfikator (oparty na modelu wytrenowanym metodą lasu losowego) do odróżniania póź natywnych od fałszywych [270]. Klasyfikator ten został sprzężony z funkcjami oceniającymi programu Glide, dając poprawę zdolności rankingowania ligandów testowych wiążących się z białkiem.

Berishvili i wsp. opublikowali modele klasyfikacyjne (do *rescoringu*) specyficzne dla kinazy PI3K i tankirazy [271]. Modele zostały uzyskane dzięki kilku różnym technikom uczenia maszynowego. Deskryptorami wejściowymi były deskryptory z funkcji programu Smina [272]. Do treningu i testowania modeli wykorzystano zestawy związków aktywne z bazy ChEMBL [257,258] (około 2000/1500 związków dla odpowiednio zbioru treningowego i testowego dla PI3K, oraz około 150/90 związków dla odpowiednio zbioru treningowego i testowego dla tankirazy), oraz wabiki wygenerowane według

metodologii DUD-E [255]. Na zbiorach testowych, niektóre z modeli wykazywały znaczną poprawę mocy przesiewania w porównaniu do funkcji AutoDock Vina i Vinardo [218].

Guedes i wsp. przygotowali zestaw nowych empirycznych funkcji DockTScore, w tym funkcje ogólne i specyficzne dla proteaz lub oddziaływań białko-białko. Funkcje zawierają deskryptory (na bazie pola siłowego MMFF94S) odpowiadające ugruntowanym fizycznym opisom oddziaływań elektrostatycznych, van der Waalsa, kontaktów lipofilowych, polarnych i niepolarnych wkładów energetycznych związanych z solwatacją oraz wkładu związanego z entropią konformacyjną liganda. Autorzy przygotowali funkcje liniowe (za pomocą regresji) oraz nieliniowe (z użyciem metod lasów losowych albo maszyny wektorów nośnych). Funkcje były trenowane na zbiorze PDBbind v2013 [273], lub w przypadku TSSF na wybranych strukturach z tego zbioru. Walidacja na zbiorach z bazy DUD-E [255] wykazała dobrą moc przesiewową nowych funkcji.

Duże badanie nad skutecznością funkcji trenowanych maszynowo w przesiewie wirtualnym przedstawili Wójcikowski i wsp. [274]. Pośród rozpatrywania innych kwestii, autorzy wytrenowali również (metodą lasów losowych) ML-TSSF (do *rescoringu*) dla 102 celów molekularnych z wykorzystaniem danych z bazy DUD-E [255]). Walidacja pokazała, że w większości przypadków ML-TSSF dawały lepszą skuteczność przesiewania, ale poprawa względem funkcji uczonych maszynowo trenowanych na szerokich zbiorach nie była zbyt duża. Dostrojone funkcje były jednak wyraźnie lepsze w przypadku celów molekularnych, które miały mało związków aktywnych w zbiorze.

Obiecujące wyniki dla TSSF opartych na sieciach neuronowych (do *rescoringu*) otrzymali ostatnio Pellicani i wsp. [275], którzy opracowali funkcje specyficzne dla 6 celów molekularnych.

Wang i wsp. zaproponowali funkcję DeepScore (do *rescoringu*) [276], która opiera się na potencjale PMF i uczeniu głębokim (sieć neuronowa). Autorzy przygotowali warianty funkcji dla 102 celów molekularnych ze zbiorów bazy DUD-E [255]. W testach walidacyjnych DeepScore pokazywał poprawę zdolności przesiewania w porównaniu do klasycznych funkcji i innych TSSF, w tym ML-TSSF.

Podobnie jak w przypadku ogólnych funkcji oceniających opartych na uczeniu maszynowym pojawiają się doniesienia kwestionujące skuteczność ML-TSSF. Shen i wsp. testowali zdolność przesiewu dla trzech TSSF opartych na uczeniu maszynowym [277]. Stwierdzili, że ML-TSSF były jedynie marginalnie lepsze niż ogólna, klasyczna funkcja oceniająca programu Glide, a gorsze w porównaniu do dwuwymiarowych modeli QSAR (nie związanych z dokowaniem).

W szerokim ujęciu, do TSSF można zaliczyć także funkcje specjalnie dopasowane do typu dokowanych ligandów. W literaturze znajdziemy znacznie mniej takich prób niż w przypadku TSSF dostrajanych do poszczególnych białek lub rodzin białek (ewentualnie innych celów molekularnych). Warto wymienić tu TSSF przeznaczone do dokowania węglowodanów [278–280] lub dużych giętkich ligandów (peptydów), w tym podejście oparte na kombinacji *scoringu* AutoDock z metodami 2D QSAR [281] czy funkcję *rescorującą* Peptogrid [282].

Badania własne

III. Cel pracy

Celem badań, prowadzonych w ramach niniejszej pracy doktoranckiej, było:

1. zaplanowanie i otrzymanie nowych bibliotek pochodnych fenyloalaniny i tryptofanu, potencjalnie wiążących się do receptora NK1, oraz
2. opracowanie nowych funkcji oceniających specyficznych dla receptora NK1.

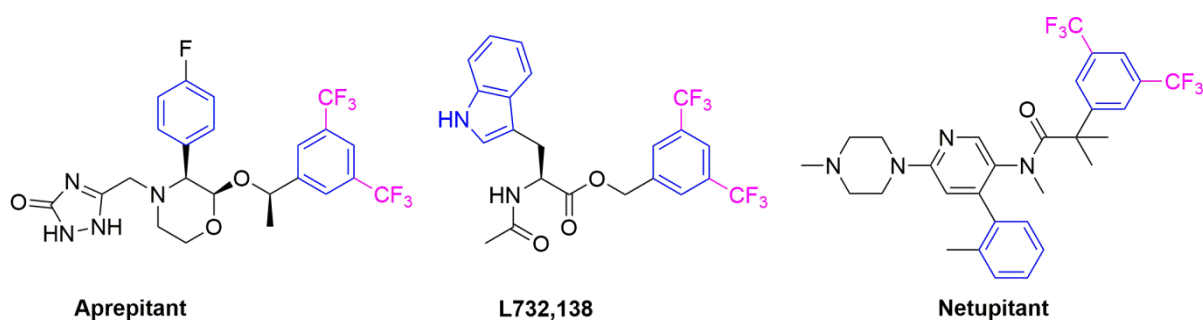
Nowe biblioteki ligandów NK1R miały posłużyć dalszym badaniom zależności struktura–aktywność w kontekście receptora NK1. W otrzymanych analogach chciano sprawdzić, czy i jak poszczególne modyfikacje w strukturze wpływają na powinowactwo do receptora NK1. Dalej, czy możliwe jest rozbudowanie (na sposób umożliwiających tworzenie koniugatów i związków wielofunkcyjnych) znanego antagonisty L-732,138 bez zmiany powinowactwa. W innej serii związków testowano czy możliwe jest zastąpienie kluczowej dla powinowactwa grupy CF_3 inną grupą (SF_5) i czy ta nowa grupa będzie kompatybilna z receptorem NK1.

Dodatkowo, otrzymane przez współpracowników dane o powinowactwie receptorowym zsyntezowanych przez związków, planowałam wykorzystać do opracowania nowych funkcji oceniających specyficznych dla receptora NK1. W ten sposób chciałam uzyskać narzędzie *in silico* wspomagające odkrywanie ligandów NK1R.

IV. Wyniki

IV.1. Projektowanie związków

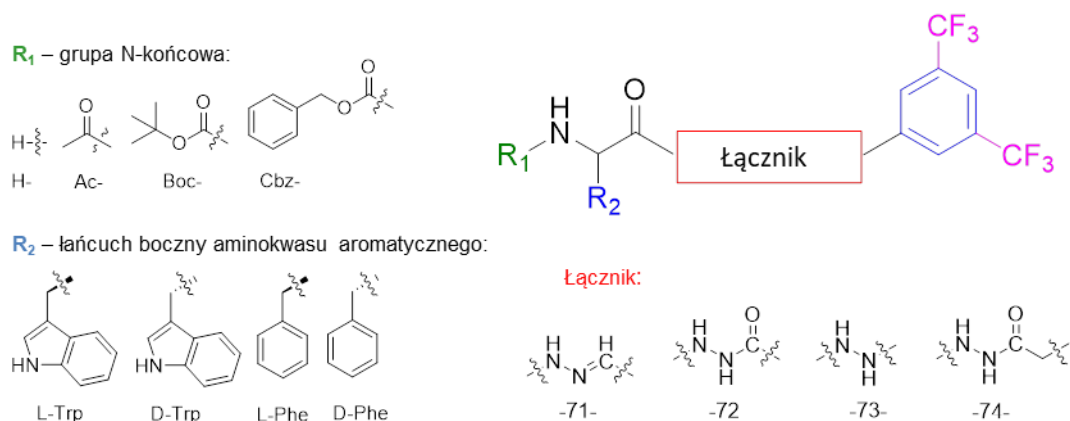
Na podstawie obszernych badań SAR prowadzonych nad antagonistami NK1R oraz wcześniejszych badań syntetycznych i teoretycznych przeprowadzonych w Zakładzie Neuropeptydów Instytutu Medycyny Doświadczalnej i Klinicznej Polskiej Akademii Nauk (ZN IMDiK PAN), wiadomo, że klasyczny małowymiarowy antagonist NK1R (Rysunek 16) składa się z dwóch ugrupowań aromatycznych (na rysunku zaznaczonych na niebiesko), oraz akceptora wiązania wodorowego pomiędzy nimi. W wielu antagonistach o wysokim powinowactwie, jeden z pierścieni aromatycznych jest podstawiony grupami CF₃ w pozycjach 3 i 5 (na rysunku zaznaczone na fioletowo). Kluczowa dla wiązania jest również odległość (zazwyczaj 4–5 atomów) oraz wzajemne położenie między ugrupowaniami aromatycznymi.



Rysunek 16. Struktury chemiczne niektórych znanych antagonistów NK1R.

W początkowym etapie projektu w ramach tworzenia bibliotek pochodnych fenyloalaniny (F) i tryptofanu (W) potencjalnie wiążących się do receptora NK1 zaprojektowano **osiem serii związków** (oznaczonych **W/F-71** do **W/F-74**, Rysunek 17). Pierwotna biblioteka obejmowała **64 nowe związki (1–64)**, w których zróżnicowanie strukturalne pochodziło od kombinatorycznego łączenia:

- 4 możliwych łączników pomiędzy grupami aromatycznymi (Łącznik),
- 4 możliwych podstawników N-końcowych (grupa R₁),
- 4 możliwych ugrupowań aromatycznych (grupa R₂).



Rysunek 17. Wzór ogólny związków serii 71-74.

Opracowanie serii tych związków miało na celu sprawdzenie czy, i jaki wpływ na powinowactwo ma rozdzielanie kluczowych elementów strukturalnych (ugrupowań aromatycznych) poprzez użycie łączników o różnej długości (od 2 do 4 atomów) oraz o różnym charakterze (o różnym stopniu swobody konformacyjnej):

- -NH-N=CH- (serie 71, *N*-acylohydrazony)
- -NH-NH-C(=O)- (serie 72, *N,N'*-diacylohydrazydy)
- -NH-NH- (serie 73, *N*-acylo-*N'*-fenylohydrazydy)
- =NH-NH-C(=O)-CH₃- (serie 74, *N*-acylo-*N'*-fenyloacetylohydrazydy).

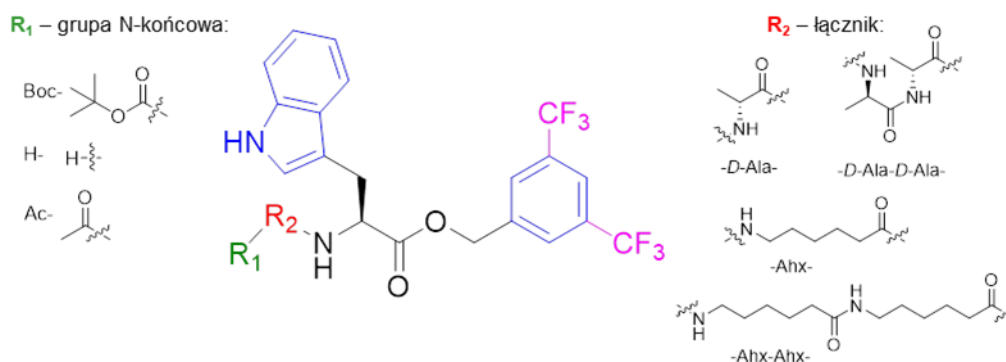
Modyfikacje dotyczyły także łańcucha bocznego aminokwasu aromatycznego R₂ (zastosowanie L- i D-Trp oraz L- i D-Phe), co miało pozwolić na przetestowanie wpływu stereochemii na powinowactwo. Zakres zmienności obejmował także różne podstawienia grupy N-końcowej R₁ (brak podstawienia (H), grupa acetylowa (Ac), grupa *tert*-butyloksykarbonylowa (Boc) oraz grupa benzyloksykarbonylowa (Cbz)). Modyfikacje te zmieniały możliwość tworzenia wiązań wodorowych, wielkość liganda oraz jego lipofilowość. Grupa Cbz miała służyć do testowania ewentualnych oddziaływań aromatycznych. Zmiany powinowactwa między analogami z R₁ = H i R₁ = Ac miały też dać wskazówkę, czy fragmenty oparte o łączniki 71–74 mogą być rozbudowywane/dołączane jako fragmenty koniugatów (przez acylowanie). Struktury wszystkich związków **1–64** przedstawione są w Tabeli 4.

Tabela 4. Struktury związków 1–64.

Lp.	Symbol związku	R ₁	R ₂	Łącznik	Lp.	Symbol związku	R ₁	R ₂	Łącznik
Seria W-71					Seria W-73				
1	Boc-W-71	Boc-	-Trp	-NH-N=CH-	33	Boc-W-73	Boc-	-Trp	-NH-NH-
2	Boc-dW-71	Boc-	-D-Trp		34	Boc-dW-73	Boc-	-D-Trp	
3	Z-W-71	Cbz-	-Trp		35	Z-W-73	Cbz-	-Trp	
4	Z-dW-71	Cbz-	-D-Trp		36	Z-dW-73	Cbz-	-D-Trp	
5	H-W-71	H-	-Trp		37	H-W-73	H-	-Trp	
6	H-dW-71	H-	-D-Trp		38	H-dW-73	H-	-D-Trp	
7	Ac-W-71	Ac-	-Trp		39	Ac-W-73	Ac-	-Trp	
8	Ac-dW-71	Ac-	-D-Trp		40	Ac-dW-73	Ac-	-D-Trp	
Seria F-71					Seria F-73				
9	Boc-F-71	Boc-	-Phe	-NH-N=CH-	41	Boc-F-73	Boc-	-Phe	-NH-NH-
10	Boc-dF-71	Boc-	-D-Phe		42	Boc-dF-73	Boc-	-D-Phe	
11	Z-F-71	Cbz-	-Phe		43	Z-F-73	Cbz-	-Phe	
12	Z-dF-71	Cbz-	-D-Phe		44	Z-dF-73	Cbz-	-D-Phe	
13	H-F-71	H-	-Phe		45	H-F-73	H-	-Phe	
14	H-dF-71	H-	-D-Phe		46	H-dF-73	H-	-D-Phe	
15	Ac-F-71	Ac-	-Phe		47	Ac-F-73	Ac-	-Phe	
16	Ac-dF-71	Ac-	-D-Phe		48	Ac-dF-73	Ac-	-D-Phe	
Seria W-72					Seria W-74				
17	Boc-W-72	Boc-	-Trp	-NH-NH-C(=O)-	49	Boc-W-74	Boc-	-Trp	-NH-NH-C(=O)-CH ₂ -
18	Boc-dW-72	Boc-	-D-Trp		50	Boc-dW-74	Boc-	-D-Trp	
19	Z-W-72	Cbz-	-Trp		51	Z-W-74	Cbz-	-Trp	
20	Z-dW-72	Cbz-	-D-Trp		52	Z-dW-74	Cbz-	-D-Trp	
21	H-W-72	H-	-Trp		53	H-W-74	H-	-Trp	
22	H-dW-72	H-	-D-Trp		54	H-dW-74	H-	-D-Trp	
23	Ac-W-72	Ac-	-Trp		55	Ac-W-74	Ac-	-Trp	
24	Ac-dW-72	Ac-	-D-Trp		56	Ac-dW-74	Ac-	-D-Trp	
Seria F-72					Seria F-74				
25	Boc-F-72	Boc-	-Phe	-NH-NH-C(=O)-	57	Boc-F-74	Boc-	-Phe	-NH-NH-C(=O)-CH ₂ -
26	Boc-dF-72	Boc-	-D-Phe		58	Boc-dF-74	Boc-	-D-Phe	
27	Z-F-72	Cbz-	-Phe		59	Z-F-74	Cbz-	-Phe	
28	Z-dF-72	Cbz-	-D-Phe		60	Z-dF-74	Cbz-	-D-Phe	
29	H-F-72	H-	-Phe		61	H-F-74	H-	-Phe	
30	H-dF-72	H-	-D-Phe		62	H-dF-74	H-	-D-Phe	
31	Ac-F-72	Ac-	-Phe		63	Ac-F-74	Ac-	-Phe	
32	Ac-dF-72	Ac-	-D-Phe		64	Ac-dF-74	Ac-	-D-Phe	

Ac – grupa acetylowa, Boc – grupa *tert*-butyloksykarbonylowa, Cbz – grupa benzyloksykarbonylowa.

W kolejnym etapie projektu zdecydowano opracować serię związków opartych na strukturze znanego antagonisty NK1R związku L-732,138 [139,140], który jest estrem 3,5-*bis*(trifluorometylo)benzylowym *N*-acetylo-L-tryptofanu. Nowe analogi planowano użyć jako wektory do radiofarmaceutyków. Ta część projektu miała na celu sprawdzenie, czy rozbudowanie L-732,138 o dodatkowe elementy od N-końca nie spowoduje utraty powinowactwa do NK1. Docelowo, te elementy rozbudowujące miałyby być *spacerami* umożliwiającymi dołączanie różnych „drugich” farmakoforów albo struktur funkcjonalnych, takich jak chelatory, ładunki molekularne (*cargo*) itp. Opracowano **15 analogów** (związki **65–79**), w tym 14 nowych związków (Rysunek 18) o różnych łącznikach R₂: brak łącznika (**seria W-80**), D-Ala (**seria W-81**), D-Ala-D-Ala (**seria W-82**), Ahx (**seria W-83**), Ahx-Ahx (**seria W-84**). Dodatkowo, na N-końcu związki miały posiadać grupę aminową niepodstawioną (R₁ = H), lub podstawioną grupą acetylową (R₁ = Ac) lub grupą *tert*-butyloksykarbonylową (R₂ = Boc). Zmienność w R₁ miała na celu sprawdzenie, czy tolerowane jest blokowanie grupy aminowej (tak jak miałyby to mieć miejsce w potencjalnych koniugatach). Struktury wszystkich związków **65–79** podane są w Tabeli 5.



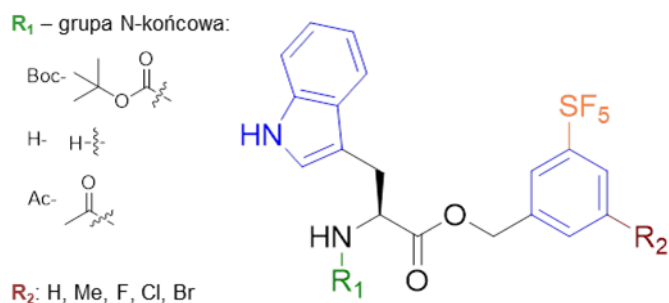
Rysunek 18. Wzór ogólny serii związków W-80–W-84.

Tabela 5. Struktury związków 65–79.

Lp.	Symbol związku	R ₁	R ₂	Lp.	Symbol związku	R ₁	R ₂
Seria W-80				Seria W-83			
65	Boc-W-80	Boc-	-	74	Boc-W-83	Boc-	-Ahx-
66	H-W-80	H-	-	75	H-W-83	H-	-Ahx-
67	Ac-W-80 (L-732,138 [139,140])	Ac-	-	76	Ac-W-83	Ac-	-Ahx-
Seria W-81				Seria W-84			
68	Boc-W-81	Boc-	-D-Ala-	77	Boc-W-84	Boc-	-Ahx-Ahx-
69	H-W-81	H-	-D-Ala-	78	H-W-84	H-	-Ahx-Ahx-
70	Ac-W-81	Ac-	-D-Ala-	79	Ac-W-84	Ac-	-Ahx-Ahx-
Seria W-82							
71	Boc-W-82	Boc-	-D-Ala-D-Ala-				
72	Ac-W-82	H-	-D-Ala-D-Ala-				
73	H-W-82	Ac-	-D-Ala-D-Ala-				

Ac – grupa acetylowa, Ahx – kwas 6-aminoheksanowy, Boc – grupa *tert*-butyloksykarbonylowa, Cbz – grupa benzyloksykarbonylowa.

W dalszych etapach, zaprojektowano kolejne **15 nowych związków (seria W-90, związki 80–94, Rysunek 19)** w oparciu o strukturę L-732,138. W tej serii chciano sprawdzić, czy możliwe jest zastąpienie grupy CF₃ inną grupą, bez utraty powinowactwa. Opracowano związki, w których jedną z grup CF₃ w pierścieniu benzylowym zastąpiono grupą pentafluorosulfanylową, SF₅. Grupa SF₅ nazywana jest grupą „super-CF₃”, ponieważ posiada właściwości podobne do CF₃, ale bardziej natężone (jest bardziej elektroujemna, bardziej elektronowyciągająca, bardziej lipofilowa, większa sterycznie) [283,284]. Ostatnio cieszy się sporym zainteresowaniem w syntezie nowych związków biologicznie czynnych [285].



Rysunek 19. Wzór ogólny serii związków W-90

W serii W-90 badano, jaki wpływ na powinowactwo (przy obecności grupy SF₅) będzie mieć drugi podstawnik w pierścieniu. W tym celu zaplanowano podstawić pierścień (R₂) atomem wodoru, grupą metylową, atomem fluoru, chloru lub bromu. Sprawdzano także, czy analogi z grupą SF₅ muszą być zaacetylowane na N-końcu (R₁ = Ac), aby wykazywać wysokie powinowactwo. Struktury związków **80–94** przedstawiono w Tabeli 6.

Tabela 6. Struktury związków **80–94**.

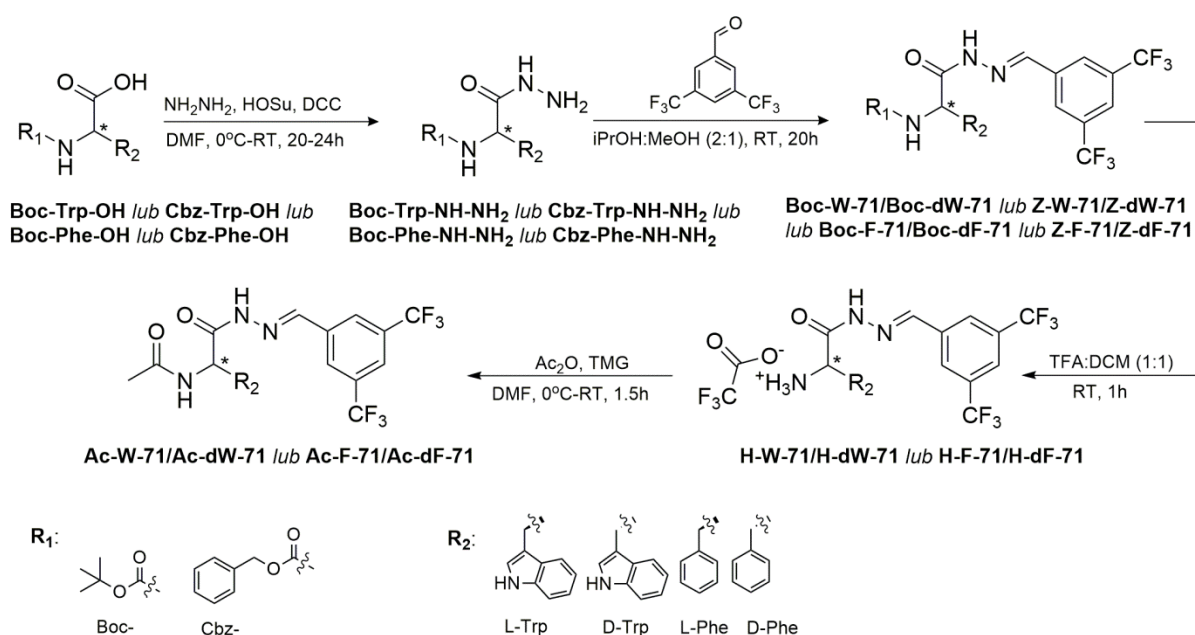
Lp.	Symbol związku	R ₁	R ₂	Lp.	Symbol związku	R ₁	R ₂	Lp.	Symbol związku	R ₁	R ₂
80	Boc-W-90-H	Boc-	-H	85	H-W-90-H	H-	-H	90	Ac-W-90-H	Ac-	-H
81	Boc-W-90-Me	Boc-	-Me	86	H-W-90-Me	H-	-Me	91	Ac-W-90-Me	Ac-	-Me
82	Boc-W-90-F	Boc-	-F	87	H-W-90-F	H-	-F	92	Ac-W-90-F	Ac-	-F
83	Boc-W-90-Cl	Boc-	-Cl	88	H-W-90-Cl	H-	-Cl	93	Ac-W-90-Cl	Ac-	-Cl
84	Boc-W-90-Br	Boc-	-Br	89	H-W-90-Br	H-	-Br	94	Ac-W-90-Br	Ac-	-Br

Ac – grupa acetylowa, Ahx – kwas 6-aminoheksanowy, Boc – grupa *tert*-butyloksykarbonylowa, Cbz – grupa benzyloksykarbonylowa.

IV.2. Syntezy

Zaprojektowane związki otrzymano w ramach kilkusetapowych syntez, wychodząc z tanich i komercyjnie dostępnych *N*^α-Boc lub *N*^α-Cbz pochodnych fenyloalaniny i tryptofanu. W zależności od struktury związku prowadzono reakcje: a) sprzęgania grupy karboksylowej i aminowej (*N,N'*-dicykloheksylokarbodiimid [DCC] jako czynnik sprzęgający z dodatkiem *N*-hydroksysukcynoimidu [HOSu] jako 1,2-dinukleofila), b) otrzymywania hydrazydów, c) otrzymywania hydrazonów, d) usuwania grupy ochronnej Boc przez acydolizę kwasem trifluorooctowym (TFA), e) acetylowania wolnej grupy aminowej za pomocą bezwodnika octowego (Ac₂O), f) *O*-alkilowania bromkami benzyłowymi (tworzenie estrów). Ścieżka syntetyczna wymagała od jednego do siedmiu kroków i była pomyślana w taki sposób, że niektóre związki pośrednie były jednocześnie związkami docelowymi i półproduktami dla innych związków docelowych. Wspólnym elementem dla syntezy wszystkich serii była sekwencja sprzęganie do Boc-pochodnych → acydoliza do soli amoniowych → acetylowanie do *N*-acetylo pochodnych. Reakcje prowadzono w skali milimolowej.

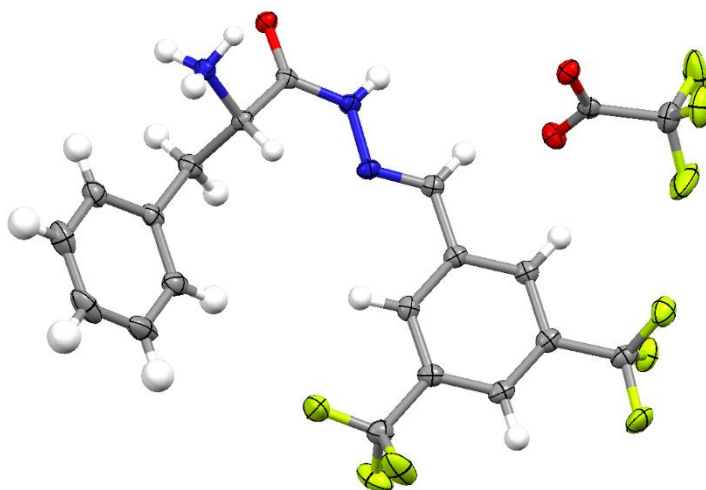
Związki z serii 71 syntezowano zgodnie ze Schematem 1. W pierwszym kroku, przekształcono N^α -Boc lub N^α -Cbz pochodne aminokwasów w hydrazydy przez reakcje z monohydratem hydrazyny (wydajności 41–96%). Hydrazydy były następnie kondensowane z 3,5-bis(trifluorometylo)benzaldehydem, przez co otrzymano hydrazony Boc-X-71 lub Z-X-71 (wydajności 50–94%). Związki posiadające grupę N^α -Boc odbezpieczono za pomocą kwasu trifluorooctowego uzyskując trifluorooctanowe sole amoniowe hydrazonów H-X-71 (wydajności 32–94%). Produkty te poddano acetylacji Ac_2O otrzymując N -acetylowane hydrazony Ac-X-71 z wydajnościami 51–95%.



Schemat 1. Schemat otrzymywania związków z serii 71.

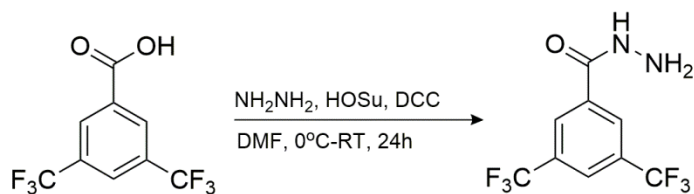
Ac_2O – bezwodnik octowy, DCC – N,N' -dicykloheksylokarbodiimid, DCM – dichlorometan, DMF – N,N -dimetyloformamid, HOSu – N -hydroksysukcynoimid, iPrOH – izopropanol, MeOH – metanol, TFA – kwas trifluorooctowy, TMG – 1,1,3,3-tetrametyloguanidyna.

Związek H-F-71 wykrystalizował z frakcji zebranych w trakcie chromatografii preparatywnej (roztwór ACN:H₂O), dając kryształy odpowiednie do rentgenowskiej analizy strukturalnej. Strukturę związku (Rysunek 20) rozwiązał dr Krzysztof Łyczko z Instytutu Chemii i Techniki Jądrowej w Warszawie (ICHTJ). Struktura została zdeponowana w bazie Cambridge CSD (kod depozycji: CCDC2096461). Szczegółowe badania strukturalne i spektroskopowe (widma VCD w ciele stałym) H-F-71 i H-dF-71 były przedmiotem publikacji [4].

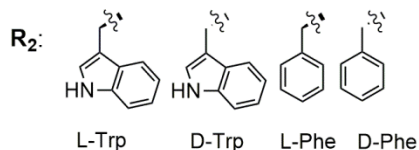
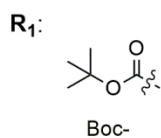
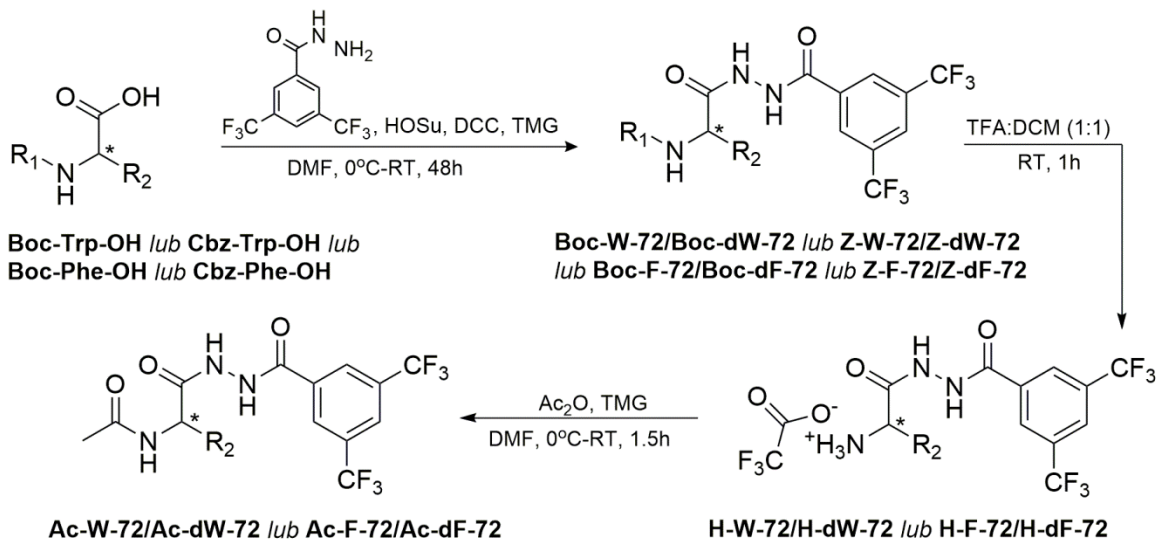


Rysunek 20. Struktura krystaliczna związku H-F-71 (CCDC2096461).

Zaprojektowane związki z serii 72 zsyntezowano zgodnie ze Schematem 2. Syntezę rozpoczęto od przekształcenia kwasu 3,5-*bis*(trifluorometylo)benzoesowego w 3,5-*bis*(trifluorometylo)benzohydrazyd z użyciem monohydratu hydrazyny (wydajność 67%). Następnie związek ten sprzęgano z odpowiednią N^α -Boc lub N^α -Cbz pochodną aminokwasu, otrzymując Boc-X-72 lub Z-X-72 (wydajności 70–90%). Analogi H-X-72 zostały uzyskane przez acydolizę (wydajności 37–69%). W ostatnim kroku, ich acetylowanie z Ac_2O dało pochodne Ac-X-72 (wydajności 68–89%).



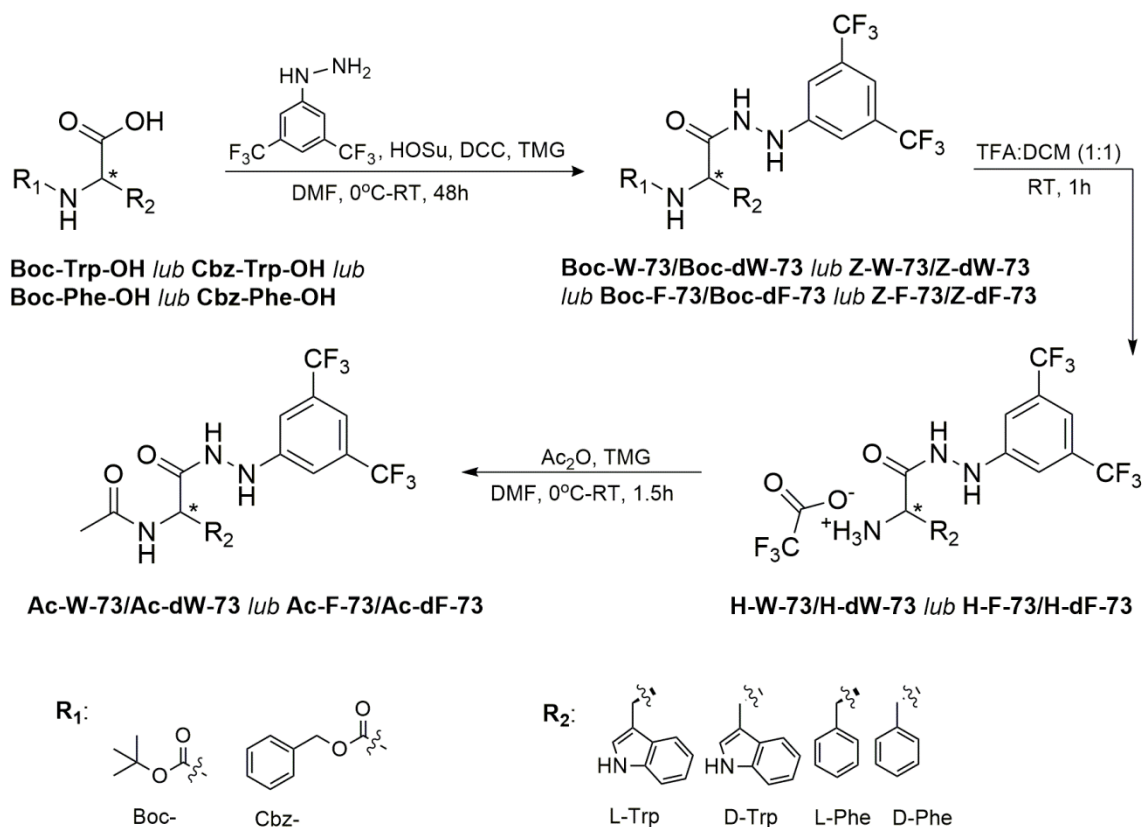
kwask 3,5-bis(trifluorometylo)benzoesowy 3,5-bis(trifluorometylo)benzohydrazyd



Schemat 2. Schemat otrzymywania związków z serii 72.

Ac₂O – bezwodnik octowy, DCC – *N,N'*-dicykloheksylokarbodiimid, DCM – dichlorometan, DMF – *N,N*-dimetyloformamid, HOSu – *N*-hydroksysukcynoimid, TFA – kwas trifluoroctowy, TMG – 1,1,3,3-tetrametyloguanidyna.

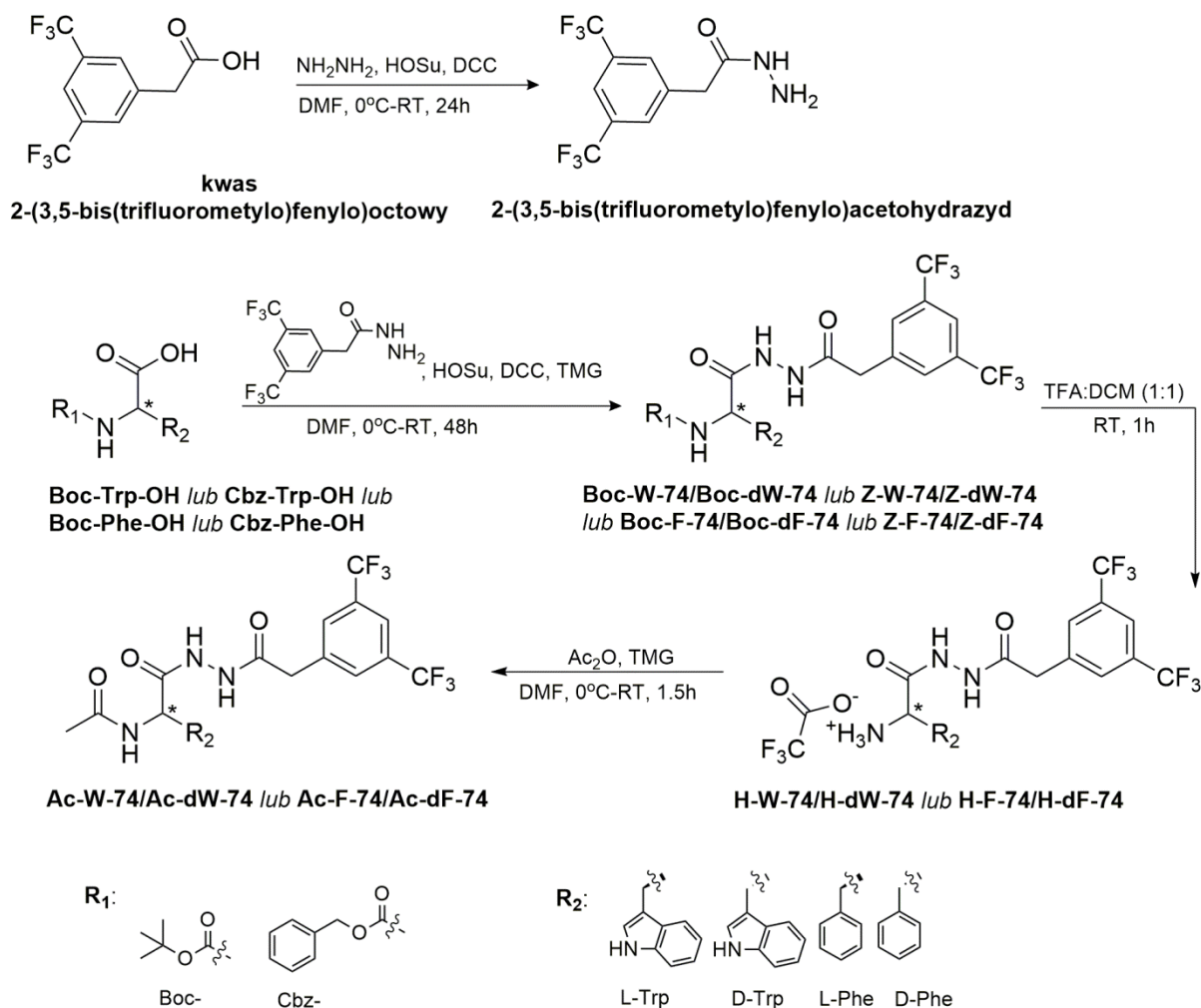
Syntezę serii 73 (Schemat 3) rozpoczęto od sprzęgania 3,5-*bis*(trifluorometylo)fenylohydrazyny z odpowiednimi *N*^α-Boc lub *N*^α-Cbz pochodnymi aminokwasów, otrzymując związki Boc-X-73 i Z-X-73. Analogicznie do poprzednich serii, *N*^α-Boc zabezpieczone analogi poddano reakcji z TFA i otrzymano trifluoroctanowe sole amoniowe (H-X-73, wydajności 77–80%), które następnie zacetylowano do pochodnych Ac-X-73 (wydajności 40–87%).



Schemat 3. Schemat otrzymywania związków z serii 73.

Ac₂O – bezwodnik octowy, DCC – *N,N'*-dicykloheksylokarbodiimid, DCM – dichlorometan, DMF – *N,N*-dimetyloformamid, HOSu – *N*-hydroksysukcynoimid, TFA – kwas trifluoroctowy, TMG – 1,1,3,3-tetrametyloguanidyna.

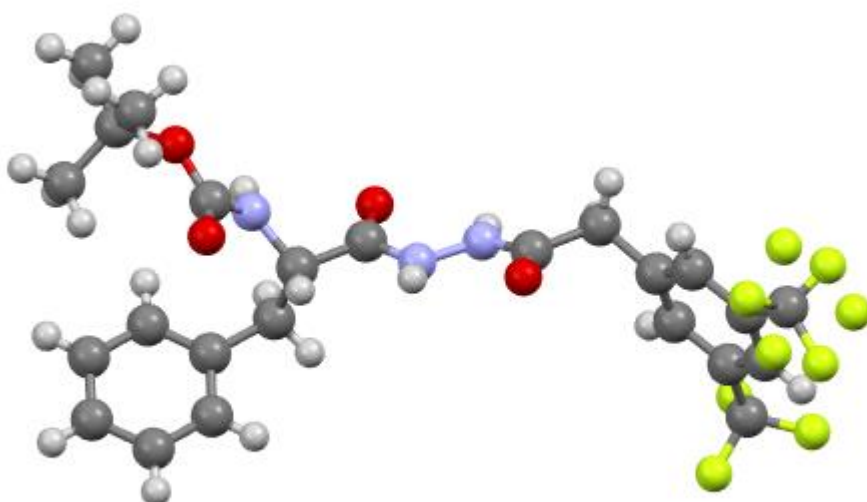
Zaprojektowane związki z serii 74 zsyntezowano zgodnie ze Schematem 4. Syntezę rozpoczęto od przekształcenia kwasu 2-(3,5-*bis*(trifluorometylo)fenylo)octowego w 2-(3,5-*bis*(trifluorometylo)fenylo)acetohydrazyd z użyciem monohydratu hydrazyny (wydajność 83%). Następnie związek ten sprzęgano z odpowiednimi *N*^α-Boc lub *N*^α-Cbz zabezpieczonymi aminokwasami, otrzymując pochodne Boc-X-74 i Z-X-74 (wydajności 30–78%). Acydoliza Boc-pochodnych do trifluoroctanowych soli amoniowych H-X-74 przebiegała z wydajnościami 80–83%. W ostatnim kroku, acetylowanie Ac₂O dało analogi Ac-X-74 (wydajność 40–87%).



Schemat 4. Schemat otrzymywania związków z serii 74.

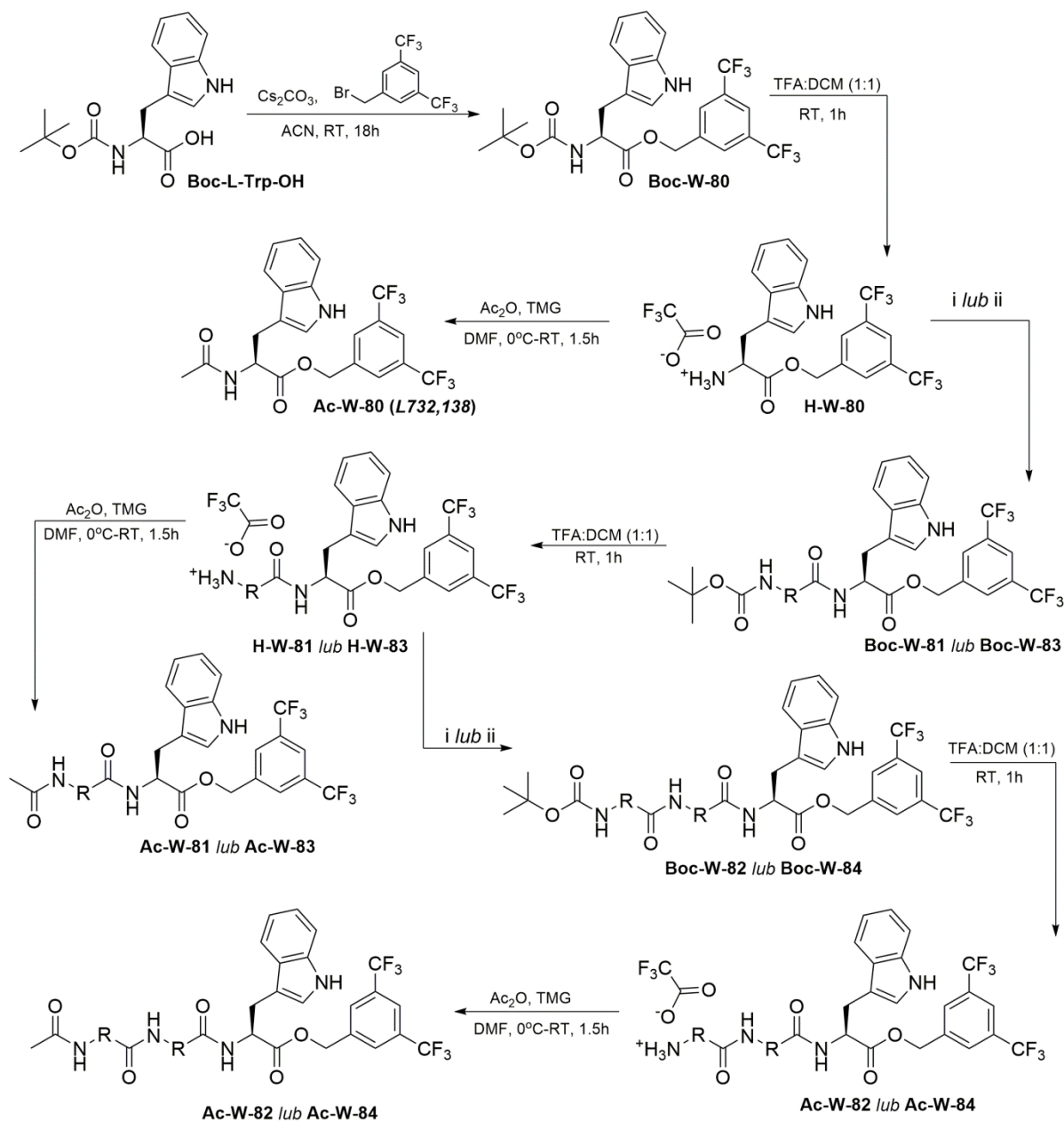
Ac₂O – bezwodnik octowy, DCC – *N,N'*-dicykloheksylokarbodiimid, DCM- dichlorometan, DMF – *N,N*-dimetyloformamid, HOSu – *N*-hydroksysukcynoimid, TFA – kwas trifluorooctowy, TMG – 1,1,3,3-tetrametyloguanidyna.

Związek Boc-F-74 wykrystalizował z roztworu zebranego w trakcie chromatografii preparatywnej (ACN:H₂O), dając kryształy nadające się do rentgenowskiej analizy strukturalnej. Strukturę związku (Rysunek 21) rozwiązał dr Krzysztof Łyczko (IChTJ).

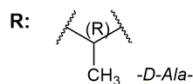


Rysunek 21. Struktura krystaliczna związku Boc-F-74.

Syntezę związków z serii W-80 do W-84 wykonano zgodnie ze Schematem 5. Syntezę rozpoczęto od reakcji *O*-alkilowania Boc-L-Trp-OH bromkiem 3,5-*bis*(trifluorometylo)benzylowym, otrzymując Boc-W-80 (wydajność 80%), który poddano reakcji z TFA i uzyskano H-W-80 (wydajność 81%). W kolejnych etapach związek ten używano w reakcji acetylowania bezwodnikiem octowym (z wytworzeniem Ac-W-80, L-732,138 [139,140], wydajność 98%) lub do dalszych syntez (przyłączenia łączników). H-W-80 sprzęgano z Boc-D-Ala-OH (droga i) lub z preformowanym estrem *N*-hydroksysukcynimidowym kwasu Boc-6-aminoheksanowego (droga ii), otrzymując odpowiednio Boc-W-81 (wydajność 38%) lub Boc-W-83 (wydajność 57%). Kolejne analogi otrzymano w wyniku tej samej sekwencji reakcji: odbezpieczanie (wydajności 32–65%), acetylowanie (wydajności 61–68%) lub sprzęganie (wydajność H-W-81 → Boc-W-82 38%; H-W-83 → Boc-W-84 55%). Wydajność sprzęgania z estrem preformowanym była wyższa.

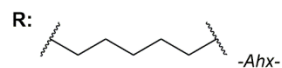


dla: Boc-W-81, H-W-81, Ac-W-81, Boc-W-82, H-W-82, Ac-W-82



i: Boc-D-Ala-OH, HOSu, DCC, TMG, DMF, 0°C-RT, 48h

dla: Boc-W-83, H-W-83, Ac-W-83, Boc-W-84, H-W-84, Ac-W-84

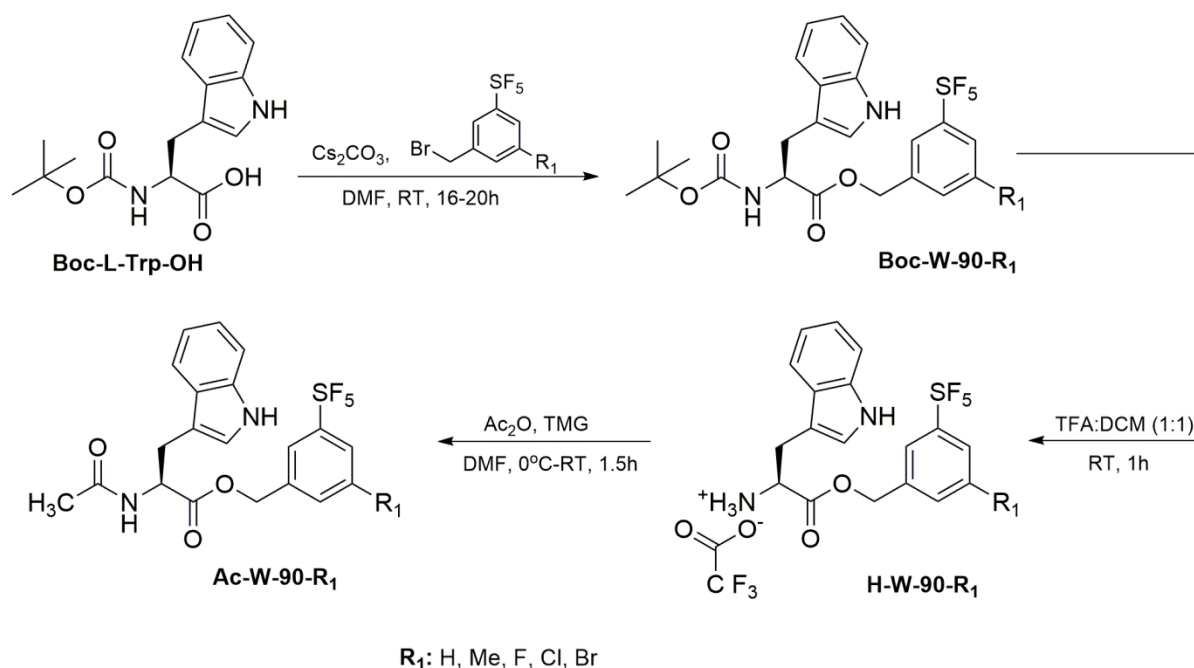


ii: Boc-6-Ahx-OSu, TMG, DMF, RT, 6h

Schemat 5. Schemat otrzymywania związków z serii W-80–84.

Ac₂O – bezwodnik octowy, DCC – *N,N'*-dicykloheksylokarbodiimid, DCM- dichlorometan, DMF – *N,N*-dimetyloformamid, HOSu – *N*-hydroksysukcynoimid, TFA – kwas trifluorooctowy, TMG – 1,1,3,3-tetrametyloguanidyna.

Synteze serii W-90 wykonano zgodnie ze Schematem 6. Analogicznie, jak w seriach W-80 do W-84, Boc-L-Trp-OH poddano reakcji *O*-alkilowania z odpowiednimi bromkami benzylovym podstawionym grupą SF₅ i grupą R₁. Wydajność *O*-alkilowania wynosiła 82–93% i nie wydaje się zależna od charakteru elektronowego podstawnika R₁ bromku benzylovego. Otrzymane *N*^α-Boc pochodne tryptofanu odbezpieczono TFA, uzyskując odpowiednie trifluorooctanowe sole amoniowe H-W-90-R₁ z wydajnościami 58–79%. W ostatnim kroku, analogi acetylowane Ac-W-90-R₁ otrzymano przez acetylowanie Ac₂O (wydajności 44–69%).



Schemat 6. Schemat otrzymywania związków z serii W-90.

Ac₂O – bezwodnik octowy, DCM – dichlorometan, DMF – *N,N*-dimetyloformamid, TFA – kwas trifluorooctowy, TMG – 1,1,3,3-tetrametyloguanidyna.

IV.3. Aktywność biologiczna

IV.3.1. Powinowactwo do receptora NK1R

Zsyntezowane związki zostały poddane ewaluacji biologicznej. Oznaczenia powinowactwa receptorowego wobec receptora NK1 wykonała dr Joanna Matalińska z Zakładu Neuropeptydów IMDiK PAN. Wyniki przedstawione są w Tabeli 7. Część omawianych danych biologicznych przedstawiono w pracach [1] i [2].

Tabela 7. Powinowactwo zeszyntezowanych związków do receptora NK1 (IC_{50}).

Lp.	Związek	IC_{50} [nM]	Lp.	Związek	IC_{50} [nM]	Lp.	Związek	IC_{50} [nM]	Lp.	Związek	IC_{50} [nM]
Seria W-71			Seria F-72			Seria W-74			Seria W-82		
1	Boc-W-71	34100	25	Boc-F-72	>> 30 μ M	49	Boc-W-74	> 100 μ M	71	Boc-W-82	14 [1]
2	Boc-dW-71	> 30 μ M	26	Boc-dF-72	>> 30 μ M	50	Boc-dW-74	> 100 μ M	72	H-W-82	81 [1]
3	Z-W-71	> 100 μ M	27	Z-F-72	>> 30 μ M	51	Z-W-74	> 100 μ M	73	Ac-W-82	14 [1]
4	Z-dW-71	> 100 μ M	28	Z-dF-72	>> 30 μ M	52	Z-dW-74	> 100 μ M	Seria W-83		
5	H-W-71	10100	29	H-F-72	>> 30 μ M	53	H-W-74	> 30 μ M	74	Boc-W-83	61 [1]
6	H-dW-71	9600	30	H-dF-72	>> 30 μ M	54	H-dW-74	77300	75	H-W-83	35 [1]
7	Ac-W-71	39900	31	Ac-F-72	>> 30 μ M	55	Ac-W-74	> 100 μ M	76	Ac-W-83	83 [1]
8	Ac-dW-71	>100 μ M	32	Ac-dF-72	>> 30 μ M	56	Ac-dW-74	> 30 μ M	Seria W-84		
Seria F-71			Seria W-73			Seria F-74			77	Boc-W-84	5 [1]
9	Boc-F-71	> 100 μ M	33	Boc-W-73	50993	57	Boc-F-74	> 30 μ M	78	H-W-84	20 [1]
10	Boc-dF-71	> 100 μ M	34	Boc-dW-73	17500	58	Boc-dF-74	> 30 μ M	79	Ac-W-84	14 [1]
11	Z-F-71	>> 30 μ M	35	Z-W-73	> 30 μ M	59	Z-F-74	>> 30 μ M	Seria W-90		
12	Z-dF-71	>> 30 μ M	36	Z-dW-73	>> 30 μ M	60	Z-dF-74	>> 30 μ M	85	H-W-90-H	14730 [2]
13	H-F-71	9400	37	H-W-73	4030	61	H-F-74	3900	86	H-W-90-Me	1020 [2]
14	H-dF-71	26600	38	H-dW-73	5323	62	H-dF-74	> 30 μ M	87	H-W-90-F	5023 [2]
15	Ac-F-71	>> 30 μ M	39	Ac-W-73	14095	63	Ac-F-74	> 30 μ M	88	H-W-90-Cl	1797 [2]
16	Ac-dF-71	> 30 μ M	40	Ac-dW-73	41300	64	Ac-dF-74	> 30 μ M	89	H-W-90-Br	1691 [2]
Seria W-72			Seria F-73			Seria W-80			90	Ac-W-90-H	3022 [2]
17	Boc-W-72	> 30 μ M	41	Boc-F-73	> 30 μ M	65	Boc-W-80	180 [1]	91	Ac-W-90-Me	85 [2]
18	Boc-dW-72	> 30 μ M	42	Boc-dF-73	> 100 μ M	66	H-W-80	359 [1]	92	Ac-W-90-F	383 [2]
19	Z-W-72	> 100 μ M	43	Z-F-73	> 30 μ M	67	Ac-W-80 (L-732, 138 [139,140])	26 [1]	93	Ac-W-90-Cl	42 [2]
20	Z-dW-72	> 30 μ M	44	Z-dF-73	> 100 μ M	Seria W-81			94	Ac-W-90-Br	34 [2]
21	H-W-72	> 100 μ M	45	H-F-73	5879	68	Boc-W-81	135 [1]	Związki odniesienia		
22	H-dW-72	> 100 μ M	46	H-dF-73	13346	69	H-W-81	25 [1]		Aprepitant	28
23	Ac-W-72	>> 30 μ M	47	Ac-F-73	> 30 μ M	70	Ac-W-81	25 [1]		Substancja P	31
24	Ac-dW-72	>> 30 μ M	48	Ac-dF-73	> 30 μ M						

IC_{50} – stężenie, przy którym ligand redukuje wiązanie radiolianda o 50%. Część wyników przedstawiono uprzednio w publikacjach [1] i [2]. Oznaczenia powinowactwa receptorowego wykonała dr Joanna Matalińska z ZN IMDiK PAN.

IV.3.1.1. Serie W/F-70–W/F-74

Analiza danych o powinowactwie dla serii W/F-71–W/F-74 prowadzi do następujących spostrzeżeń:

- Większość związków z tych serii nie ma znaczącego powinowactwa do receptora NK1, żaden związek nie wykazuje powinowactwa na poziomie submikromolowym.
- Spośród wszystkich serii wyróżnić można siedem związków (H-F-74, H-W-73, H-dW-73, H-F-73, H-F-71, H-dW-71, H-W-71), które można uznać za „najlepsze” (IC_{50} między 4 a 10 μM). Wszystkie te związki posiadają niepodstawioną grupę aminową. Analogi *N*-podstawione, zarówno Ac, Boc i Z nie wykazują żadnej mierzalnej aktywności albo bardzo słabą.
- W mierzonym zakresie stężeń, *N,N'*-diacylohydrazydy H-F-72, H-dF-72, H-W-72, H-dW-72 (czyli odpowiedniki „aktywnych” *N*-acylohydrazonów z serii 71) nie wykazywały nawet szczątkowego powinowactwa ($\gg 30 \mu M$ lub $> 100 \mu M$).
- Wydaje się, że związki z krótszymi łącznikami wiążą się z receptorem NK1 przeciętnie lepiej, niż te posiadające dłuższe łączniki. Najlepszą aktywność wykazywały pojedyncze związki z serii W-73 i F-73, a następnie z serii W-71 i F-71. Są to serie z najkrótszymi łącznikami (odpowiednio 2 lub 3 atomy). Z drugiej strony, spośród związków z serii 72, które posiadają równie krótki łącznik jak związki z serii 71, nie znaleziono żadnej aktywnej pochodnej.
- Nie można jednoznacznie stwierdzić, że związki z aminokwasami o konfiguracji L są aktywniejsze od tych z aminokwasami o konfiguracji D. Jednakże analiza wybranych, najbardziej aktywnych związków w seriach sugeruje, że związki posiadające L-aminokwasy są zazwyczaj nieznacznie lepsze.
- Szczególnym przykładem bardzo „ostrych” zależności SAR może być porównanie związków Ac-W-73 ($IC_{50} = 14\ 095 \text{ nM}$) oraz Ac-W-80 (L-732,138, $IC_{50} = 26 \text{ nM}$). Strukturalnie różnią się one tylko dwoma atomami w łączniku (Ac-W-73: -NH-NH-, a Ac-W-80: -O-CH₂-), a jednocześnie powinowactwo Ac-W-73 jest o około 540 razy słabsze niż Ac-W-80.

IV.3.1.2. Serie W-80 – W-84

Następną zsyntezowaną grupę związków stanowiły analogi z serii W-80–W-84, które zostały zaprojektowane jako potencjalne wektory radionuklidów w celowanej terapii radionuklidowej. W tej grupie związków zaobserwowano następujące fakty odnośnie do powinowactwa receptorowego:

- Macierzysty związek serii, Ac-W-80, (w literaturze znany jako L-732,138 [139,140]), charakteryzuje się powinowactwem na poziomie nanomolowym (spójnym z wcześniejszą literaturą [139,140]), porównywalnym do referencyjnego aprepitantu. Analog z wolną grupą aminową wiąże się około 14 razy słabiej, a podstawiony grupą Boc około 7 razy słabiej, co jest spójne z danymi literaturowymi dla analogów 3,5-dimetylo-podstawionych [139,140].
- Rozbudowanie szkieletu W-80 o dodatkową resztę D-Ala (seria W-81) nie powoduje spadku powinowactwa. Pochodne z wolną grupą aminową oraz *N*-acetylowane wykazują podobne powinowactwo do związku macierzystego (Ac-W-80), natomiast związek podstawiony grupą Boc charakteryzuje się około 5-krotnie słabszym powinowactwem.
- Wprowadzenie kolejnej reszty D-Ala (seria W-82) powoduje poprawę powinowactwa ($IC_{50} = 14$ nM) w przypadku związków *N*-acetylowanego i Boc-podstawionego. W tej grupie najslabsze powinowactwo wykazuje analog z wolną grupą aminową (około 5 razy słabsze niż podstawione analogi i około 3 razy słabsze niż jego odpowiednik z jedną resztą D-Ala).
- Dołączenie reszty Ahx do związku macierzystego (seria W-83) skutkuje nieznacznym spadkiem powinowactwa (IC_{50} dla Ac-W-83 jest około 3 razy mniejsze niż dla L-732,138). Jednocześnie analogi z wolną grupą aminową oraz Boc-podstawiony wykazują lepszą aktywność niż ich odpowiedniki bez dodatkowego łącznika. W tej serii związków (z dołączoną jedną resztą Ahx) najlepszym powinowactwem charakteryzuje się pochodna z wolną grupą aminową (H-W-83), która jest około 10 razy bardziej aktywna niż nierozbudowany odpowiednik (H-W-80).
- Rozbudowanie o kolejną resztę Ahx (seria W-84) powoduje wzrost powinowactwa, wszystkie związki z serii W-84 są aktywniejsze niż związek macierzysty. Najaktywniejszym związkiem, we wszystkich seriach, jest Boc-W-84 ($IC_{50} = 5$ nM, około 5 razy wyższe niż dla L-732,138). *N*-acetylowany analog wykazuje nieco

słabsze powinowactwo, zaś pochodna z wolną grupą aminową charakteryzuje się aktywnością zbliżoną do związku macierzystego Ac-W-80.

Otrzymane wyniki powinowactwa wskazują, że możliwe jest rozbudowanie znanego antagonisty L-732,138 o dodatkowe elementy od N-końca bez utraty powinowactwa do NK1R (z wyjątkiem rozbudowania o pojedynczą grupę Ahx, gdzie można zaobserwować nieznaczny spadek aktywności). Poprawę aktywności względem macierzystego związku zaobserwowano dla serii, w których przyłączano podwójny łącznik (W-82 i W-84), zarówno w przypadku pochodnych z wolną grupą aminową i *N*-acetylowanych. Najsilniejszym związkiem spośród wszystkich badanych w tych seriach jest najdłuższy analog (Boc-W-84).

IV.3.1.3. Seria W-90

Kolejno opracowano analogi z serii W-90, które zaprojektowano również w oparciu o strukturę L-732,138. W tych związkach jedną z grup CF_3 w pierścieniu benzyłowym zastąpiono grupą SF_5 , a drugą innym podstawnikiem. Sprawdzano także, czy związki muszą być zaacetylowane na N-końcu, aby wykazywać wysoką aktywność. Analiza danych dotyczących powinowactwa prowadzi do następujących wniosków:

- Kluczowa dla aktywności jest obecność grupy acetylowej na N-końcu. *N*-acetylowane analogi wykazują lepsze powinowactwo niż ich odpowiedniki z wolną grupą aminową (mają nawet około 50 razy lepsze powinowactwo niż analogi z wolną grupą aminową, dla $R_2 = H$ około 5 razy, $R_2 = Me$ około 12 razy, $R_2 = F$ około 13 razy, $R_2 = Cl$ około 43 razy, $R_2 = Br$ około 50 razy). Wynik ten jest spójny z trendami obserwowanymi dla innych estrów benzyłowych tryptofanu [139,140].
- Rodzaj podstawnika (R_2) w pierścieniu aromatycznym (wielkość) wpływa na aktywność związków. W całej serii i pośród *N*-acetylowanych pochodnych najlepszym powinowactwem charakteryzuje się związek z największym (o największej objętości) podstawnikiem R_2 , Ac-W-90-Br ($IC_{50} = 34$ nM, wartość nieznacznie gorsza niż dla referencyjnych ligandów aprepitantu i SP). Dalej, niewielki spadek aktywności można zaobserwować dla związku o nieco mniejszych podstawinkach $R_2 = Cl$, Ac-W-90-Cl ($IC_{50} = 42$ nM) oraz $R_2 = Me$, Ac-W-90-Me ($IC_{50} = 85$ nM). Natomiast, znaczący spadek powinowactwa widoczny jest dla analogu z najmniejszym podstawnikiem $R_2 = H$, Ac-W-90-H, który jest około 100 razy słabszy niż odpowiednik z grupą $R_2 = Br$. Podobne zależności można zaobserwować wśród związków z wolną grupą aminową, w których najsłabsze powinowactwo wykazuje

analog z najmniejszym podstawnikiem $R_2 = H$ ($IC_{50} = 14.73 \mu M$). Najbardziej aktywnym jest natomiast związek o $R_2 = CH_3$, H-W-90-CH₃ ($IC_{50} \sim 1 \mu M$), i kolejno analogi o $R_2 = Br$ oraz Cl. Jednakże żaden ze związków w tej podserii nie wykazuje powinowactwo w zakresie nanomolowym.

Na podstawie analizy otrzymanych wyników powinowactwa można stwierdzić, że możliwe jest zastąpienie w strukturze znanego antagonisty L-732,138 kluczowej dla powinowactwa grupy CF_3 inną grupą SF_5 bez znaczącego spadku powinowactwa. Istotnymi czynnikami dla aktywności jest rodzaj (rozmiar) drugiego podstawnika w pierścieniu aromatycznym oraz obecność *N*-acetylowej grupy na *N*-końcu.

IV.3.2. Oddziaływanie W-90 z NK1R

W celu zrozumienia zależności SAR nowych ligandów z serii H-W-90 i Ac-W-90 na poziomie oddziaływań ligand-receptor, przeprowadzono dokowanie molekularne do struktur receptora NK1 pobranych z bazy PDB (6HLO, 6HLL i 6HLP) [27] za pomocą programu AutoDock Vina. Wizualna inspekcja póż z dokowania [222] oraz uwzględnieniem wiedzy SAR, pozwoliły założyć, że analizowane związki wiążą się z NK1 jak na Rysunku 22. Sposób oddziaływania zostanie szczegółowo omówione na przykładzie związku Ac-W-90-Br (związek **94**) [2].

Według dokowania, pierścień 3-Br-5- SF_5 -fenyłowy skierowany jest do dołu kieszeni wiążącej. Ugrupowanie to tworzy wiele oddziaływań hydrofobowych (van der Waalsa) z Met81, Asn85, Pro112, Ile113, Ile116, Phe264, Met291, Ala294 i Met295. Położenie grupy SF_5 pokrywa się z pozycją jednej z grup CF_3 aprepitantu w strukturze eksperymentalnej 6HLO (Rysunek 22C) [27] oraz w L-732,138 (według modelowania [1]). Z atomami fluoru z grupy SF_5 mogą oddziaływać wodory z ugrupowania amidowego łańcucha bocznego Asn89 poprzez oddziaływania $H \cdots F$. Inną resztą w pobliżu SF_5 jest Trp261. Atom bromu znajduje się blisko Ile113 i Val200, nieco powyżej pozycji jednej z grup CF_3 aprepitantu w 6HLO [27] (Rysunek 22C) i L-732,138 (według modelowania [1]). Tlen karboksylowy ugrupowania estrowego znajduje się w odległości wiązania wodorowego od wodoru hydroksylowego Tyr287, jednakże geometria jest raczej nieoptymalna dla powstawania silnego wiązania wodorowego. Fragment $-CH_2-O-$ liganda sąsiaduje z łańcuchami bocznymi Asn89 i His108. Pierścień indolowy bierze udział w oddziaływaniach $\pi-\pi$ z pierścieniami aromatycznymi Phe268 i His197. Dodatkowo jest otoczony łańcuchami bocznymi Ile182 i Gln165. *N*-acetyłowy fragment liganda skierowany jest w kierunku wyjścia z kieszeni wiążącej i

wystawiony jest na działanie rozpuszczalnika. Ogólne pozycje wiązania przewidywane dla Ac-W-90-Br i wcześniej opisane dla L-732,138 (na podstawie modelowania w ZN IMDIK PAN [1]) są bardzo podobne.

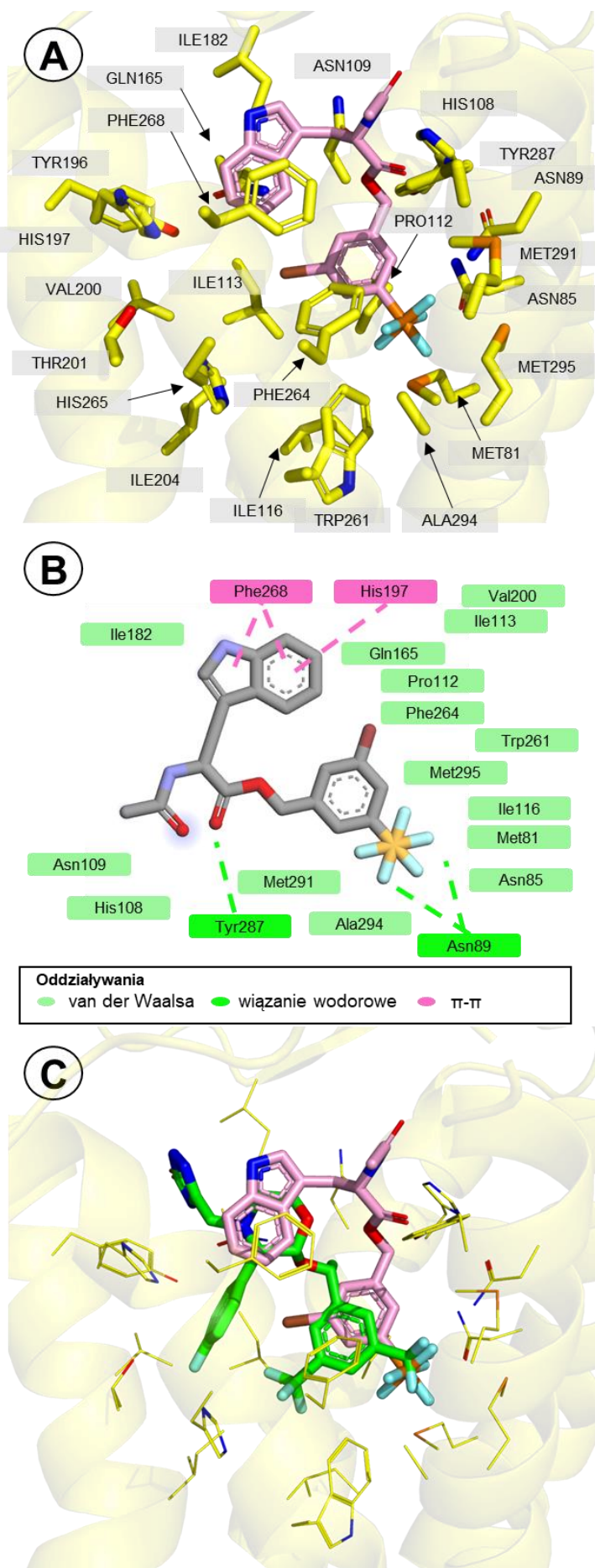
Inne związki z serii Ac-W-90 przyjmują pozy wiążące dość podobne do omawianej dla Ac-W-90-Br [2]. Warto zauważyć, że dokowanie z konformacjami *cis* lub *trans* wiązania amidowego daje bardzo podobne pozy, różniące się jedynie orientacją grupy *N*-acetylowej (która mimo to nadal znajduje się w podobnym miejscu i jest wystawiona na działanie rozpuszczalnika). Analogiczne są pozy uzyskane dla związków z serii H-W-90. Naładowana grupa aminowa jest także wystawiona na działanie rozpuszczalnika.

Przedstawione sposoby oddziaływania wydają się wyjaśniać obserwowane zależności SAR w tej grupie związków. Zgodnie z modelem, podstawniki R_2 znajdują się na dnie hydrofobowej kieszeni wiążącej. Można oczekiwać, że większe podstawniki, takie jak $R_2 = Cl$ lub $R_2 = Br$, zapewnią lepsze wypełnienie tego hydrofobowego obszaru i zwiększą aktywność poprzez oddziaływania dyspersyjne [286,287]. Jest to zgodne z eksperymentem, ponieważ trend powinowactwa jest w przybliżeniu zgodny ze zmianą objętości podstawnika (w serii Ac-W-90 porządek powinowactwa jest następujący: $R_2 = Br > Cl > Me > F > H$, zaś w serii H-W-90 $R_2 = Me \sim Br \sim Cl > F > H$). Hipotetycznie, jeszcze większy (dłuższy) podstawnik hydrofobowy R_2 mógłby zapewnić jeszcze lepsze wypełnienie obszaru hydrofobowego i jeszcze bardziej poprawić powinowactwo wiązania. Dodatkowo, to że Ac-W-90-Cl ($R_2 = Cl$) i Ac-W-90-Br ($R_2 = Br$) mają zdecydowanie wyższe powinowactwo niż Ac-W-90-F ($R_2 = F$) może wskazywać na obecność wiązania halogenowego, jednakże nie jest to obserwowane w wynikach dokowania. Potwierdzenie tej hipotezy mógłby dostarczyć analog z $R_2 = I$, gdyby wykazywał dodatkowo znaczący wzrost powinowactwa.

Zgodnie z moim modelem grupa *N*-acetylowa w serii Ac-W-90 jest skierowana na zewnątrz kieszeni wiążącej. Podobna jest orientacja tego fragmentu w analogu Ac-W-80 w modelu (opracowanym przez dr hab. Piotra Lipińskiego w ZN IMDIK PAN [1]) dla związków z serii W-80–W-84. W świetle tego zrozumiała jest możliwość rozbudowy Ac-W-90 przez dobudowywanie fragmentów do *N*-końca.

Brak specyficznych oddziaływań między białkiem a grupą aminową w analogach H-W-90 może korelować ze zmniejszonym powinowactwem serii H-W-90 w porównaniu z *N*-acetylowanymi analogami Ac-W-90. Hipotetycznie, koszt energetyczny częściowej desolvatacji naładowanej grupy aminowej jest wyższy niż grupy acetylowej, a ponieważ nie tworzą się dodatkowe konkretne oddziaływania z białkiem, aby zrekompensować tę karę,

obserwuje się spadek powinowactwa. Efekty kosztu częściowej desolvatacji w miejscach wystawionych na działanie rozpuszczalnika były systematycznie obserwowane wcześniej [288]. W przypadku moich związków jest to prawdopodobna hipoteza, oparta na wyniku dokowania, eksperymentalnym powinowactwie dla serii H-W-90 oraz wcześniejszych danych SAR dla ligandów NK1R na bazie tryptofanu. W wielu wypadkach dla takich związków obserwowano tolerancję na *N*-acylację z szerokim zakresem podstawników i jednocześnie znaczne spadki powinowactwa w przypadku obecności wolnych grup aminowych [139–141].



Rysunek 22. Wyniki dokowania do NK1R dla związku Ac-W-90-Br.

(A) Sposób oddziaływania związku Ac-W-90-Br (94) z receptorem NK1 w miejscu wiążącym przewidziany na podstawie dokowania molekularnego.

(B) Dwuwymiarowy schemat oddziaływań. Oddziaływania zaznaczone zgodnie z legendą.

(C) Nałożenie sposobów wiązania Ac-W-90-Br (różowy, kompleks z dokowania) i aprepitantu (zielony, struktura eksperymentalna 6HLO [27]). Kompleks jest zorientowany w ten sam sposób jak na rysunku (A).

Ligand zaznaczono na różowo. Heteroatomy mają kolory zgodnie z przyjętą powszechnie konwencją (O – czerwony, N – ciemnoniebieski, F – jasnoniebieski, S – pomarańczowy, Br – brązowy). Receptor (kolor żółty) przedstawiono w sposób uproszczony, uwzględniono jedynie wybrane helisy (wstążki) i łańcuchy boczne (patyczki). Dla przejrzystości ukryto atomy wodoru.

Dla przejrzystości na rysunku (C) wyłączono wyświetlanie etykiet. Łańcuchy boczne receptora przedstawiono jako cienkie linie.

IV.3.3. Cytotoksyczność

Drugim rodzajem badania *in vitro*, w którym testowana była część zsyntezowanych przeze mnie związków było badanie wpływu na żywotność komórek (cytotoksyczności). Testy te dotyczyły tylko analogów z serii W/F-70–W/F-74 i zostały przeprowadzone przed wykonaniem testów powinowactwa receptorowego. Motywacją do wykonania tego badania był fakt, że antagoniści receptorów NK1 mogą mieć działanie przeciwnowotworowe [31,165], a we wczesnych etapach prac w ZN IMDiK PAN nie funkcjonowała sprawna metoda testowania powinowactwa wobec NK1R. Stąd też badanie cytotoksyczności miało dawać pośrednio wskazówkę co do ewentualnego powinowactwa receptorowego.

Badanie wykonano za pomocą testu MTS na komórkach nowotworowych czerniaka linii COLO679 i prawidłowych fibroblastach BJ. Oznaczenia cytotoksyczności wykonała dr Joanna Matalińska z Zakładu Neuropeptydów IMDiK PAN.

Analiza otrzymanych wyników aktywności cytotoksycznej prowadzi do następujących spostrzeżeń:

- Przedstawiane związki cechują się zróżnicowaną aktywnością wobec komórek nowotworowych COLO679 i prawidłowych BJ.
- *N,N'*-diacylohydrazydy z serii W-72 i F-72 nie wykazują znaczącego wpływu na żywotność komórek zarówno prawidłowych jak i nowotworowych ($IC_{50} > 200 \mu M$).
- Nieaktywne lub mało aktywne są również *N*-acylo-*N'*-fenylohydrazydy z serii W-73 i większość z serii F-73 (z wyjątkiem H-dF-73).
- Podobnie brak aktywności obserwuje się w *N*-acylo-*N'*-fenyloacetylohydrazydach serii F-74 i większości W-74 (z wyjątkiem H-W-74 i H-dW-74).
- Nieaktywne są również *N*-acylohydrazony F-71 (z wyjątkiem H-F-71 i H-dF-71).
- W wysokich stężeniach ($IC_{50} \sim 140 \mu M$) związek H-W-74 działa cytotoksycznie na obie linie komórkowej, bez wyraźnej selektywności.
- *N*-acylohydrazony H-F-71 i H-dF-71 wykazują znaczące działanie cytotoksyczne wobec linii komórkowej COLO679 (IC_{50} w zakresie 50–100 μM) i jednocześnie słabsze wobec linii komórek prawidłowych BJ. W tym przykładzie związków można zaobserwować stereoselektywność efektu, ponieważ izomer *S* (związek H-F-71) jest bardziej selektywny niż izomer *R* (H-dW-71).
- Silne działanie cytotoksyczne (IC_{50} w zakresie 15–85 μM) wykazują związki Boc-dW-71, Z-W-71, H-W-71, H-dF-73, Boc-W-74, H-dW-74, przy czym aktywność ta jest nieselektywna lub nieznacznie selektywna.

- Innymi związkami o działaniu silnie cytotoksycznym, ale jednocześnie selektywnym wobec linii komórkowej COLO679, są *N*-acylohydrazony H-dW-71 oraz Z-dW-71.
- Szczególnie korzystnym profilem aktywności cechuje się związek Z-dW-71, który w stężeniu około 20 μM redukuje żywotność komórek COLO679 o 50% względem kontroli, a inkubowany z komórkami prawidłowymi BJ w stężeniach nawet do 100 μM nie ma żadnego wpływu na żywotność ($\text{IC}_{50} > 150 \mu\text{M}$). Co więcej również w tym przypadku można zaobserwować stereoselektywność efektu, ponieważ związek Z-W-71 wykazuje aktywność cytotoksyczną względem komórek COLO679 na zbliżonym poziomie co Z-dW-71, ale nie jest selektywny jak jego odpowiednik. Co ciekawe, związek Z-dW-71 we wcześniej omówionych badaniach powinowactwa nie wykazuje znaczącego powinowactwa do receptora NK1. Także związek H-dW-71 nie jest najsilniejszym ligandem NK1R z serii W/F-71–74, co sugeruje, że mechanizm działania cytotoksycznego tych związków jest zapewne inny niż przez receptor NK1.

Tabela 8. Badanie wpływu związków z serii W/F-71 – W/F-74 na żywotność (cytotoksyczność) linii komórkowych COLO679 oraz BJ (IC_{50}).

Lp.	Związek	IC_{50} [nM]		Lp.	Związek	IC_{50} [nM]		Lp.	Związek	IC_{50} [nM]	
		COLO679	BJ			COLO679	BJ			COLO679	BJ
Seria W-71				Seria F-72				Seria W-74			
1	Boc-W-71	> 100	~ 150	25	Boc-F-72	> 200	> 200	49	Boc-W-74	86	85
2	Boc-dW-71	16,6	36,5	26	Boc-dF-72	> 200	> 200	50	Boc-dW-74	> 200	> 200
3	Z-W-71	22	33	27	Z-F-72	> 200	> 200	51	Z-W-74	> 200	> 200
4	Z-dW-71	21	> 150	28	Z-dF-72	> 200	> 200	52	Z-dW-74	> 200	> 200
5	H-W-71	16,6	40	29	H-F-72	> 200	> 200	53	H-W-74	142	138
6	H-dW-71	15,9	69,6	30	H-dF-72	> 200	> 200	54	H-dW-74	58	70
7	Ac-W-71	> 200	~ 150	31	Ac-F-72	> 200	> 200	55	Ac-W-74	> 200	> 200
8	Ac-dW-71	> 150	> 200	32	Ac-dF-72	> 200	> 200	56	Ac-dW-74	> 200	> 200
Seria F-71				Seria W-73				Seria F-74			
9	Boc-F-71	> 200	~ 200	33	Boc-W-73	> 200	> 200	57	Boc-F-74	> 200	> 200
10	Boc-dF-71	> 150	> 150	34	Boc-dW-73	> 200	> 200	58	Boc-dF-74	> 200	> 200
11	Z-F-71	> 150	> 200	35	Z-W-73	> 200	> 200	59	Z-F-74	> 200	> 200
12	Z-dF-71	> 200	> 200	36	Z-dW-73	> 200	> 200	60	Z-dF-74	> 200	> 200
13	H-F-71	50-100	> 200	37	H-W-73	> 200	> 200	61	H-F-74	> 200	> 200
14	H-dF-71	50-100	> 100	38	H-dW-73	> 200	> 200	62	H-dF-74	> 200	> 200
15	Ac-F-71	> 200	> 200	39	Ac-W-73	> 200	> 200	63	Ac-F-74	> 200	> 200
16	Ac-dF-71	> 200	> 200	40	Ac-dW-73	> 200	> 200	64	Ac-dF-74	> 200	> 200
Seria W-72				Seria F-73							
17	Boc-W-72	> 200	> 200	41	Boc-F-73	> 200	> 200				
18	Boc-dW-72	> 200	> 200	42	Boc-dF-73	> 200	> 200				
19	Z-W-72	> 200	> 200	43	Z-F-73	> 200	> 200				
20	Z-dW-72	> 200	> 200	44	Z-dF-73	> 200	> 200				
21	H-W-72	> 200	> 200	45	H-F-73	> 200	> 200				
22	H-dW-72	> 200	> 200	46	H-dF-73	31	53				
23	Ac-W-72	> 200	> 200	47	Ac-F-73	> 200	> 200				
24	Ac-dW-72	> 200	> 200	48	Ac-dF-73	> 200	> 200				

IC_{50} – stężenie, przy którym odczyt żywotności metabolicznej jest zmniejszony o 50%.

Oznaczenia cytotoksyczności wykonała dr Joanna Matalińska z ZN IMDiK PAN.

IV.3.4. Hipoteza celu molekularnego dla związków z serii W-71

Otrzymane wyniki aktywności komórkowej dla związku Z-dW-71 i innych związków z serii W-71, oraz wyniki powinowactwa do NK1R wskazują, że mechanizm działania cytotoksycznego tych związków jest prawdopodobnie inny niż przez NK1R. W celu zgłębienia tego zagadnienia postanowiłam opracować hipotezę dotyczącą potencjalnego celu molekularnego tych związków, korzystając z metod *in silico*.

W tym celu skorzystałam z powszechnie dostępnego narzędzia SwissTargetPrediction [289,290]. Serwer ten przewiduje potencjalne cele molekularne badanych związków poprzez obliczanie ich podobieństwa do znanych ligandów różnych celów molekularnych. Warto podkreślić, że wynik działania SwissTarget jest jedynie predykcją, która wymaga potwierdzenia eksperymentalnego. Predykcje tego typu mogą być jednak bardzo przydatne w planowaniu pracy eksperymentalnej.

W ramach badań przeprowadzono analizę wszystkich hydrazonów z serii W-71 za pomocą SwissTarget. Pomimo tego, że dla większości przypadków receptor NK1 został przewidziany jako pierwszy potencjalny cel molekularny (najwyższe prawdopodobieństwo według SwissTarget), to na podstawie naszych danych eksperymentalnych wiadomo, że wiązalność serii W-71 do NK1R, jeśli istnieje, jest znikoma. Przeanalizowano więc inne sugerowane cele molekularne, odrzucając te związane na przykład z metabolizmem, przewodnictwem nerwowym, czy białka transporterowe. Dodatkowo, odrzucono również cele molekularne, w przypadku których program był trenowany na małej liczbie punktów eksperymentalnych.

Wyniki przeprowadzonej analizy sugerują, że potencjalnym celem molekularnym związanym z aktywnością komórkową wobec linii czerniaka może być receptor CCK2 (CCK2R lub inaczej CCK_BR). W Tabeli 9 przedstawiono niektóre znane ligandy CCK2, na podstawie których SwissTarget przewiduje prawdopodobieństwo oddziaływania z CCK2R dla Z-dW-71. Wszystkie te związki zawierają (przynajmniej) trzy ugrupowania aromatyczne). W każdej pochodnej jeden z pierścieni aromatycznych podstawiony jest grupą CF₃. We wszystkich związkach występuje wiązanie amidowe.

Za tym, że CCK2R może być celem molekularnym związanym z cytotoksycznością przemawia np. to, że:

- CCK2R ulega ekspresji w różnych typach nowotworów [291]. Zaangażowany jest w przeżywanie i proliferację komórek nowotworowych.

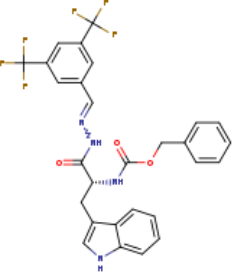
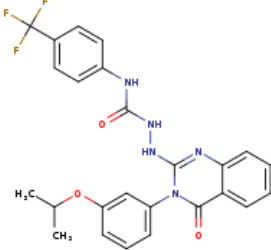
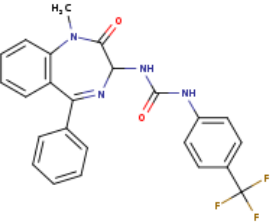
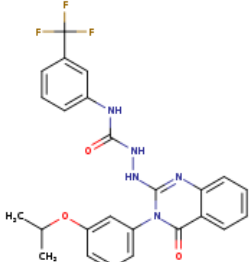
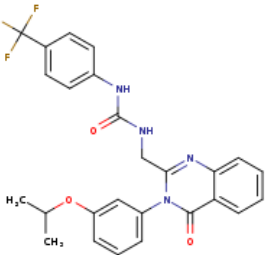
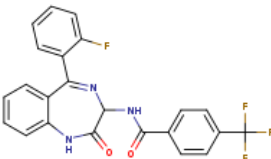
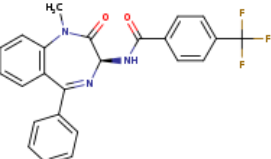
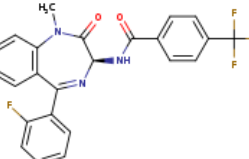
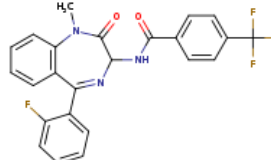
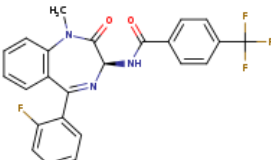
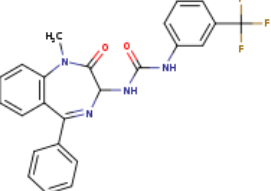
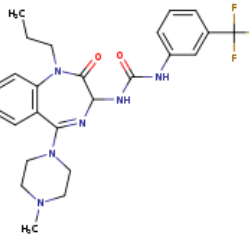
- Ostatnio wykazano, że CCK2R jest ekspresowany na niektórych liniach komórkowych czerniaka [292].
- Wcześniejsze badania pokazały także, że *in vivo* długotrwałe podawanie selektywnego antagonisty CCK2R spowalnia wzrost czerniaka [293].

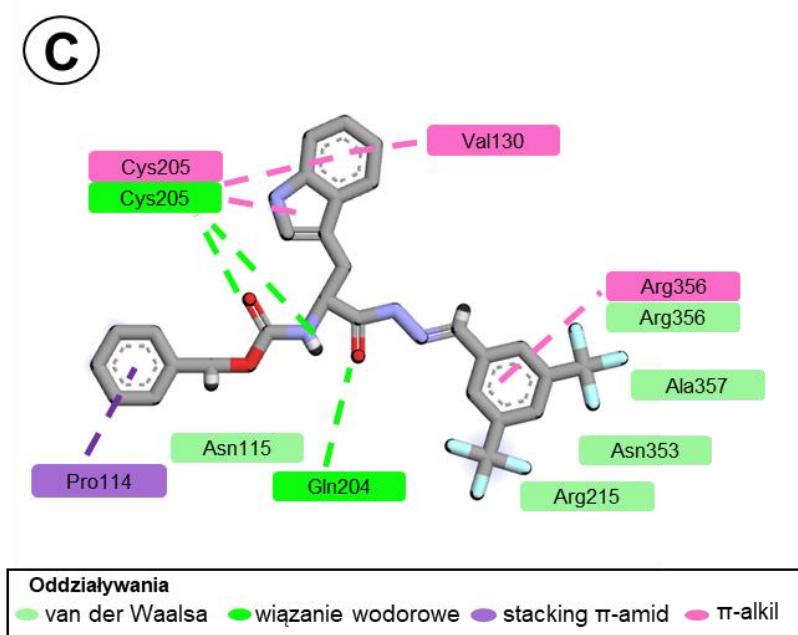
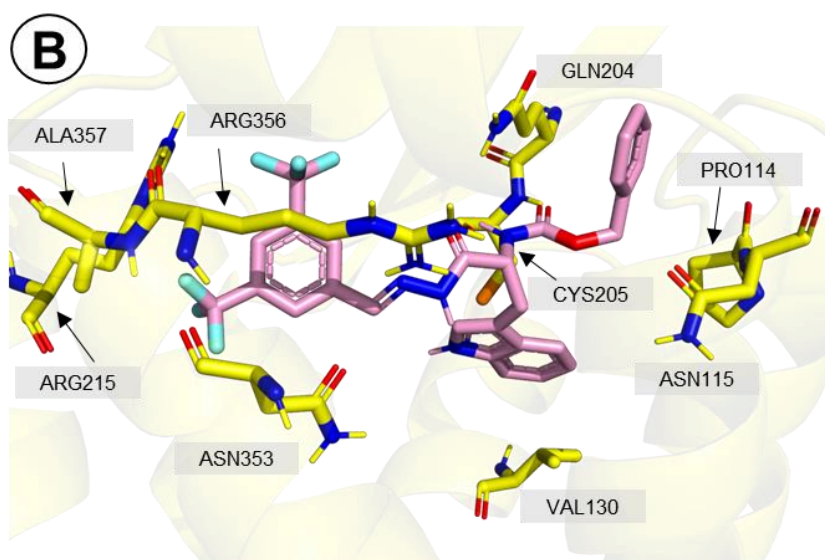
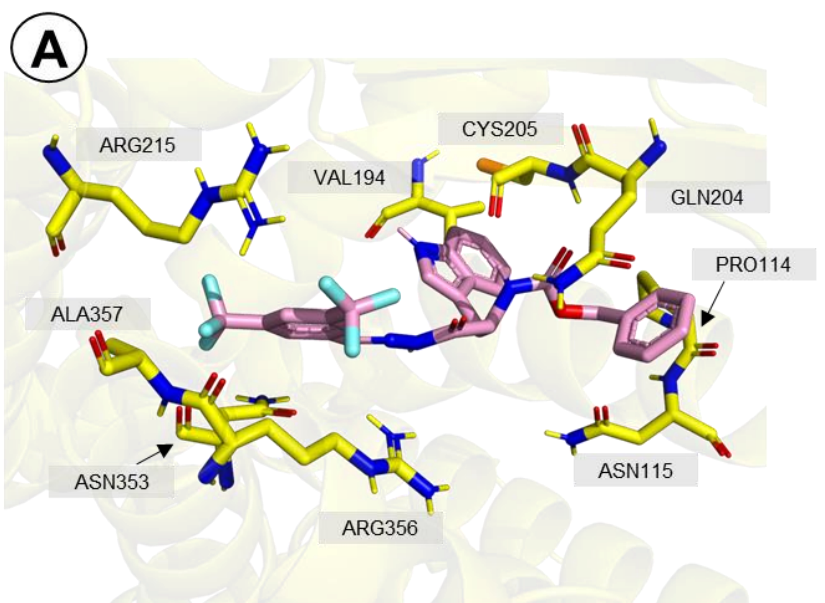
Dodatkowo, aby sprawdzić, czy oddziaływanie Z-dW-71 z CCK2R jest prawdopodobne z punktu widzenia strukturalnego, zadokowałam ten związek do receptora za pomocą programu AutoDock Vina. Program przewiduje wiązanie Z-dW-71 od CCK2R z entalpią swobodną równą -10.3 kcal/mol. Poza wiążącą przedstawiona jest na Rysunku 23.

Ligand w zadokowanym kompleksie znajduje się w górnej części kieszeni wiążącej, przy czym dół tej kieszeni pozostaje niewypełniony. Kompleks stabilizowany jest przez trzy wiązania wodorowe. Tworzą się one pomiędzy 1) tlenem z ugrupowania *N*-acylohydrazonu liganda a wodorem z grupy amidowej łańcucha bocznego Gln204, 2) wodorem z ugrupowania karbaminowego liganda a tlenem z ugrupowania amidowego (łańcuch główny) Cys205 oraz 3) tlenem z ugrupowania karbaminowego a wodorem z ugrupowania amidowego (łańcuch główny) Cys205. Pierścień fenyłowy Cbz jest wsunięty między Pro114 (stacking π -amid) a łańcuch boczny Gln204. Grupa -CH₂- Cbz jest położona w pobliżu łańcucha bocznego Asn115. Pierścień fenyłowy dipodstawiony grupami CF₃ oddziałuje z Arg356 poprzez oddziaływania π -alkil. Grupy CF₃ tworzą również kontakty hydrofobowe z resztami Arg215, Asn353, Arg356 i Ala357. W kierunku dna kieszeni wiążącej skierowany jest pierścień indolowy, który uczestniczy w oddziaływaniach π -alkil z łańcuchami bocznymi Val130 oraz Cys205.

Uzyskane podczas dokowania molekularnego predykcje sugerują dość dobry scoring oraz wiarygodną pozę. Dane te uprawdopodobniają przewidywania uzyskane za pomocą SwissTarget o prawdopodobieństwie oddziaływania Z-dW-71 z CCK2R, o którym wiadomo, że jest zaangażowany w procesy nowotworowe. Jednakże warto zauważyć, że wyniki scoringu są jedynie przybliżone i mogą być niedokładne, a SwissTarget bazuje na analogiach. Dlatego też, pełna weryfikacja tej hipotezy wymaga dalszych badań eksperymentalnych, co stanowi obiecujący kierunek dla przyszłych badań nad tym związkiem.

Tabela 9. Wybrane związki, na podstawie których SwissTarget przewiduje prawdopodobieństwo oddziaływania Z-dW-71 z receptorem CCK2.

Z-dW-71	CHEMBL415965	CHEMBL135538	CHEMBL8954
	Podobieństwo: 0.861 	Podobieństwo: 0.858 	Podobieństwo: 0.854 
CHEMBL167169 Podobieństwo: 0.827	CHEMBL73561 Podobieństwo: 0.822	CHEMBL2092865 Podobieństwo: 0.809	CHEMBL1907841 Podobieństwo: 0.803
			
CHEMBL305951 Podobieństwo: 0.803	CHEMBL1907842 Podobieństwo: 0.765	CHEMBL337055 Podobieństwo: 0.759	CHEMBL105273 Podobieństwo: 0.755
			



Rysunek 23. Wyniki dokowania do CCK2R dla związku Z-dW-71.

A) Sposób oddziaływania związku Z-dW-71 (4) z receptorem CCK2 w miejscu wiążącym przewidziany na podstawie dokowania molekularnego (widok z góry)

B) Sposób oddziaływania związku Z-dW-71 (4) z receptorem CCK2 w miejscu wiążącym przewidziany na podstawie dokowania molekularnego (widok z boku)

C) Dwuwymiarowy schemat oddziaływań. Oddziaływania zaznaczone zgodnie z legendą.

Ligand zaznaczono na różowo. Heteroatomy mają kolory zgodnie z przyjętą powszechnie konwencją (O – czerwony, N – ciemnoniebieski, F – jasnoniebieski, S – pomarańczowy). Receptor (kolor żółty) przedstawiono w sposób uproszczony, uwzględniono jedynie wybrane helisy (wstążki) i łańcuchy boczne (patyczki). Dla przejrzystości ukryto atomy wodoru (z wyjątkiem polarnych).

IV.4. Badanie skuteczności funkcji oceniającej programu AutoDock Vina

Druga część moich badań dotyczyła aspektów metodologicznych, związanych ze skutecznością funkcji oceniających stosowanych w dokowaniu molekularnym. Dokowanie molekularne jest jedną z ważnych technik *in silico* stosowanych w badaniach nad nowymi lekami. Celem symulacji dokowania jest przewidywanie oddziaływań między potencjalnymi cząsteczkami leków a wybranymi celami molekularnymi. Dokowanie molekularne jest metodą stosunkowo szybką i tanią obliczeniowo. Dzięki temu zyskało ono bardzo dużą popularność jako narzędzie racjonalnego projektowania leków. Z drugiej jednak strony, stosowanie tej techniki wiąże się z różnymi ograniczeniami, z których być może najważniejszym jest niedokładność predykcji siły wiązania przez funkcje oceniające.

W praktyce ZN IMDiK PAN zaobserwowano, że domyślna funkcja oceniająca programu AutoDock Vina daje niskiej jakości szacowania energii wiązania ligandów z NK1R. Postanowiłam więc systematycznie sprawdzić zdolność tego programu do rozróżniania związków wiążących się od niewiążących się (słabowiążących się) z NK1R na szerszym zbiorze danych eksperymentalnych.

Przebieg badania był następujący:

1. dokowałam ligandy ze Zbiorów I, II i III (opisane poniżej) do kilku struktur NK1R,
2. sortowałam listę rankingową według przewidywanej energii wiązania,
3. przypisywałam zadokowanym ligandom kategorię aktywny/nieaktywny (na podstawie danych eksperymentalnych),
4. dla posortowanego rankingu obliczałam metryki sukcesu (opisane poniżej).

Moje badanie pozwoliło zatem ocenić tzw. „moc skryningu”, czyli zdolność programu do trafnego selekcjonowania silnie wiążących się ligandów od związków nieaktywnych.

Badanie skuteczności funkcji oceniającej przeprowadziłam z wykorzystaniem trzech zbiorów danych eksperymentalnych:

- Zbiór I – związki, które zsyntezowałam w ramach mojego projektu doktoranckiego (oraz kilka innych związków otrzymanych w ZN IMDiK PAN), których powinowactwo receptorowe do NK1R zostało oznaczone w ZN IMDiK PAN,

- Zbiór II – związki, dla których dane o powinowactwie do receptora NK1 pochodzą z literatury naukowej (pobrane z bazy ChEMBL [257]), z wyłączeniem związków obecnych w Zbiorze III,
- Zbiór III – związki, których powinowactwo do receptora NK1 było badane w dużym badaniu przesiewowym DRUGMATRIX [294,295] (pobrane z bazy ChEMBL [257]) zbiór bardzo zróżnicowany strukturalnie, zawierający m.in. zatwierdzone leki.

Zbiór I bardzo dobrze nadaje się do „trudnego” testowania funkcji oceniającej, ponieważ większość zawartych w nim związków jest nieaktywna wobec NK1R, a jednocześnie jest strukturalnie podobna do typowych antagonistów NK1R (związki te posiadają dwa lub trzy ugrupowania aromatyczne, oraz jeden z pierścieni aromatycznych podstawiony grupami CF₃ w pozycji 3 i 5). Aby rozszerzyć zakres możliwości testowania, zdecydowałam się także dołączyć zbiory związków pochodzących z literatury, pobrane z bazy ChEMBL (Zbiór II i Zbiór III). W Zbiorze II znajdują się związki strukturalnie podobne do znanego farmakoforu NK1R, z których większość wykazuje znaczące powinowactwo do NK1R. Zbiór III zawiera związki o zróżnicowanej budowie strukturalnej, nieprzypominającej klasycznych antagonistów NK1R, i nieaktywne wobec NK1R.

W ramach przeprowadzonych badań oceny skuteczności funkcji oceniającej, wybrałam cztery scenariusze testowania (A-D, podsumowane w Tabeli 10), różniące się:

- używanym zestawem danych eksperymentalnych – Zbiór I i II lub Zbiór I i II i III, [Połączenie Zbioru I i II skutkowało otrzymaniem zestawu testowego składającego się ze związków o podobnej budowie strukturalnej, ale zróżnicowanych pod względem aktywności wobec NK1R. W wyniku połączenia Zbioru I i II i III otrzymano zestaw testowy bardziej zróżnicowany strukturalnie, w których wiele nieaktywnych związków różni się strukturalnie od związków aktywnych.]
- wybranymi progami odcięcia dla przypisywania kategorii aktywny/nieaktywny (pAff¹ = 6.0 lub 6.5).

Dodatkowo w obrębie każdego scenariusza, wyróżniono „warianty” różniące się:

- sposobem traktowania danych z dokowania do różnych struktur NK1R,

[ponieważ w bazie PDB dostępnych jest kilka struktur NK1R [23–28] nie jest z góry oczywiste, która struktura powinna zostać użyta do skryningu; w związku z tym wybrano pięć struktur (nie dublujących się konformacyjnie) i testowano oddzielnie *scoring* uzyskany w dokowaniu do jednej z pięciu struktur albo *scoring* będący średnią lub minimum z dokowania do tych pięciu struktur].

¹ pAff – ujemny logarytm powinowactwa eksperymentalnego (próg odcięcia pAff = 6.0 oznacza, że związki o aktywności pAff < 6.0 (1 uM) uznawane są za nieaktywne)

Tabela 10. Scenariusze eksperymentu wirtualnego przesiewu.

Scenariusz	Funkcja oceniająca	Zestaw danych (ligandy)	Próg odcięcia pAff ¹ aktywny-nieaktywny	Jak traktowane są dane z dokowania do różnych struktur receptora?
A	Oryginalna funkcja VINA	Wylosowane podzbiory Zbiorów I + II ²	6.0	Każda struktura receptora traktowana osobno
				Jako scoring przyjmuje się średnią energię z dokowań do wszystkich struktur receptora
B		6.5	Każda struktura receptora traktowana osobno	
			Jako scoring przyjmuje się średnią energię z dokowań do wszystkich struktur receptora	
C	Wylosowane podzbiory Zbiorów I + II + III ²	6.0	Każda struktura receptora traktowana osobno	
			Jako scoring przyjmuje się średnią energię z dokowań do wszystkich struktur receptora	
D		6.5	Każda struktura receptora traktowana osobno	
			Jako scoring przyjmuje się średnią energię z dokowań do wszystkich struktur receptora	

¹ pAff – ujemny logarytm powinowactwa eksperymentalnego (próg odcięcia pAff = 6.0 oznacza, że związki o aktywności pAff < 6.0 (1 uM) uznawane są za nieaktywne)

² 500 podzbiorów po 410 ligandów (w tym: 10/400 aktywnych/nieaktywnych)

W każdym scenariuszu, testowanie skuteczności wykonałam wielokrotnie. Do testowania wylosowano pięćset podzbiorów rozważanych Zbiorów eksperymentalnych. Podzbiory zawierały losowo wybrane związki aktywne i nieaktywne, tak aby w każdym znajdowało się dziesięć związków aktywnych, a na każdy aktywny związek przypadało czterdzieści związków nieaktywnych. Losowanie pozwalało na stworzenie sytuacji testowej zbliżonej do rzeczywistych warunków, w których w bibliotekach przesiewowych występuje zauważalna dysproporcja między liczbą nieaktywnych a aktywnych związków, a biblioteki te są często niebilansowane pod względem właściwości i aktywności. Dodatkowo, wykorzystanie losowania i wielokrotnego testowania pozwala na oszacowanie statystyki oraz błędów związanych z przeprowadzanymi badaniami. Analogiczna praktyka losowania zbiorów testowych jest powszechnie stosowana, m.in. w walidacji metod QSAR [296].

Do oceny skuteczności funkcji oceniającej użyto powszechnie stosowanych metryk skuteczności przesiewu wirtualnego (odróżnianie związków aktywnych od nieaktywnych):

- logAUC – pole pod krzywą charakterystyki operacyjnej odbiornika ROC (ang. *receiver operating characteristic curve*) wykreślonej w skali semilogarytmicznej [297],
[wartości koło 0 oznaczają brak skuteczności wyższej niż losowa klasyfikacja („rzut monetą”), maksymalna wartość to 86.3]
- $EF_{fracx\%}$ – znormalizowany współczynnik wzbogacenia przy pewnym górnym x% bazy
[im wartość EF_{frac} jest bliższa 1, tym metoda przesiewu jest bardziej skuteczna (EF_{frac} równy 1 oznacza pełne możliwe wzbogacenie w porównaniu z maksymalnym teoretycznym wzbogaceniem)].

W celu porównania średnich wyników wartości logAUC otrzymanych dla różnych wariantów *scoringu* w każdym scenariuszu przeprowadzono jednokierunkową analizę ANOVA z testem *post hoc* wielokrotnych porównań Sidaka.

Dodatkowo, w omówieniu wyników zastosowano opisową kategoryzację wyników skryningu według wartości logAUC zaproponowaną przez Jaiteha i wsp. [298]. Według tych autorów (dalej w tekście: „kryterium Jaiteha”), moc przesiewania metody można określić jako:

- słabą ($\log AUC < 10.0$),
- niezłą (ang. *fair*; $10.0 < \log AUC < 15.0$),
- dobrą ($15.0 < \log AUC < 20.0$),
- bardzo dobrą ($20.0 < \log AUC < 25.0$),
- doskonałą ($\log AUC > 25.0$).

Parametry statystyczne dla scenariusza A przedstawiono w Tabeli 11. Średnie wartości logAUC dla poszczególnych wariantów wynoszą między 5.4 a 8.5 (słaba moc przesiewania). Najlepszy wynik uzyskano dla dokowania z wykorzystaniem struktury receptora 6HLO (logAUC = 8.5) i 6HLL (logAUC = 7.6, różnica między nimi jest niestotna statystycznie). Najniższą średnią wartość logAUC wykazał wariant, w którym za *scoring* przyjmowano najniższą wartość spośród zadokowanych do różnych receptorów (minimum; logAUC = 5.4). Wartość $EF_{frac1\%}$ przeważnie mieści się w zakresie 0.0–0.1 (51.0–75.6% obserwacji, średnio 63% obserwacji), zaś pozostała część przypada głównie na przedział od 0.1 do 0.2 (21.6–38.2% obserwacji). Wartość $EF_{frac5\%}$ w zdecydowanej większości obserwacji (86.0–94.2%) znajduje się w zakresie 0.0–0.1. Przy obu EF_{frac} , *scoring* z 6HLO i 6HLL daje większy odsetek obserwacji w wyższych przedziałach.

Parametry statystyczne dla scenariusza B przedstawiono w Tabeli 12. Średnie wartości logAUC dla poszczególnych wariantów wynoszą między 4.8 a 7.9 (słaba moc przesiewania). Najlepszy wynik uzyskano dla dokowania z wykorzystaniem struktury receptora 6HLO (logAUC = 7.9) oraz 6HLL (logAUC = 7.0, różnica między nimi jest niestotna statystycznie). Najniższą średnią wartość logAUC wykazał wariant, w którym za *scoring* przyjmowano najniższą wartość spośród zadokowanych do różnych receptorów (minimum; logAUC = 4.8). Wartość $EF_{frac1\%}$ dla większej liczby obserwacji mieści się w zakresie 0.0–0.1 (49.6–75.4%, średnio 67% obserwacji), zaś pozostała część przypada głównie na przedział od 0.1 do 0.2 (20.2–36.8% obserwacji). Wartość $EF_{frac5\%}$ dla większości obserwacji (85.0–95.6%) znajduje się w zakresie 0.0–0.1. Przy obu EF_{frac} , *scoring* z 6HLO i 6HLL daje większy odsetek obserwacji w wyższych przedziałach.

Tabela 11. Średnie wartości parametru logAUC i rozkład wartości EF_{frac} dla dokowań w Scenariuszu A.

Scenariusz	Wariant metody	logAUC ¹ [%]	Rozkład wartości EF _{frac} ² przy górnym 1% bazy [%]					Rozkład wartości EF _{frac} ² przy górnych 5% bazy [%]				
			0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0	0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0
A	6E59	6.2 ± 5.1	75.6	21.6	0.2	2.6	0.0	91.8	8.0	0.2	0.0	0.0
	6HLL	7.6 ± 5.6	51.2	38.2	1.2	9.4	0.0	86.0	13.8	0.2	0.0	0.0
	6HLO	8.5 ± 5.1	58.2	34.4	1.0	6.4	0.0	91.6	8.4	0.0	0.0	0.0
	7P00	5.9 ± 5.1	62.0	33.0	0.2	4.8	0.0	93.8	6.2	0.0	0.0	0.0
	7RMG	5.9 ± 5.3	67.0	30.8	0.2	1.8	0.2	92.6	7.4	0.0	0.0	0.0
	średnia	7.3 ± 5.8	64.8	30.0	0.4	4.4	0.4	87.4	11.8	0.8	0.0	0.0
	minimum	5.4 ± 5.0	64.2	31.8	0.2	3.8	0.0	94.2	5.8	0.0	0.0	0.0
Porównanie wartości logAUC dla poszczególnych wariantów												
<p>← wykres skrzypcowy („skrzypce” przedstawiają gęstość rozkładu [jądrowy estymator gęstości]. czerwona linia oznacza medianę)</p> <p>próg pAff>6.0 zestaw I+II</p>			<p>N = 500 (500 losowanych podzbiorów. każdy po 10 związków aktywnych i 10*40 związków nieaktywnych) Jednokierunkowa ANOVA (F = 22.42, p < 0.0001) z testem wielokrotnych porównań Sidaka, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>									
			6E59	6HLL	6HLO	7P00	7RMG	średnia	minimum			
			6E59									
			6HLL	***								
			6HLO	****	ns							
			7P00	ns	****	****						
			7RMG	ns	****	****	ns					
			średnia	*	ns	**	***	***				
			minimum	ns	****	****	ns	ns	****			
<p>¹ logAUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² EF_{frac} – znormalizowany współczynnik wzbogacenia</p>												

Tabela 12. Średnie wartości parametru $\log AUC$ i rozkład wartości EF_{frac} dla dokowań w Scenariuszu B.

Scenariusz	Wariant metody	$\log AUC^1$ [%]	Rozkład wartości EF_{frac}^2 przy górnym 1% bazy [%]					Rozkład wartości EF_{frac}^2 przy górnych 5% bazy [%]				
			0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0	0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0
B	6E59	6.1 ± 5.4	67.2	25.4	2.6	4.2	0.6	92.0	8.0	0.0	0.0	0.0
	6HLL	7.0 ± 5.7	49.6	36.8	3.4	8.6	1.6	85.0	15.0	0.0	0.0	0.0
	6HLO	7.9 ± 5.4	60.0	30.4	2.2	6.4	1.0	91.6	8.0	0.4	0.0	0.0
	7P00	5.3 ± 4.9	75.0	20.6	1.8	2.0	0.6	94.8	5.2	0.0	0.0	0.0
	7RMG	5.4 ± 5.0	72.6	22.8	2.2	2.0	0.4	93.4	6.6	0.0	0.0	0.0
	średnia	6.0 ± 5.0	66.2	28.2	1.8	3.8	0.0	90.8	9.2	0.0	0.0	0.0
	minimum	4.8 ± 4.9	75.4	20.2	2.0	1.8	0.6	95.6	4.4	0.0	0.0	0.0
Porównanie wartości $\log AUC$ dla poszczególnych wariantów												
<p>próg $pAff > 6.5$ zestaw I+II</p> <p>← wykres skrzypcowy („skrzypce” przedstawiają gęstość rozkładu [jądrowy estymator gęstości], czerwona linia oznacza medianę)</p> <p>N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA (F = 20.32, p < 0.0001) z testem wielokrotnych porównań Sidaka, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, **** – p ≤ 0.0001.</p>												
			6E59	6HLL	6HLO	7P00	7RMG	średnia	minimum			
			6E59									
			6HLL	ns								
			6HLO	****	ns							
			7P00	ns	****	****						
			7RMG	ns	****	****	ns					
			średnia	ns	ns	****	ns	ns				
			minimum	**	****	****	ns	ns	**			
<p>¹ $\log AUC$ – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe;</p> <p>² EF_{frac} – znormalizowany współczynnik wzbogacenia</p>												

Parametry statystyczne dla scenariusza C przedstawiono w Tabeli 13. Średnie wartości logAUC dla poszczególnych wariantów wynoszą między 14.5 a 18.6 (według kryterium Jaiteha, wynik niezły lub dobry). Najlepszy wynik uzyskano dla dokowania z wykorzystaniem struktury receptora 6HLL (logAUC = 18.6) i 6HLO (logAUC = 18.5, różnica między nimi jest nieistotna statystycznie), zaś najniższą średnią wartość logAUC otrzymano dla 7RMG (logAUC = 14.4). Wartości $EF_{frac1\%}$ w zakresie 0.0–0.1 przypadają dla 37.6–48.4% obserwacji, w zakresie 0.1–0.2 dla 16.0–19.8%, w zakresie 0.2–0.3 znajduje się 20.0–24.6% obserwacji, a 8.8–14.0% wykazuje wartość w przedziale od 0.4 do 1.0. Wartość $EF_{frac5\%}$ dla większości obserwacji znajduje się w zakresie 0.0–0.1 (54.2–69.2%) lub 0.1–0.2 (27.8–40.0%). Przy obu EF_{frac} , *scoring* z 6HLO i 6HLL daje większy odsetek obserwacji w wyższych przedziałach.

Parametry statystyczne dla scenariusza D przedstawiono w Tabeli 14. Średnie wartości logAUC dla poszczególnych wariantów wynoszą między 12.3 a 15.2 (według kryterium Jaiteha, wynik niezły lub dobry). Najlepszy wynik uzyskano dla dokowań z wykorzystaniem struktur receptora 6HLL i 6HLO (dla obu logAUC = 15.2), zaś najniższą średnią wartość logAUC otrzymano dla 7RMG (logAUC = 12.3). Wartości $EF_{frac1\%}$ w zakresie 0.0–0.1 przypadają dla 37.6–59.8% obserwacji, w zakresie 0.1–0.2 dla 13.2–19.0%, w zakresie 0.2–0.3 znajduje się 18.4–26.4% obserwacji, a 5.3–10.6% wykazuje wartość w przedziale od 0.4 do 1.0. Wartość $EF_{frac5\%}$ dla większości obserwacji znajduje się w zakresie 0.0–0.1 (65.4–78.2%) lub 0.1–0.2 (19.4–27.6%). Przy obu EF_{frac} , *scoring* z 6HLO i 6HLL daje większy odsetek obserwacji w wyższych przedziałach.

Tabela 13. Średnie wartości parametru $\log AUC$ i rozkład wartości EF_{frac} dla dokowań w Scenariuszu C.

Scenariusz	Wariant metody	$\log AUC^1$ [%]	Rozkład wartości EF_{frac}^2 przy górnym 1% bazy [%]					Rozkład wartości EF_{frac}^2 przy górnych 5% bazy [%]																																																																				
			0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0	0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0																																																																
C	6E59	16.2 ± 5.9	46.8	15.0	23.2	6.2	8.8	54.2	40.0	5.8	0.0	0.0																																																																
	6HLL	18.6 ± 6.3	37.6	19.8	24.4	4.2	14.0	55.0	38.2	6.4	0.4	0.0																																																																
	6HLO	18.5 ± 5.9	39.4	19.8	24.6	5.2	11.0	57.4	36.0	6.4	0.2	0.0																																																																
	7P00	15.7 ± 5.9	47.0	18.2	21.6	4.0	9.2	62.8	32.8	4.0	0.4	0.0																																																																
	7RMG	14.4 ± 6.0	46.4	17.6	22.0	5.4	8.6	69.2	27.8	2.8	0.2	0.0																																																																
	średnia	17.4 ± 6.3	38.8	17.8	23.4	7.4	12.6	61.8	31.6	6.2	0.4	0.0																																																																
	minimum	15.5 ± 5.8	48.4	18.2	20.0	4.4	9.0	66.4	29.0	4.4	0.2	0.0																																																																
Porównanie wartości $\log AUC$ dla poszczególnych wariantów																																																																												
<p>← wykres skrzypcowy („skrzypce” przedstawiają gęstość rozkładu [jądrowy estymator gęstości], czerwona linia oznacza medianę)</p> <p>próg $p_{Aff} > 6.0$ zestaw I+II+III</p> <p>logAUC</p> <p>warianty metody</p>			<p>N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA (F = 35.71, $p < 0.0001$) z testem wielokrotnych porównań Sidaka, oznaczenia istotności: ns – $p > 0.05$ (brak istotnej statystycznie różnicy), * – $p \leq 0.05$, ** – $p \leq 0.01$, *** – $p \leq 0.001$, **** – $p \leq 0.0001$.</p> <table border="1"> <thead> <tr> <th></th> <th>6E59</th> <th>6HLL</th> <th>6HLO</th> <th>7P00</th> <th>7RMG</th> <th>średnia</th> <th>minimum</th> </tr> </thead> <tbody> <tr> <td>6E59</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>6HLL</td> <td>****</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>6HLO</td> <td>****</td> <td>ns</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>7P00</td> <td>ns</td> <td>****</td> <td>****</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>7RMG</td> <td>****</td> <td>****</td> <td>****</td> <td>*</td> <td></td> <td></td> <td></td> </tr> <tr> <td>średnia</td> <td>*</td> <td>*</td> <td>ns</td> <td>***</td> <td>****</td> <td></td> <td></td> </tr> <tr> <td>minimum</td> <td>ns</td> <td>****</td> <td>****</td> <td>ns</td> <td>*</td> <td>****</td> <td></td> </tr> </tbody> </table>											6E59	6HLL	6HLO	7P00	7RMG	średnia	minimum	6E59								6HLL	****							6HLO	****	ns						7P00	ns	****	****					7RMG	****	****	****	*				średnia	*	*	ns	***	****			minimum	ns	****	****	ns	*	****	
	6E59	6HLL	6HLO	7P00	7RMG	średnia	minimum																																																																					
6E59																																																																												
6HLL	****																																																																											
6HLO	****	ns																																																																										
7P00	ns	****	****																																																																									
7RMG	****	****	****	*																																																																								
średnia	*	*	ns	***	****																																																																							
minimum	ns	****	****	ns	*	****																																																																						
<p>¹ $\log AUC$ – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe;</p> <p>² EF_{frac} – znormalizowany współczynnik wzbogacenia</p>																																																																												

Tabela 14. Średnie wartości parametru $\log AUC$ i rozkład wartości EF_{frac} dla dokowań w Scenariuszu D.

Scenariusz	Wariant metody	$\log AUC^1$ [%]	Rozkład wartości EF_{frac}^2 przy górnym 1% bazy [%]					Rozkład wartości EF_{frac}^2 przy górnym 5% bazy [%]				
			0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0	0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–1.0
D	6E59	13.9 ± 6.2	46.8	16.4	23.6	4.8	8.4	68.8	27.6	3.6	0.0	0.0
	6HLL	15.2 ± 6.2	37.0	19.0	26.4	7.6	10.0	65.4	29.0	5.2	0.4	0.0
	6HLO	15.2 ± 6.3	47.2	13.4	24.0	4.8	10.6	68.6	27.8	3.4	0.2	0.0
	7P00	13.0 ± 5.6	59.0	13.8	18.4	3.2	5.6	75.8	21.2	3.0	0.0	0.0
	7RMG	12.3 ± 5.9	52.8	14.8	22.2	3.2	7.0	74.0	22.6	3.2	0.2	0.0
	średnia	13.5 ± 5.4	52.0	18.0	19.4	4.0	6.6	72.8	24.6	2.6	0.0	0.0
	minimum	12.5 ± 5.5	59.8	13.2	19.4	2.2	5.4	78.2	19.4	2.4	0.0	0.0
Porównanie wartości $\log AUC$ dla poszczególnych wariantów												
<p>← wykres skrzypcowy („skrzypce” przedstawiają gęstość rozkładu [jądrowy estymator gęstości], czerwona linia oznacza medianę)</p> <p>próg $p_{Aff} > 6.5$ zestaw I+II+III</p>			<p>N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA (F = 20.70, p < 0.0001) z testem wielokrotnych porównań Sidaka, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, **** – p ≤ 0.0001.</p>									
			6E59	6HLL	6HLO	7P00	7RMG	średnia	minimum			
			6E59									
			6HLL	**								
			6HLO	**	ns							
			7P00	ns	****	****						
			7RMG	***	****	****	ns					
			średnia	ns	***	***	ns	*				
			minimum	**	****	****	ns	ns	ns			
<p>¹ $\log AUC$ – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe;</p> <p>² EF_{frac} – znormalizowany współczynnik wzbogacenia</p>												

Otrzymane wartości logAUC we wszystkich scenariuszach są stosunkowo niskie. Maksymalna teoretycznie możliwa wartość logAUC wynosi około 86. W niniejszym badaniu otrzymano najwyższą średnią wartość logAUC równą 18.6 (według kryterium Jaiteha, wynik dobry), dla Scenariusza C, w którym $p_{Aff} > 6.0$ i testowano zestaw składający się ze Zbiorów I i II i III. Pokazuje to, że w naszym eksperymencie Vina względnie słabo radzi sobie z odróżnianiem związków nieaktywnych od aktywnych.

Niższe („gorsze”; słabe według kryterium Jaiteha) wartości logAUC otrzymano w scenariuszach, w których do przesiewu brano „trudniejszy” zestaw testowy składający się tylko ze Zbiorów I i II. Rozszerzenie zestawu testowego o związki ze Zbioru III skutkowało około 2–3-krotnym wzrostem wartości logAUC (wyniki w kategoriach niezły lub dobry według kryterium Jaiteha). Wynika to z faktu, że w Zbiorach I i II są związki, które są podobne do ligandów NK1R, a w Zbiorze III są różne ligandy, często nieprzypominające ligandów NK1R. Stąd AutoDock Vina mając więcej ligandów, które nie są podobne do ligandów NK1R, myli się rzadziej.

Wyższe („lepsze”) wartości logAUC otrzymano w scenariuszach, gdy próg odcięcia był równy 6.0. W przypadku, gdy próg odcięcia $p_{Aff} = 6.0$ klasyfikator uznaje za nieaktywne związki o powinowactwie gorszym niż $1\mu\text{M}$, a w przypadku $p_{Aff} = 6.5$ jako nieaktywne klasyfikowane są już związki z powinowactwem powyżej $0.3\mu\text{M}$. Ta druga sytuacja jest w oczywisty sposób „trudniejszym” testem dla *scoringu*.

Najwyższe („najlepsze”) wartości logAUC we wszystkich scenariuszach otrzymano dla *scoringów* z dokowań do struktur 6HLL i 6HLO. W każdym scenariuszu różnice średnich pomiędzy tymi dwoma strukturami są nieistotne statystycznie. Wskazuje to, że oba te warianty (przy pięciuset próbach) są tak samo skuteczne w przesiewaniu. Ich wyższa skuteczność w porównaniu do pozostałych wariantów może być związana z tym, że są to struktury związane z antagonistami NK1R, i być może konformacje receptora uchwycone w 6HLL i 6HLO są konformacjami natywnymi dla większej grupy ligandów w bazie ChEMBL.

Najniższe („najgorsze”) wartości logAUC otrzymano dla struktur 7RMG i 7P00, które są strukturami receptora związanego z agonistą peptydowym (konformacja receptora jest w stanie aktywowanym) [24]. W większości scenariuszy różnice średnich pomiędzy tymi dwoma strukturami są nieistotne statystycznie ($p > 0.05$), co może odzwierciedlać podobieństwo konformacyjne obu wariantów. Ich niższa skuteczność w tym testowaniu wiąże się prawdopodobnie z tym, że większość związków w bazie ChEMBL jest antagonistami NK1R.

Wyniki tego eksperymentu sugerują, że do przesiewu wirtualnego antagonistów powinno stosować się raczej struktury 6HLO lub 6HLL. Wydaje się, że warto byłoby zbadać skuteczność funkcji oceniającej dla agonistów, jednakże istnieje mało znanych agonistów (nie licząc peptydowych) i wymagałoby to oddzielnego systematycznego badania.

Obserwowane we wszystkich scenariuszach stosunkowo relatywnie duże odchylenia standardowe logAUC podkreślają konieczność randomizacji zbiorów testowych, dla wiarygodnego szacowania różnic pomiędzy wariantami metod wirtualnego przesiewania.

IV.5. Opracowanie funkcji oceniających specyficznych dla receptora NK1

Jak wykazano w poprzednim rozdziale, domyślna funkcja oceniająca programu AutoDock Vina nie najlepiej radzi sobie w szacowaniu energii wiązania ligandów z NK1R i odróżnianiu związków aktywnych od nieaktywnych, wykazując względnie niską moc przesiewania. Odzwierciedlone to jest w niskich wartościach logAUC przypadających do kategorii „słaby” według kryterium Jaiteha w przypadku trudniejszych scenariuszy testowych, lub „niezły” i czasami „dobry” w przypadku łatwiejszych scenariuszy testowych. W związku z tym podjęłam próbę opracowania nowych funkcji oceniających, specyficznych dla receptora NK1, które będzie można wykorzystywać do ponownej oceny (*rescoringu*) wyników dokowania w celu poprawy jakości klasyfikacji. Specyficzność moich funkcji dla NK1R, chciałam uzyskać dzięki trenowaniu tych funkcji na eksperymentalnych danych o powinowactwie wobec NK1R. Trenowanie nowych funkcji przeprowadzono za pomocą zestawu skryptów *VinaTuner* opracowanych w ZN IMDIK PAN przez dr. hab. Piotra Lipińskiego.

Przebieg tego badania był następujący:

TRENOWANIE

1. dokowałam ligandy z odpowiednich zbiorów treningowych (opisane poniżej) do wybranych struktur NK1R (opisane poniżej),
2. uzyskiwałam składowe *scoringu* dla pól o najniższej energii (por. Równanie 3-5),
3. dostrajałam wagi funkcji (por. Tabela 3) do danych eksperymentalnych za pomocą skryptów *VinaTuner*, [ten krok powtarzany był wielokrotnie dla różnych parametrów skryptów, por. Rodział VII.8,]
4. wybierałam najlepszych 5 funkcji z każdego przebiegu do walidacji,

WALIDACJA

5. dokowałam ligandy z odpowiednich zbiorów testowych (opisane poniżej),
6. dokonywałam *rescoringu* ligandów testowych za pomocą nowych funkcji,
7. sortowałam listę rankingową według energii wiązania przewidywanych za pomocą nowych funkcji,
8. przypisywałam zadokowanym ligandom kategorię aktywny/nieaktywny (na podstawie danych eksperymentalnych),
9. dla posortowanego rankingu obliczałam metryki sukcesu (opisane poniżej).

Trenowanie nowych funkcji przeprowadzono początkowo na Zbiorze I jako zestawie treningowym, według scenariuszy trenowania T1–T3 podsumowanych w Tabeli 15.

Zaletą Zbioru I jako zbioru do trenowania funkcji jest jego homogenne pochodzenie. Wszystkie wyniki uzyskano w jednym laboratorium, według jednego protokołu i na tym samym materiale błonowym. Ostatnio wskazano, że duże zbiory danych biologicznych o heterogennym pochodzeniu posiadają bardzo wysoki poziom „szumu” informacyjnego [299], który może negatywnie wpływać na jakość trenowania metod *in silico*.

Ponieważ podczas badania skuteczności funkcji oceniającej AutoDock Vina najlepsze wyniki uzyskano dla dokowania do struktur 6HLO i 6HLL, w trenowaniu jako pozę wiążącą ligandów wybrano pozę o najniższej energii ze zbioru póz z dokowania do wszystkich struktur NK1R, pozę o najniższej energii z dokowania do struktury 6HLO lub pozę o najniższej energii z dokowania do struktury 6HLL.

Trenowanie, polegające na dostrajeniu wag funkcji przeprowadzono za pomocą algorytmu genetycznego zaprogramowanego w zestawie skryptów *VinaTuner*. Celem dostrajania było uzyskanie jak najniższej wartości średniego błędu bezwzględnego (MAE, ang. *Mean Absolute Error*) między predykcją entalpii swobodnej wiązania uzyskaną za pomocą nowych funkcji a danymi eksperymentalnymi. Wybrane funkcje zostały użyte do *rescoringu* na zbiorze walidacyjnym. Walidacje przeprowadzono według różnych scenariuszy (W1–W4), opisanych w Tabeli 15, uwzględniając różne zestawy walidacyjne oraz progi odcięcia pAff aktywny/nieaktywny. Zbiór testowy był randomizowany, wygenerowano pięćset podzbiorów o proporcji 10/400 aktywnych/nieaktywnych związków. Miarą sukcesu nowych funkcji w eksperymencie walidacyjnym były wartości logAUC przesiewu wirtualnego. Porównanie z wartościami dla oryginalnej funkcji AutoDock Vina przeprowadzono z użyciem jednokierunkowej ANOVY z testem wielokrotnych porównań Dunnetta (wynik funkcji oryginalnej jako „kontrola”).

W Tabelach 16–27 przedstawiono dziesięć najlepszych (górných pozycji) funkcji wytrenowanych w danym scenariuszu (T1–T3) w porównaniu do funkcji oryginalnej AutoDock Vina i walidowanych według danego scenariusza (W1–W4).

Parametry statystyczne dla trenowania według scenariusza T1 (gdzie jako pozę wiążącą ligandów wybrano pozę o najniższej energii ze zbioru póz z dokowania do wszystkich struktur NK1R) przedstawiono w Tabelach 16–19. Poprawa skuteczności nowych funkcji (wzrost wartości logAUC dla najlepszych funkcji o około 3.0 jednostki; około 50%) względem oryginalnej funkcji oceniającej jest zauważalna w scenariuszach walidacji W1 oraz W2, w których jako zestaw walidacyjny używano wylosowanego podzbioru ze Zbioru II. Dla progu odcięcia $p_{\text{Aff}} = 6.0$ wartości logAUC był nieznacznie wyższe. Oryginalna funkcja AutoDock Vina zajmowała stosunkowo niskie pozycje w rankingu logAUC (odpowiednio 339 i 236). Mimo wyraźnej (w skali względnej) poprawy parametru logAUC, wartości metryki dla najlepszych funkcji przypadają jednak w kategorii „słaby” według kryterium Jaiteha. Nieznaczną poprawę względem oryginalnej funkcji AutoDock Vina (wzrost wartości logAUC o około 2.0 jednostki, czyli około 15%) zaobserwować można w przypadku scenariuszy walidacji W3 oraz W4, w których jako zestaw walidacyjny używano wylosowanych podzbiorów ze Zbiorów II i III. W tych scenariuszach oryginalna funkcja zajmuje 144 i 120 pozycję w rankingu logAUC. W scenariuszu T1W3, najlepsze funkcje są na granicy kategorii „dobry/bardzo dobry” (według kryterium Jaiteha), zaś w T1W4 są w kategorii „dobry”. Trzy funkcje z rodziny T1 (NF-T1-1484, NF-T1-1486 i NF-T1-1219) pojawiły się w górnej dziesiątce rankingu w każdym scenariuszu walidacji (W1–W4).

Parametry statystyczne dla trenowania według scenariusza T2 (w trenowaniu jako pozę wiążącą ligandów wybrano pozę o najniższej energii z dokowania do struktury 6HLO) zamieszczono w Tabelach 20–23. Dla scenariuszy walidacji W1 oraz W2 (walidacja na wylosowanym podzbiórze ze Zbioru II) obserwowany jest wzrost skuteczności nowych funkcji o około 30% (wzrost wartości logAUC o około 2.0-3.0 jednostki) względem funkcji oryginalnej, co wystarcza do osiągnięcia kategorii „dobry” według kryterium Jaiteha. W obu tych scenariuszach oryginalna funkcja zajmuje 176 pozycję w rankingu logAUC (kategoria „słaby” według Jaiteha). Podobnie, jak w wyżej opisanym przypadku testowania nowych funkcji w scenariuszach walidacji W3 i W4, zauważyć można mniejszą względną poprawę skuteczności nowych funkcji (wzrost wartości logAUC o około 15%, czyli o około 2.5 jednostki) względem funkcji oryginalnej. Nie skutkuje to zmianą kategorii według kryterium Jaiteha, nawet dla najlepszych funkcji. W scenariuszach tych oryginalna funkcja zajmuje odpowiednio 178 (kategoria „bardzo dobry” według Jaiteha) i 156 (kategoria „dobry” według

Jaiteha) pozycję w rankingu logAUC. Dwie funkcje z rodziny T2 (NF-T2-1136 oraz NF-T2-1462) pojawiły się w górnej dziesiątce rankingu w każdym scenariuszu walidacji (W1–W4).

Parametry statystyczne dla trenowania według scenariusza T3 (w trenowaniu jako pozę wiążącą ligandów wybrano pozę o najniższej energii z dokowania do struktury 6HLL) przedstawiono w Tabelach 24–27. Umiarkowaną poprawę skuteczności nowo opracowanych funkcji względem oryginalnej funkcji można zaobserwować w scenariuszu W1 (wzrost wartości logAUC o około 23%; o 2.0 jednostki). W przypadku najlepszych funkcji skutkuje to zmianą kategorii według kryterium Jaiteha z „słaby” na „niezły”. W przypadku scenariusza walidacji W2 tylko dla dwóch pierwszych funkcji w rankingu logAUC zaobserwowano statystycznie istotną różnicę w porównaniu do oryginalnej funkcji VINY (NF-T3-1283, NF-T3-1285; poprawa logAUC o około 1.5 jednostki). Dla walidacji w scenariuszu W3 i W4 maksymalny wzrost wartości logAUC względem wartości dla oryginalnej funkcji wynosi około 11% (około 1.5-2.0 jednostek). Dla tych walidacji jedynie pierwsze osiem (W3) lub pierwsze sześć (W4) nowych funkcji uszeregowanych w rankingu logAUC różni się istotnie statystycznie od funkcji oryginalnej. W T3W3 zarówno funkcja oryginalna, jak i najlepsze funkcje trenowane dają wyniki należące do kategorii „bardzo dobry” według kryterium Jaiteha. Podobnie w T3W4, zarówno funkcja wyjściowa jak i funkcje modyfikowane można określić jako „dobre” według kryterium Jaiteha.

Wyniki te wskazują, że otrzymane w trenowaniu T3 funkcje w większości przypadków w poszczególnych scenariuszach walidacji zbliżone są skutecznością do funkcji oryginalnej (oryginalna funkcja zajmuje odpowiednio pozycję: W1 – 113, W2 – 14, W3 – 33, W4 – 20 w rankingu logAUC).

Dwie funkcje z rodziny T3 (NF-T3-1283 i NF-T3-1285), we wszystkich scenariuszach walidacji, zajmują dwa pierwsze miejsca rankingowe.

Tabela 15. Scenariusze trenowania i walidacji funkcji rescoreingowych (trenowanie z wykorzystaniem wyłącznie Zestawu I).

KOD	Scenariusz trenowania	Wybór pozy wiążącej ligandów	Zestaw danych do trenowania	Scenariusz walidacji	Zestawy danych do walidacji	Próg odcięcia pAff aktywny-nieaktywny	Wyniki w Tabeli
T1W1	T1	Poza o najniższej energii ze zbioru póż z dokowania do wszystkich struktur NK1R	Zbiór I	W1	Wylosowane podzbiory Zbioru II	6.0	16
T1W2				W2		6.5	17
T1W3				W3	Wylosowane podzbiory Zbiorów II + III	6.0	18
T1W4				W4		6.5	19
T2W1	T2	Poza o najniższej energii z dokowania do struktury 6HLO		W1	Wylosowane podzbiory Zbioru II	6.0	20
T2W2				W2		6.5	21
T2W3				W3	Wylosowane podzbiory Zbiorów II + III	6.0	22
T2W4				W4		6.5	23
T3W1	T3	Poza o najniższej energii z dokowania do struktury 6HLL		W1	Wylosowane podzbiory Zbiorów II	6.0	24
T3W2				W2		6.5	25
T3W3				W3	Wylosowane podzbiory Zbiorów II + III	6.0	26
T3W4				W4		6.5	27

¹ pAff – ujemny logarytm powinowactwa eksperymentalnego (próg odcięcia pAff = 6.0 oznacza, że związki o aktywności pAff < 6.0 (1 μM) uznawane są za nieaktywne)

² 500 podzbiorów po 410 ligandów (w tym: 10/400 aktywnych/nieaktywnych)

Tabela 16. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W1 (T1W1). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T1-1486	9.5 ± 5.2	2.2	-0.03337	-0.00276	0.60640	-0.00669	-0.05064	0.05204	<p>scenariusz T1W1</p> <p>logAUC</p> <p>funkcja</p> <p>czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T1-1484	9.4 ± 5.3	2.1	-0.02931	-0.00308	0.60640	-0.00718	-0.05064	0.05204	
3.	NF-T1-1361	9.2 ± 5.4	2.1	-0.03034	-0.00325	0.74279	-0.01327	-0.06249	0.06053	
4.	NF-T1-1219	9.2 ± 5.0	2.1	-0.03259	-0.00261	0.64492	-0.00712	-0.08096	0.04395	
5.	NF-T1-1292	9.1 ± 5.1	2.2	-0.03428	-0.00294	0.76324	-0.00710	-0.11789	0.05556	
6.	NF-T1-1499	9.1 ± 5.0	2.1	-0.02412	-0.00336	0.63957	-0.00494	-0.06440	0.04681	
7.	NF-T1-1359	9.1 ± 5.2	2.1	-0.02635	-0.00325	0.74279	-0.01327	-0.05242	0.06053	
8.	NF-T1-1032	9.0 ± 5.0	2.0	-0.03116	-0.00307	0.78408	-0.00612	-0.11831	0.04940	
9.	NF-T1-1037	9.0 ± 5.1	2.1	-0.03177	-0.00343	0.68667	-0.00076	-0.03535	0.05466	
10.	NF-T1-1294	9.0 ± 5.1	2.2	-0.03380	-0.00252	0.72000	-0.00890	-0.16738	0.05556	
339.	VINA	6.3 ± 4.6	5.7	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.,05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T1-1486	NF-T1-1484	NF-T1-1361	NF-T1-1219	NF-T1-1292	NF-T1-1499	NF-T1-1359	NF-T1-1032	NF-T1-1037	NF-T1-1294
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 17. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W2 (T1W2). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T1-1485	8.3 ± 5.1	2.1	-0.03282	-0.00286	0.63402	-0.00327	-0.04338	0.03624	<p>scenariusz T1W2</p> <p>logAUC</p> <p>funkcja</p> <p>czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T1-1486	8.3 ± 5.2	2.2	-0.03337	-0.00276	0.60640	-0.00669	-0.05064	0.05204	
3.	NF-T1-1483	8.2 ± 5.1	2.1	-0.02836	-0.00292	0.61409	-0.00255	-0.03355	0.03624	
4.	NF-T1-1063	8.2 ± 5.0	2.1	-0.02792	-0.00306	0.64352	-0.00122	-0.04314	0.03671	
5.	NF-T1-1219	8.2 ± 5.1	2.1	-0.03259	-0.00261	0.64492	-0.00712	-0.08096	0.04395	
6.	NF-T1-1484	8.1 ± 5.2	2.1	-0.02931	-0.00308	0.60640	-0.00718	-0.05064	0.05204	
7.	NF-T1-1380	8.1 ± 5.0	2.1	-0.03284	-0.00282	0.79120	-0.00064	-0.02987	0.03268	
8.	NF-T1-1037	8.1 ± 5.0	2.1	-0.03177	-0.00343	0.68667	-0.00076	-0.03535	0.05466	
9.	NF-T1-1384	8.1 ± 5.0	2.0	-0.03284	-0.00247	0.82042	-0.00082	-0.06470	0.03268	
10.	NF-T1-1499	8.0 ± 5.1	2.1	-0.02412	-0.00336	0.63957	-0.00494	-0.06440	0.04681	
236.	VINA	5.5 ± 4.8	5.7	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T1-1485	NF-T1-1486	NF-T1-1483	NF-T1-1063	NF-T1-1219	NF-T1-1484	NF-T1-1380	NF-T1-1037	NF-T1-1384	NF-T1-1499
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 18. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W3 (T1W3). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T1-1038	20.0 ± 5.7	2.2	-0.02914	-0.00221	0.68408	-0.01840	-0.17410	0.05118	<p style="text-align: center;">scenariusz T1W3</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerwona linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T1-1162	20.0 ± 5.6	2.2	-0.02801	-0.00244	0.58991	-0.01395	-0.16544	0.04604	
3.	NF-T1-1219	19.8 ± 5.7	2.1	-0.03259	-0.00261	0.64492	-0.00712	-0.08096	0.04395	
4.	NF-T1-1486	19.7 ± 5.8	2.2	-0.03337	-0.00276	0.60640	-0.00669	-0.05064	0.05204	
5.	NF-T1-1484	19.4 ± 5.7	2.2	-0.02931	-0.00308	0.60640	-0.00718	-0.05064	0.05204	
6.	NF-T1-1470	19.3 ± 5.6	2.2	-0.01351	-0.00232	0.38148	-0.02841	-0.33539	0.04829	
7.	NF-T1-1485	19.3 ± 5.7	2.2	-0.03282	-0.00286	0.63402	-0.00327	-0.04338	0.03624	
8.	NF-T1-1214	19.3 ± 5.5	2.2	-0.01598	-0.00320	0.49843	-0.01181	-0.12056	0.04328	
9.	NF-T1-1120	19.2 ± 5.3	2.1	-0.02167	-0.00223	0.81669	-0.01742	-0.00573	0.02216	
10.	NF-T1-1131	19.1 ± 5.6	2.2	-0.02020	-0.00291	0.69428	-0.01918	-0.23950	0.05393	
140.	VINA	17.4 ± 5.7	5.7	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T1-1038	NF-T1-1162	NF-T1-1219	NF-T1-1486	NF-T1-1484	NF-T1-1470	NF-T1-1485	NF-T1-1214	NF-T1-1120	NF-T1-1131
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 19. Najlepsze funkcje wytrenowane w scenariuszu T1 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W4 (T1W4). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T1-1485	15.9 ± 5.6	2.2	-0.03282	-0.00286	0.63402	-0.00327	-0.04338	0.03624	<p style="text-align: center;">scenariusz T1W4</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czzerwona linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T1-1219	15.9 ± 5.5	2.1	-0.03259	-0.00261	0.64492	-0.00712	-0.08096	0.04395	
3.	NF-T1-1162	15.9 ± 5.3	2.2	-0.02801	-0.00244	0.58991	-0.01395	-0.16544	0.04604	
4.	NF-T1-1486	15.7 ± 5.6	2.2	-0.03337	-0.00276	0.60640	-0.00669	-0.05064	0.05204	
5.	NF-T1-1483	15.6 ± 5.5	2.1	-0.02836	-0.00292	0.61409	-0.00255	-0.03355	0.03624	
6.	NF-T1-1229	15.6 ± 5.6	2.1	-0.03851	-0.00221	0.71683	-0.00169	-0.04632	0.02976	
7.	NF-T1-1094	15.5 ± 5.5	2.2	-0.02811	-0.00263	0.35449	-0.00544	-0.15728	0.03905	
8.	NF-T1-1038	15.5 ± 5.3	2.2	-0.02914	-0.00221	0.68408	-0.01840	-0.17410	0.05118	
9.	NF-T1-1484	15.5 ± 5.6	2.2	-0.02931	-0.00308	0.60640	-0.00718	-0.05064	0.05204	
10.	NF-T1-1214	15.4 ± 5.3	2.2	-0.01598	-0.00320	0.49843	-0.01181	-0.12056	0.04328	
122.	VINA	14.0 ± 5.2	5.7	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T1-1485	NF-T1-1219	NF-T1-1162	NF-T1-1486	NF-T1-1483	NF-T1-1229	NF-T1-1094	NF-T1-1038	NF-T1-1484	NF-T1-1214
VINA	****	****	****	****	****	****	****	****	***	***

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 20. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W1 (T2W1). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T2-1136	12.6 ± 6.0	2.3	-0.03107	-0.00246	0.77437	-0.02239	-0.31375	0.05472	<p style="text-align: center;">scenariusz T2W1</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czernowa linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T2-1342	12.2 ± 5.7	2.2	-0.02899	-0.00351	0.66829	-0.01180	-0.16138	0.05752	
3.	NF-T2-1123	12.1 ± 5.7	2.2	-0.02527	-0.00351	0.75289	-0.01291	-0.25277	0.05844	
4.	NF-T2-1464	12.1 ± 6.2	2.3	-0.02335	-0.00302	0.95409	-0.02008	-0.40268	0.04618	
5.	NF-T2-1465	12.1 ± 5.9	2.3	-0.03535	-0.00266	0.83419	-0.02105	-0.42026	0.05976	
6.	NF-T2-1111	12.1 ± 6.0	2.3	-0.02458	-0.00295	0.71132	-0.02797	-0.33116	0.06638	
7.	NF-T2-1462	12.1 ± 6.2	2.3	-0.02335	-0.00302	0.94170	-0.02008	-0.40268	0.04618	
8.	NF-T2-1113	12.1 ± 5.9	2.3	-0.02458	-0.00294	0.68807	-0.02684	-0.33116	0.06638	
9.	NF-T2-1035	12.0 ± 6.0	2.2	-0.03058	-0.00332	1.06143	-0.01141	-0.38519	0.04858	
10.	NF-T2-1270	12.0 ± 5.9	2.3	-0.02696	-0.00326	0.87365	-0.02043	-0.42618	0.05620	
176.	VINA	9.4 ± 5.0	4.5	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T2-1136	NF-T2-1342	NF-T2-1123	NF-T2-1464	NF-T2-1465	NF-T2-1111	NF-T2-1462	NF-T2-1113	NF-T2-1035	NF-T2-1270
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 21. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W2 (T2W2). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T2-1035	10.2 ± 5.9	2.2	-0.03058	-0.00332	1.06143	-0.01141	-0.38519	0.04858	<p style="text-align: center;">scenariusz T2W2</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerwona linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T2-1342	10.0 ± 5.4	2.2	-0.02899	-0.00351	0.66829	-0.01180	-0.16138	0.05752	
3.	NF-T2-1464	9.9 ± 5.9	2.3	-0.02335	-0.00302	0.95409	-0.02008	-0.40268	0.04618	
4.	NF-T2-1123	9.9 ± 5.4	2.2	-0.02527	-0.00351	0.75289	-0.01291	-0.25277	0.05844	
5.	NF-T2-1462	9.9 ± 5.9	2.3	-0.02335	-0.00302	0.94170	-0.02008	-0.40268	0.04618	
6.	NF-T2-1136	9.9 ± 5.6	2.3	-0.03107	-0.00246	0.77437	-0.02239	-0.31375	0.05472	
7.	NF-T2-1122	9.7 ± 5.6	2.2	-0.04572	-0.00194	0.80105	-0.00919	-0.30539	0.04674	
8.	NF-T2-1019	9.7 ± 5.9	2.2	-0.01662	-0.00324	0.77694	-0.01455	-0.23422	0.03341	
9.	NF-T2-1272	9.6 ± 5.7	2.3	-0.02608	-0.00311	0.87365	-0.01947	-0.43110	0.05620	
10.	NF-T2-1270	9.6 ± 5.6	2.3	-0.02696	-0.00326	0.87365	-0.02043	-0.42618	0.05620	
176.	VINA	8.0 ± 5.2	4.5	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T2-1035	NF-T2-1342	NF-T2-1464	NF-T2-1123	NF-T2-1462	NF-T2-1136	NF-T2-1122	NF-T2-1019	NF-T2-1272	NF-T2-1270
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 22. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W3 (T2W3). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T2-1136	23.7 ± 6.2	2.3	-0.03107	-0.00246	0.77437	-0.02239	-0.31375	0.05472	<p style="text-align: center;">scenariusz T2W3</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czarna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T2-1319	23.1 ± 5.9	2.2	-0.05966	-0.00137	0.68582	-0.01856	-0.31237	0.06138	
3.	NF-T2-1374	23.0 ± 6.2	2.3	-0.02264	-0.00289	0.77778	-0.02866	-0.44487	0.05672	
4.	NF-T2-1376	23.0 ± 6.2	2.3	-0.02306	-0.00289	0.77778	-0.02866	-0.44487	0.05672	
5.	NF-T2-1110	23.0 ± 6.0	2.3	-0.03448	-0.00200	0.76107	-0.03053	-0.44546	0.05724	
6.	NF-T2-1318	23.0 ± 5.9	2.2	-0.05815	-0.00142	0.63715	-0.01856	-0.31237	0.06138	
7.	NF-T2-1465	22.9 ± 6.2	2.3	-0.03535	-0.00266	0.83419	-0.02105	-0.42026	0.05976	
8.	NF-T2-1462	22.9 ± 6.3	2.3	-0.02335	-0.00302	0.94170	-0.02008	-0.40268	0.04618	
9.	NF-T2-1342	22.9 ± 6.1	2.2	-0.02899	-0.00351	0.66829	-0.01180	-0.16138	0.05752	
10.	NF-T2-1111	22.9 ± 6.3	2.3	-0.02458	-0.00295	0.71132	-0.02797	-0.33116	0.06638	
178.	VINA	20.5 ± 5.8	4.5	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T2-1136	NF-T2-1319	NF-T2-1374	NF-T2-1376	NF-T2-1110	NF-T2-1318	NF-T2-1465	NF-T2-1462	NF-T2-1342	NF-T2-1111
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 23. Najlepsze funkcje wytrenowane w scenariuszu T2 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W4 (T2W4). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T2-1464	19.2 ± 6.4	2.3	-0.02335	-0.00302	0.95409	-0.02008	-0.40268	0.04618	<p style="text-align: center;">scenariusz T2W4</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerwona linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T2-1035	19.2 ± 6.5	2.2	-0.03058	-0.00332	1.06143	-0.01141	-0.38519	0.04858	
3.	NF-T2-1462	19.2 ± 6.4	2.3	-0.02335	-0.00302	0.94170	-0.02008	-0.40268	0.04618	
4.	NF-T2-1019	19.1 ± 6.3	2.2	-0.01662	-0.00324	0.77694	-0.01455	-0.23422	0.03341	
5.	NF-T2-1017	18.9 ± 6.3	2.2	-0.01662	-0.00350	0.77694	-0.01455	-0.23422	0.03341	
6.	NF-T2-1381	18.8 ± 6.4	2.2	-0.02491	-0.00307	0.95804	-0.01523	-0.41114	0.03762	
7.	NF-T2-1196	18.8 ± 6.3	2.2	-0.01603	-0.00326	0.80562	-0.01518	-0.34034	0.03831	
8.	NF-T2-1136	18.7 ± 6.0	2.3	-0.03107	-0.00246	0.77437	-0.02239	-0.31375	0.05472	
9.	NF-T2-1122	18.6 ± 6.2	2.2	-0.04572	-0.00194	0.80105	-0.00919	-0.30539	0.04674	
10.	NF-T2-1194	18.6 ± 6.2	2.2	-0.01604	-0.00376	0.80562	-0.01518	-0.34034	0.03831	
156.	VINA	16.5 ± 5.8	4.5	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T2-1464	NF-T2-1035	NF-T2-1462	NF-T2-1019	NF-T2-1017	NF-T2-1381	NF-T2-1196	NF-T2-1136	NF-T2-1122	NF-T2-1194
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 24. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W1 (T3W1). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T3-1285	10.4 ± 6.3	2.2	-0.02931	-0.00303	0.44301	-0.01393	-0.21636	0.05742	<p style="text-align: center;">scenariusz T3W1</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T3-1283	10.2 ± 6.2	2.2	-0.02931	-0.00303	0.38336	-0.01393	-0.18805	0.05742	
3.	NF-T3-1503	10.2 ± 6.2	2.2	-0.01766	-0.00333	0.59580	-0.02104	-0.34567	0.04967	
4.	NF-T3-1290	10.1 ± 6.4	2.3	-0.03350	-0.00304	0.87107	-0.02781	-0.49960	0.07140	
5.	NF-T3-1288	10.1 ± 6.4	2.3	-0.03350	-0.00301	0.90475	-0.02802	-0.50644	0.07140	
6.	NF-T3-1289	10.1 ± 6.5	2.3	-0.03156	-0.00310	0.84773	-0.02868	-0.59619	0.07153	
7.	NF-T3-1287	10.0 ± 6.4	2.3	-0.03269	-0.00303	0.89372	-0.02868	-0.59619	0.07153	
8.	NF-T3-1320	9.9 ± 6.4	2.2	-0.02937	-0.00299	0.78565	-0.01614	-0.49044	0.05327	
9.	NF-T3-1030	9.9 ± 6.4	2.2	-0.03270	-0.00317	0.97442	-0.01920	-0.53489	0.06264	
10.	NF-T3-1318	9.9 ± 6.4	2.2	-0.02937	-0.00299	0.78565	-0.01605	-0.49044	0.05327	
113.	VINA	8.4 ± 5.4	4.6	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T3-1285	NF-T3-1283	NF-T3-1503	NF-T3-1290	NF-T3-1288	NF-T3-1289	NF-T3-1287	NF-T3-1320	NF-T3-1030	NF-T3-1318
VINA	****	****	***	***	***	***	***	**	**	**

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 25. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W2 (T3W2). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC	
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆		
1.	NF-T3-1283	8.7 ± 6.1	2.2	-0.02931	-0.00303	0.38336	-0.01393	-0.18805	0.05742	<p style="text-align: center;">scenariusz T3W2</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerowna linia – mediana, VINA – funkcja oryginalna</p>	
2.	NF-T3-1285	8.6 ± 6.1	2.2	-0.02931	-0.00303	0.44301	-0.01393	-0.21636	0.05742		
3.	NF-T3-1503	8.0 ± 5.8	2.2	-0.01766	-0.00333	0.59580	-0.02104	-0.34567	0.04967		
4.	NF-T3-1170	7.9 ± 5.8	2.1	-0.02064	-0.00349	0.49510	-0.00971	-0.35383	0.04801		
5.	NF-T3-1268	7.9 ± 5.9	2.2	-0.03586	-0.00351	0.72224	-0.00557	-0.31248	0.06015		
6.	NF-T3-1479	7.8 ± 5.6	2.1	-0.02195	-0.00290	0.45594	-0.01481	-0.29727	0.04290		
7.	NF-T3-1270	7.7 ± 5.8	2.2	-0.03178	-0.00335	0.64460	-0.00518	-0.35203	0.06015		
8.	NF-T3-1482	7.7 ± 5.6	2.2	-0.02360	-0.00272	0.45594	-0.01481	-0.29727	0.04290		
9.	NF-T3-1460	7.7 ± 5.7	2.2	-0.01939	-0.00399	0.70005	-0.01141	-0.31126	0.05604		
10.	NF-T3-1402	7.5 ± 5.7	2.2	-0.03680	-0.00322	0.76274	-0.00726	-0.44569	0.05757		
14.	VINA	7.2 ± 5.7	4.6	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846		
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>											
vs	NF-T3-1283	NF-T3-1285	NF-T3-1503	NF-T3-1170	NF-T3-1268	NF-T3-1479	NF-T3-1270	NF-T3-1482	NF-T3-1460		NF-T3-1402
VINA	***	**	ns	ns	ns	ns	ns	ns	ns		ns

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 26. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W3 (T3W3). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC	
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆		
1.	NF-T3-1285	23.2 ± 6.8	2.2	-0.02931	-0.00303	0.44301	-0.01393	-0.21636	0.05742	<p style="text-align: center;">scenariusz T3W3</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerowna linia – mediana, VINA – funkcja oryginalna</p>	
2.	NF-T3-1283	23.1 ± 6.8	2.2	-0.02931	-0.00303	0.38336	-0.01393	-0.18805	0.05742		
3.	NF-T3-1503	23.0 ± 6.3	2.2	-0.01766	-0.00333	0.59580	-0.02104	-0.34567	0.04967		
4.	NF-T3-1482	22.3 ± 6.3	2.3	-0.02360	-0.00272	0.45594	-0.01481	-0.29727	0.04290		
5.	NF-T3-1479	22.3 ± 6.3	2.1	-0.02195	-0.00290	0.45594	-0.01481	-0.29727	0.04290		
6.	NF-T3-1320	22.0 ± 6.5	2.2	-0.02937	-0.00299	0.78565	-0.01614	-0.49044	0.05327		
7.	NF-T3-1318	22.0 ± 6.5	2.2	-0.02937	-0.00299	0.78565	-0.01605	-0.49044	0.05327		
8.	NF-T3-1290	22.0 ± 6.7	2.3	-0.03350	-0.00304	0.87107	-0.02781	-0.49960	0.07140		
9.	NF-T3-1425	21.9 ± 6.5	2.2	-0.02444	-0.00298	0.66867	-0.01889	-0.53801	0.05504		
10.	NF-T3-1288	21.8 ± 6.6	2.3	-0.03350	-0.00301	0.90475	-0.02802	-0.50644	0.07140		
33.	VINA	20.9 ± 6.1	4.6	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846		
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>											
vs	NF-T3-1285	NF-T3-1283	NF-T3-1503	NF-T3-1482	NF-T3-1479	NF-T3-1320	NF-T3-1318	NF-T3-1290	NF-T3-1425		NF-T3-1288
VINA	****	****	****	**	**	*	*	*	ns	ns	

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 27. Najlepsze funkcje wytrenowane w scenariuszu T3 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W4 (T3W4). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T3-1283	18.8 ± 6.6	2.2	-0.02931	-0.00303	0.38336	-0.01393	-0.18805	0.05742	<p style="text-align: center;">scenariusz T3W4</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T3-1285	18.7 ± 6.6	2.2	-0.02931	-0.00303	0.44301	-0.01393	-0.21636	0.05742	
3.	NF-T3-1479	18.4 ± 6.2	2.1	-0.02195	-0.00290	0.45594	-0.01481	-0.29727	0.04290	
4.	NF-T3-1482	18.3 ± 6.2	2.2	-0.02360	-0.00272	0.45594	-0.01481	-0.29727	0.04290	
5.	NF-T3-1503	18.3 ± 6.5	2.2	-0.01766	-0.00333	0.59580	-0.02104	-0.34567	0.04967	
6.	NF-T3-1170	18.2 ± 6.5	2.1	-0.02064	-0.00349	0.49510	-0.00971	-0.35383	0.04801	
7.	NF-T3-1268	17.7 ± 6.7	2.2	-0.03586	-0.00351	0.72224	-0.00557	-0.31248	0.06015	
8.	NF-T3-1142	17.6 ± 5.9	2.1	-0.00756	-0.00358	0.31057	-0.01292	-0.24103	0.03453	
9.	NF-T3-1144	17.6 ± 6.0	2.1	-0.00756	-0.00378	0.31057	-0.01156	-0.24103	0.03453	
10.	NF-T3-1460	17.6 ± 6.6	2.2	-0.01939	-0.00399	0.70005	-0.01141	-0.31126	0.05604	
20.	VINA	17.0 ± 6.3	4.6	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 10 związków aktywnych i 10*40 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T3-1283	NF-T3-1285	NF-T3-1479	NF-T3-1482	NF-T3-1503	NF-T3-1170	NF-T3-1268	NF-T3-1142	NF-T3-1144	NF-T3-1460
VINA	***	***	**	**	*	*	ns	ns	ns	ns

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Początkowo planowano wykorzystać jako zestaw treningowy tylko Zbiór I (czyli w większości związki zsyntezowane przeze mnie w ramach projektu doktoranckiego; scenariusze T1–T3). Walidacja nowootrzymanyh funkcji pokazała, że trenowanie na tym zbiorze daje systematyczną poprawę wartości logAUC względem wartości dla oryginalnej funkcji VINY. Niemniej jednak poprawa ta jest umiarkowana (rzędu kilku jednostek logAUC), a wartości logAUC wynoszące około 20 („dobre” lub „bardzo dobre” według Jaiteha) w „łatwiejszych” walidacjach lub około 5–10 („słabe” lub „niezłe” według Jaiteha) w „trudniejszych” scenariuszach wydają się jedynie umiarkowanie satysfakcjonujące. W związku z tym postanowiłam opracować dodatkowe funkcje na rozszerzonym (większym i bardziej zróżnicowanym) zbiorze treningowym. W scenariuszach T4–T6 zestaw do trenowania stanowiło wylosowane 25% połączonych Zbiorów I i II. Walidacje przeprowadzono według różnych scenariuszy (W5–W8), opisanych w Tabeli 28, uwzględniając różne zestawy walidacyjne oraz progi odcięcia pAff aktywny/nieaktywny. Zbiory testowe były randomizowane, w każdym przypadku wygenerowano po pięćset podzbiorów o proporcji 7/210 aktywnych/nieaktywnych związków. Podobnie jak dla trenowania w scenariuszach T1–T3, miarą sukcesu z użyciem nowych funkcji były otrzymane wartości logAUC w przesiewie wirtualnym, a porównanie z wartościami dla oryginalnej funkcji AutoDock Vina przeprowadzono z użyciem jednokierunkowej ANOVY z testem wielokrotnych porównań Dunnetta (wynik funkcji oryginalnej jako „kontrola”).

Parametry statystyczne dla trenowania według scenariusza T4 (gdzie jako pozę wiążącą ligandów wybrano pozę o najniższej energii ze zbioru póz z dokowania do wszystkich struktur NK1R) przedstawiono w Tabelach 29–32. Znaczącą poprawę skuteczności nowych funkcji można zaobserwować w scenariuszach walidacyjnych W5 i W6, w których jako zbiór walidacyjny stosowano wylosowane podzbiory z 75% Zbiorów I+II. W scenariuszu W5, wartości logAUC pierwszych dziesięciu funkcji w rankingu logAUC w obu przypadkach są wyższe o około 110–120% od kontroli (5.0–6.0 jednostek logAUC; poprawa kategorii z „słaby” na „niezły” według kryterium Jaiteha). Dla progu odcięcia pAff = 6.5 (W6) wartości logAUC są około 2 razy niższe niż dla pAff = 6.0 (W5), a najlepsze funkcje mają bardzo niską wartość logAUC (logAUC = 4.8; wszystkie funkcje w kategorii „słaby” według Jaiteha). Niemniej jednak tutaj również obserwuje się pewną poprawę względem funkcji oryginalnej, dla której wartość logAUC = 2.2. Oryginalna funkcja umieszczona jest na 601 pozycji w rankingu logAUC. Dla walidacji w scenariuszach W7 i W8, w których jako zestaw walidacyjny stosowano wylosowane podzbiory z 75% Zbiorów I+II oraz Zbioru III, wzrost wartości logAUC względem funkcji oryginalnej wynosi około 25% (około 3.0–4.0

jednostek), co pozwala na poprawę kategorii najlepszych funkcji (w porównaniu do funkcji wyjściowej) z „dobry” na „bardzo dobry” (W7) i z „niezły” na „dobry” (W8). W tych scenariuszach oryginalna funkcja zajmuje odpowiednio 601 i 600 pozycję w rankingu logAUC. Cztery funkcje z rodziny T4 (NF-T4-1055, NF-T4-1385, NF-T4-1535 i NF-T4-1537) trzykrotnie pojawiają się w górnej dziesiątce rankingów logAUC różnych scenariuszy walidacji.

Parametry statystyczne dla trenowania według scenariusza T5 (gdzie jako pozę wiążącą ligandów wybrano pozę o najniższej energii z dokowania do struktury 6HLO) zamieszczono w Tabelach 33-36. Dla walidacji w scenariuszach W5 i W6 (jako zbiór walidacyjny stosowano wylosowane podzbiory z wylosowanego 75% Zbiorów I+II) obserwowany jest wzrost skuteczności nowych funkcji o około 40% (wzrost wartości logAUC) względem funkcji oryginalnej. Przy czym lepsze wyniki uzyskano dla $p_{\text{Aff}} = 6.0$ (W5; wzrost 44–47%; ponad 3.0 jednostki; poprawa z kategorii z „słaby” do „niezły” według kryterium Jaiteha) niż dla $p_{\text{Aff}} = 6.5$ (W6; wzrost o 35–43%; do 2.0 jednostek; wszystkie funkcje w kategorii „słaby” według kryterium Jaiteha). W obu tych scenariuszach oryginalna funkcja zajmuje 601 lub 598 pozycję w rankingu logAUC. W przypadku scenariuszy walidacyjnych W7 i W8 (jako zestaw walidacyjny stosowano wylosowane podzbiory z wylosowanego 75% Zbiorów I+II oraz Zbioru III) poprawa skuteczności nowych funkcji jest umiarkowana (dla W7 wzrost wartości logAUC o około 22–23%, czyli o ponad 4.0 jednostki; dla W8 o około 16–19%, czyli do 3.0 jednostek) względem funkcji oryginalnej. Pomimo poprawy wartości logAUC, najlepsze funkcje utrzymują kategorię tę samą, co funkcja wyjściowa, czyli „bardzo dobry” (W7) lub „dobry” (W8). W scenariuszach tych oryginalna funkcja zajmuje 601 pozycję w rankingu logAUC. Jedna funkcja z rodziny T5 (NF-T5-1239) pojawia się w pierwszej dziesiątce w każdym scenariuszu walidacyjnym.

Parametry statystyczne dla trenowania według scenariusza T6 (gdzie jako pozę wiążącą ligandów wybrano pozę o najniższej energii z dokowania do struktury 6HLL) zamieszczono w Tabelach 37–40. Największą względną poprawę skuteczności można zaobserwować w walidacji według scenariusza W5 – wzrost wartości logAUC o 62–66% (około 4.0 jednostek; przejście z kategorii „słaby” do „niezły” według kryterium Jaiteha) dla pierwszych dziesięciu nowych funkcji w rankingu. Poprawę skuteczności zaobserwowano również w walidacji według scenariusza W6 – wzrost wartości logAUC o 43–51% (o około 3.0 jednostki) dla pierwszych dziesięciu nowych funkcji w rankingu, jednak tutaj nawet najlepsze funkcje utrzymują kategorię „słaby” według kryterium Jaiteha. W obu tych przypadkach jako zbiór walidacyjny stosowano wylosowane podzbiory z 75% Zbiorów I+II, a scenariusze różniły się

progami odcięcia aktywny/nieaktywny (lepsze wyniki dla progów $p_{Aff} = 6.0$). Dla walidacji według scenariuszy W7 i W8 otrzymano nieznaczną poprawę skuteczności względem funkcji oryginalnej – wzrost wartości logAUC średnio o 15-16% (o 3.0-3.5 jednostki) dla najlepszych dziesięciu pozycji w rankingu logAUC. Ponownie, pomimo poprawy wartości logAUC, najlepsze funkcje utrzymują kategorię tę samą, co funkcja wyjściowa, czyli „bardzo dobry” (W7) lub „dobry” (W8). We wszystkich tych scenariuszach Vina znajduje się na dalekim miejscu w rankingu logAUC (601 lub 598 dla W7). Co ciekawe, żadna z funkcji z rodziny T6 nie pojawiła się w górnej dziesiątce rankingów według różnych scenariuszy walidacji więcej niż dwukrotnie.

Podsumowując tę część, można stwierdzić, że w przypadku funkcji trenowanych na rozszerzonym zbiorze (scenariusze T4–T6) ponownie obserwuje się systematyczną poprawę wartości logAUC dla najlepszych funkcji w porównaniu do funkcji oryginalnej VINA. Ponownie również, obserwowana poprawa logAUC, choć w niektórych scenariuszach walidacji sięga nawet 50–100% w porównaniu do oryginalnej funkcji VINA, dotyczy jednak względnie niskich wartości logAUC. Przykładowo, największą (względnie) poprawę skuteczności zaobserwowano dla funkcji z rodziny T5 w scenariuszach walidacji W5 i W6. Obserwowany wzrost wartości logAUC względem oryginalnej funkcji oceniającej AutoDock Vina był na poziomie 110–120% dla pierwszych dziesięciu funkcji w rankingu logAUC. Należy zauważyć, że wartości logAUC są jednak jedynie na poziomie „niezłym” (10.7-11.0) lub „słabym” (4.6-4.8).

Najwyższe (numerycznie) wartości logAUC obserwowano dla funkcji z rodzin T5 i T6 w walidacji według „łatwych” scenariuszy W7 (najwyższa wartość logAUC około 24.5, kategoria „bardzo dobre” według Jaiteha) i W8 (najwyższa wartość logAUC około 19.0; kategoria „dobre” według Jaiteha). We wszystkich przypadkach nie ma statystycznie istotnych różnic między najlepszymi funkcjami z rodziny T5 i T6, lub T2 i T3. W niektórych przypadkach, najlepsze funkcje z rodzin T5 i T6 są skuteczniejsze niż funkcje z rodzin T4 (statystycznie istotnie), a także najlepsze funkcje z rodzin T2 i T3 są skuteczniejsze niż funkcje z rodzin T1 (statystycznie istotnie). Widać więc, że lepszą skuteczność uzyskują funkcje wytrenowane (i stosowane) na wynikach z dokowań do struktur 6HLO i 6HLL (w porównaniu do funkcji wytrenowanych na najniższych wynikach z dokowań do wielu struktur). Bezpośrednie, ścisłe porównanie między funkcjami z rodzin T5 i T6 a funkcjami z rodzin T2 i T3, jest utrudnione z powodu zastosowania innych zbiorów walidacyjnych.

Na podstawie uzyskanych wyników nie da się wyróżnić żadnej pojedynczej funkcji jako „najlepszej”.

Analizując zmiany w wagach funkcji w_1 – w_6 w najlepszych modelach, można poczynić następujące spostrzeżenia:

- wagi w_1 - w_3 odpowiadające za dopasowanie geometryczne zmieniają się w zróżnicowanych zakresach w porównaniu do funkcji oryginalnej,
- waga w_4 odpowiadająca za wkład oddziaływań hydrofobowych jest w zdecydowanej większości wypadków mniej ujemna niż w funkcji oryginalnej,
- waga w_5 odpowiadająca za wkład wiązań wodorowych jest systematycznie mniej ujemna niż w funkcji oryginalnej,
- waga w_6 odpowiadająca za wielkość „kary” za swobodę konformacyjną liganda zmienia się w zróżnicowanych zakresach.

Można to orientacyjnie interpretować w ten sposób, że domyślna funkcja Viny przeszacowuje siłę wiązań wodorowych obecnych w oddziaływaniach ligandów NK1R z receptorem. Dodatkowo, większą rolę niż niespecyficzne kontakty między powierzchniami odgrywa dobre dopasowanie steryczne liganda i receptora.

Tabela 28. Scenariusze trenowania i walidacji funkcji rescoringowych (trenowanie z wykorzystaniem podzbioru Zestawów I i II).

KOD	Scenariusz trenowania	Wybór pozy wiążącej ligandów	Zestaw danych do trenowania	Scenariusz walidacji	Zestawy danych do walidacji	Próg odcięcia pAff aktywny-nieaktywny	Wyniki w Tabeli
T4W5	T4	Poza o najniższej energii ze zbioru póż z dokowania do wszystkich struktur NK1R	Wylosowane 25% Zbiorów I+II	W5	Wylosowane podzbiory z wylosowanego 75% Zbiorów I+II	6.0	29.
T4W6				W6		6.5	30.
T4W7				W7	Wylosowane podzbiory z wylosowanego 75% Zbiorów I+II oraz Zbioru III	6.0	31.
T4W8				W8		6.5	32.
T5W5	T5	Poza o najniższej energii z dokowania do struktury 6HLO		W5	Wylosowane podzbiory z wylosowanego 75% Zbiorów I+II	6.0	33.
T5W6				W6		6.5	34.
T5W7				W7	Wylosowane podzbiory z wylosowanego 75% Zbiorów I+II oraz Zbioru III	6.0	35.
T5W8				W8		6.5	36.
T6W5	T6	Poza o najniższej energii z dokowania do struktury 6HLL		W5	Wylosowane podzbiory z wylosowanego 75% Zbiorów I+II	6.0	37.
T6W6				W6		6.5	38.
T6W7				W7	Wylosowane podzbiory z wylosowanego 75% Zbiorów I+II oraz Zbioru III	6.0	39.
T6W8				W8		6.5	40.

¹ pAff – ujemny logarytm powinowactwa eksperymentalnego (próg odcięcia pAff = 6.0 oznacza, że związki o aktywności pAff < 6.0 (1 uM) uznawane są za nieaktywne)

² 500 podzbiorów po 217 ligandów (w tym: 7/210 aktywnych/nieaktywnych)

Tabela 29. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W5 (T4W5). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T4-1415	11.0 ± 7.2	2.4	-0.06983	-0.00300	0.62335	-0.02100	-0.01836	0.06439	<p style="text-align: center;">scenariusz T4W5</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T4-1475	10.9 ± 7.2	2.4	-0.07966	-0.00234	0.62956	-0.02485	-0.07591	0.06850	
3.	NF-T4-1265	10.9 ± 7.2	2.4	-0.07677	-0.00269	0.66075	-0.02117	-0.04140	0.06652	
4.	NF-T4-1267	10.9 ± 7.3	2.4	-0.07677	-0.00268	0.65957	-0.01729	-0.04151	0.06652	
5.	NF-T4-1325	10.8 ± 7.1	2.4	-0.06726	-0.00314	0.64812	-0.01863	-0.01334	0.05994	
6.	NF-T4-1357	10.8 ± 7.2	2.4	-0.06440	-0.00308	0.51983	-0.02012	-0.02159	0.05693	
7.	NF-T4-1140	10.8 ± 7.2	2.4	-0.06285	-0.00346	0.63369	-0.02224	-0.03761	0.06647	
8.	NF-T4-1082	10.7 ± 7.2	2.4	-0.06306	-0.00314	0.60954	-0.02305	-0.05897	0.06494	
9.	NF-T4-1445	10.7 ± 7.2	2.4	-0.06284	-0.00317	0.60353	-0.02465	-0.03895	0.06341	
10.	NF-T4-1176	10.7 ± 7.2	2.4	-0.06755	-0.00311	0.68206	-0.02501	-0.03677	0.06771	
601.	VINA	5.0 ± 6.2	2.8	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T4-1415	NF-T4-1475	NF-T4-1265	NF-T4-1267	NF-T4-1325	NF-T4-1357	NF-T4-1140	NF-T4-1082	NF-T4-1445	NF-T4-1176
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 30. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W6 (T4W6). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T4-1446	4.8 ± 6.6	2.4	-0.06251	-0.00304	0.54793	-0.01591	-0.04512	0.05105	<p style="text-align: center;">scenariusz T4W6</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T4-1447	4.8 ± 6.6	2.4	-0.06251	-0.00304	0.54755	-0.01591	-0.04564	0.05105	
3.	NF-T4-1449	4.8 ± 6.6	2.4	-0.06306	-0.00303	0.56895	-0.01591	-0.05559	0.05105	
4.	NF-T4-1535	4.6 ± 6.4	2.4	-0.06019	-0.00280	0.60618	-0.02586	-0.01535	0.05216	
5.	NF-T4-1357	4.6 ± 6.5	2.4	-0.06440	-0.00308	0.51983	-0.02012	-0.02159	0.05693	
6.	NF-T4-1537	4.6 ± 6.4	2.4	-0.06019	-0.00266	0.60618	-0.02603	-0.01484	0.05216	
7.	NF-T4-1385	4.6 ± 6.4	2.4	-0.05067	-0.00307	0.55627	-0.02830	-0.02272	0.04911	
8.	NF-T4-1355	4.6 ± 6.5	2.4	-0.06440	-0.00308	0.62352	-0.02012	-0.00333	0.05693	
9.	NF-T4-1055	4.6 ± 6.4	2.4	-0.05971	-0.00308	0.61909	-0.02322	-0.01065	0.05435	
10.	NF-T4-1359	4.6 ± 6.6	2.4	-0.06440	-0.00308	0.50197	-0.01922	-0.05404	0.05693	
601.	VINA	2.2 ± 5.8	2.8	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T4-1446	NF-T4-1447	NF-T4-1449	NF-T4-1535	NF-T4-1357	NF-T4-1537	NF-T4-1385	NF-T4-1355	NF-T4-1055	NF-T4-1359
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 31. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W7 (T4W7). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T4-1385	21.7 ± 7.4	2.4	-0.05067	-0.00307	0.55627	-0.02830	-0.02272	0.04911	<p style="text-align: center;">scenariusz T4W7</p> <p style="text-align: center;">logAUC</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T4-1537	21.7 ± 7.5	2.4	-0.06019	-0.00266	0.60618	-0.02603	-0.01484	0.05216	
3.	NF-T4-1535	21.7 ± 7.5	2.4	-0.06019	-0.00280	0.60618	-0.02586	-0.01535	0.05216	
4.	NF-T4-1055	21.7 ± 7.6	2.4	-0.05971	-0.00308	0.61909	-0.02322	-0.01065	0.05435	
5.	NF-T4-1290	21.7 ± 7.5	2.4	-0.05138	-0.00311	0.54603	-0.02929	-0.04976	0.05339	
6.	NF-T4-1590	21.6 ± 7.6	2.4	-0.05815	-0.00314	0.61171	-0.02509	-0.03346	0.05623	
7.	NF-T4-1387	21.6 ± 7.4	2.4	-0.05067	-0.00302	0.48613	-0.02914	-0.05682	0.04911	
8.	NF-T4-1495	21.6 ± 7.5	2.4	-0.05554	-0.00333	0.61132	-0.02449	-0.03571	0.05617	
9.	NF-T4-1345	21.6 ± 7.6	2.4	-0.06054	-0.00304	0.61770	-0.02455	-0.05232	0.05675	
10.	NF-T4-1200	21.6 ± 7.5	2.4	-0.05349	-0.00310	0.60185	-0.02914	-0.05483	0.05442	
601.	VINA	17.5 ± 6.9	2.8	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p style="text-align: center;">Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p style="text-align: center;">Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T4-1385	NF-T4-1537	NF-T4-1535	NF-T4-1055	NF-T4-1290	NF-T4-1590	NF-T4-1387	NF-T4-1495	NF-T4-1345	NF-T4-1200
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 32. Najlepsze funkcje wytrenowane w scenariuszu T4 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W8 (T4W8). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T4-1385	16.4 ± 6.8	2.4	-0.05067	-0.00307	0.55627	-0.02830	-0.02272	0.04911	<p>scenariusz T4W8</p> <p>logAUC</p> <p>funkcja</p> <p>czerwona linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T4-1537	16.4 ± 6.9	2.4	-0.06019	-0.00266	0.60618	-0.02603	-0.01484	0.05216	
3.	NF-T4-1535	16.3 ± 6.9	2.4	-0.06019	-0.00280	0.60618	-0.02586	-0.01535	0.05216	
4.	NF-T4-1449	16.3 ± 6.9	2.4	-0.06306	-0.00303	0.56895	-0.01591	-0.05559	0.05105	
5.	NF-T4-1447	16.3 ± 6.9	2.4	-0.06251	-0.00304	0.54755	-0.01591	-0.04564	0.05105	
6.	NF-T4-1446	16.3 ± 6.9	2.4	-0.06251	-0.00304	0.54793	-0.01591	-0.04512	0.05105	
7.	NF-T4-1387	16.3 ± 6.8	2.4	-0.05067	-0.00302	0.48613	-0.02914	-0.05682	0.04911	
8.	NF-T4-1111	16.2 ± 6.7	2.4	-0.04759	-0.00326	0.60708	-0.02686	-0.11621	0.04788	
9.	NF-T4-1055	16.2 ± 6.9	2.4	-0.05971	-0.00308	0.61909	-0.02322	-0.01065	0.05435	
10.	NF-T4-1113	16.2 ± 6.7	2.4	-0.04759	-0.00326	0.60708	-0.02734	-0.13859	0.04788	
600.	VINA	13.2 ± 6.1	2.8	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
v/s	NF-T4-1385	NF-T4-1537	NF-T4-1535	NF-T4-1449	NF-T4-1447	NF-T4-1446	NF-T4-1387	NF-T4-1111	NF-T4-1055	NF-T4-1113
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 33. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W5 (T5W5). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T5-1378	10.6 ± 6.7	2.3	-0.05225	-0.00526	1.02051	-0.02271	-0.24449	0.06336	<p style="text-align: center;">scenariusz T5W5</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerwona linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T5-1041	10.6 ± 6.8	2.3	-0.03844	-0.00571	0.90021	-0.02865	-0.10792	0.06592	
3.	NF-T5-1403	10.6 ± 6.8	2.3	-0.02345	-0.00629	0.97914	-0.03278	-0.28162	0.06141	
4.	NF-T5-1148	10.5 ± 6.5	2.3	-0.01449	-0.00689	0.93789	-0.02786	-0.26661	0.05897	
5.	NF-T5-1051	10.5 ± 6.7	2.3	-0.02668	-0.00597	0.93427	-0.02865	-0.27924	0.05556	
6.	NF-T5-1239	10.4 ± 6.6	2.3	-0.03840	-0.00535	0.86691	-0.02774	-0.03114	0.05203	
7.	NF-T5-1568	10.4 ± 6.7	2.3	-0.06275	-0.00533	0.85929	-0.01166	-0.13019	0.07061	
8.	NF-T5-1592	10.4 ± 6.5	2.3	-0.04783	-0.00509	0.73291	-0.01736	-0.12537	0.05423	
9.	NF-T5-1027	10.4 ± 6.7	2.3	-0.01787	-0.00632	0.67094	-0.02988	-0.04213	0.05698	
10.	NF-T5-1497	10.4 ± 6.7	2.3	-0.02879	-0.00609	0.72813	-0.0289	-0.12087	0.06289	
601.	VINA	7.2 ± 6.0	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
v/s	NF-T5-1378	NF-T5-1041	NF-T5-1403	NF-T5-1148	NF-T5-1051	NF-T5-1239	NF-T5-1568	NF-T5-1592	NF-T5-1027	NF-T5-1497
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 34. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W6 (T5W6). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T5-1239	6.9 ± 6.9	2.3	-0.03840	-0.00535	0.86691	-0.02774	-0.03114	0.05203	<p style="text-align: center;">scenariusz T5W6</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czarowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T5-1554	6.7 ± 7.0	2.3	-0.03761	-0.00484	0.75172	-0.02846	-0.09112	0.04809	
3.	NF-T5-1527	6.7 ± 6.8	2.3	-0.02777	-0.00555	0.67602	-0.02282	-0.14989	0.04649	
4.	NF-T5-1592	6.6 ± 6.8	2.3	-0.04783	-0.00509	0.73291	-0.01736	-0.12537	0.05423	
5.	NF-T5-1177	6.6 ± 6.9	2.3	-0.03062	-0.00560	0.72723	-0.02392	-0.13921	0.05272	
6.	NF-T5-1378	6.6 ± 6.9	2.3	-0.05225	-0.00526	1.02051	-0.02271	-0.24449	0.06336	
7.	NF-T5-1593	6.6 ± 6.9	2.3	-0.04783	-0.00509	0.69802	-0.01710	-0.12272	0.05423	
8.	NF-T5-1051	6.5 ± 6.8	2.3	-0.02668	-0.00597	0.93427	-0.02865	-0.27924	0.05560	
9.	NF-T5-1546	6.5 ± 6.9	2.3	-0.03113	-0.00538	0.74841	-0.02682	-0.23746	0.04989	
10.	NF-T5-1539	6.5 ± 6.9	2.3	-0.06044	-0.0044	0.92207	-0.01721	-0.34399	0.05286	
598.	VINA	4.8 ± 6.5	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T5-1239	NF-T5-1554	NF-T5-1527	NF-T5-1592	NF-T5-1177	NF-T5-1378	NF-T5-1593	NF-T5-1051	NF-T5-1546	NF-T5-1539
VINA	****	****	****	***	***	***	***	***	***	***

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 35. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W7 (T5W7). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T5-1239	24.7 ± 7.8	2.3	-0.03840	-0.00535	0.86691	-0.02774	-0.03114	0.05203	<p style="text-align: center;">scenariusz T5W7</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T5-1389	24.6 ± 7.6	2.3	-0.06148	-0.00402	0.52969	-0.02806	-0.02565	0.05804	
3.	NF-T5-1117	24.6 ± 7.7	2.3	-0.04239	-0.00476	0.61672	-0.02851	-0.02750	0.05604	
4.	NF-T5-1048	24.6 ± 7.7	2.3	-0.04161	-0.00487	0.75937	-0.03604	-0.03072	0.06018	
5.	NF-T5-1115	24.6 ± 7.6	2.3	-0.04239	-0.00476	0.52590	-0.02851	-0.00805	0.05604	
6.	NF-T5-1046	24.6 ± 7.7	2.3	-0.03937	-0.00500	0.73259	-0.03604	-0.03072	0.06018	
7.	NF-T5-1118	24.5 ± 7.6	2.3	-0.04399	-0.00476	0.54100	-0.02851	-0.01679	0.05604	
8.	NF-T5-1385	24.5 ± 7.6	2.3	-0.05542	-0.00406	0.48927	-0.02735	-0.01862	0.05804	
9.	NF-T5-1535	24.5 ± 7.6	2.3	-0.04696	-0.00445	0.48794	-0.02485	-0.02269	0.05286	
10.	NF-T5-1310	24.5 ± 7.7	2.3	-0.04998	-0.00448	0.54339	-0.02613	-0.05680	0.05790	
601.	VINA	20.1 ± 7.0	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p style="text-align: center;">Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p style="text-align: center;">Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T5-1239	NF-T5-1389	NF-T5-1117	NF-T5-1048	NF-T5-1115	NF-T5-1046	NF-T5-1118	NF-T5-1385	NF-T5-1535	NF-T5-1310
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 36. Najlepsze funkcje wytrenowane w scenariuszu T5 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W8 (T5W8). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T5-1239	18.7 ± 7.3	2.3	-0.03840	-0.00535	0.86691	-0.02774	-0.03114	0.05203	<p style="text-align: center;">scenariusz T5W8</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T5-1554	18.6 ± 7.2	2.3	-0.03761	-0.00484	0.75172	-0.02846	-0.09112	0.04809	
3.	NF-T5-1539	18.6 ± 7.4	2.3	-0.06044	-0.00440	0.92207	-0.01721	-0.34399	0.05286	
4.	NF-T5-1469	18.4 ± 7.2	2.3	-0.03777	-0.00442	0.69463	-0.03198	-0.22314	0.04503	
5.	NF-T5-1527	18.4 ± 7.3	2.3	-0.02777	-0.00555	0.67602	-0.02282	-0.14989	0.04649	
6.	NF-T5-1546	18.3 ± 7.3	2.3	-0.03113	-0.00538	0.74841	-0.02682	-0.23746	0.04989	
7.	NF-T5-1592	18.3 ± 7.3	2.3	-0.04783	-0.00509	0.73291	-0.01736	-0.12537	0.05423	
8.	NF-T5-1593	18.3 ± 7.3	2.3	-0.04783	-0.00509	0.69802	-0.01710	-0.12272	0.05423	
9.	NF-T5-1566	18.3 ± 7.2	2.3	-0.03067	-0.00461	0.61239	-0.03114	-0.23546	0.04352	
10.	NF-T5-1412	18.3 ± 7.2	2.3	-0.02877	-0.00495	0.86407	-0.03891	-0.30450	0.05001	
601.	VINA	15.7 ± 6.7	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p style="text-align: center;">Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p style="text-align: center;">Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T5-1239	NF-T5-1554	NF-T5-1539	NF-T5-1469	NF-T5-1527	NF-T5-1546	NF-T5-1592	NF-T5-1593	NF-T5-1566	NF-T5-1412
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 37. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W5 (T6W5). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T6-1266	11.3 ± 7.6	2.3	-0.01720	-0.00652	0.34766	-0.01619	-0.04411	0.05482	<p style="text-align: center;">scenariusz T6W5</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T6-1268	11.2 ± 7.6	2.3	-0.01720	-0.00652	0.36004	-0.01619	-0.04411	0.05482	
3.	NF-T6-1495	11.2 ± 7.7	2.3	-0.02831	-0.00602	0.40082	-0.01507	-0.10357	0.05660	
4.	NF-T6-1059	11.2 ± 7.6	2.3	-0.02107	-0.00652	0.38354	-0.01757	-0.13398	0.05704	
5.	NF-T6-1508	11.1 ± 7.5	2.3	-0.02149	-0.00600	0.38699	-0.01569	-0.10317	0.05331	
6.	NF-T6-1595	11.1 ± 7.5	2.3	-0.01669	-0.00656	0.37617	-0.01723	-0.11411	0.05647	
7.	NF-T6-1119	11.1 ± 7.6	2.3	-0.02257	-0.00652	0.40745	-0.01790	-0.16381	0.05874	
8.	NF-T6-1293	11.1 ± 7.5	2.3	-0.01489	-0.00678	0.36627	-0.01741	-0.14706	0.05640	
9.	NF-T6-1355	11.1 ± 7.5	2.3	-0.02089	-0.00614	0.37337	-0.01793	-0.09861	0.05344	
10.	NF-T6-1520	11.0 ± 7.5	2.3	-0.02089	-0.00618	0.37888	-0.01831	-0.12306	0.05475	
601.	VINA	6.8 ± 6.7	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T6-1266	NF-T6-1268	NF-T6-1495	NF-T6-1059	NF-T6-1508	NF-T6-1595	NF-T6-1119	NF-T6-1293	NF-T6-1355	NF-T6-1520
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 38. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W6 (T6W6). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T6-1200	8.6 ± 8.1	2.3	-0.02389	-0.00619	0.42315	-0.00420	-0.02926	0.04237	<p style="text-align: center;">scenariusz T6W6</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czterowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T6-1202	8.5 ± 8.1	2.3	-0.02242	-0.00619	0.42315	-0.00420	-0.02926	0.04237	
3.	NF-T6-1149	8.3 ± 7.9	2.3	-0.01961	-0.00598	0.39762	-0.01412	-0.09523	0.04914	
4.	NF-T6-1495	8.3 ± 8.0	2.3	-0.02831	-0.00602	0.40082	-0.01507	-0.10357	0.05660	
5.	NF-T6-1084	8.3 ± 8.1	2.3	-0.02168	-0.00663	0.52560	-0.00740	-0.19080	0.04858	
6.	NF-T6-1082	8.3 ± 8.2	2.3	-0.02168	-0.00663	0.56268	-0.00740	-0.19080	0.04858	
7.	NF-T6-1508	8.3 ± 8.0	2.3	-0.02149	-0.00600	0.38699	-0.01569	-0.10317	0.05331	
8.	NF-T6-1145	8.2 ± 7.9	2.3	-0.01992	-0.00607	0.35640	-0.01623	-0.07149	0.04914	
9.	NF-T6-1357	8.2 ± 8.0	2.3	-0.03087	-0.00579	0.54429	-0.01747	-0.23276	0.05344	
10.	NF-T6-1499	8.2 ± 8.3	2.3	-0.03585	-0.00591	0.43085	-0.0087	-0.29618	0.05660	
601.	VINA	5.7 ± 6.9	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T6-1200	NF-T6-1202	NF-T6-1149	NF-T6-1495	NF-T6-1084	NF-T6-1082	NF-T6-1508	NF-T6-1145	NF-T6-1357	NF-T6-1499
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 39. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W7 (T6W7). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T6-1473	24.6 ± 8.0	2.3	-0.02956	-0.00522	0.67925	-0.03200	-0.25167	0.05164	<p style="text-align: center;">scenariusz T6W7</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czerowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T6-1253	24.6 ± 8.0	2.3	-0.02425	-0.00535	0.49632	-0.03105	-0.17429	0.05117	
3.	NF-T6-1472	24.4 ± 8.0	2.3	-0.02918	-0.00510	0.45951	-0.02499	-0.20444	0.05164	
4.	NF-T6-1116	24.4 ± 8.2	2.3	-0.04141	-0.00521	0.70126	-0.02244	-0.11218	0.05874	
5.	NF-T6-1479	24.3 ± 8.3	2.3	-0.04049	-0.00518	0.73611	-0.02584	-0.22231	0.06103	
6.	NF-T6-1322	24.3 ± 8.1	2.3	-0.02401	-0.00567	0.63395	-0.02902	-0.24221	0.05370	
7.	NF-T6-1251	24.3 ± 8.0	2.3	-0.01695	-0.00571	0.49632	-0.03105	-0.17429	0.05117	
8.	NF-T6-1356	24.3 ± 8.2	2.3	-0.02890	-0.00547	0.35792	-0.02621	-0.08925	0.05820	
9.	NF-T6-1358	24.3 ± 8.3	2.3	-0.02890	-0.00547	0.35792	-0.02546	-0.08925	0.05820	
10.	NF-T6-1234	24.3 ± 8.4	2.3	-0.04219	-0.00544	0.67576	-0.02262	-0.17461	0.06129	
598.	VINA	21.1 ± 7.4	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p style="text-align: center;">Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p style="text-align: center;">Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T6-1473	NF-T6-1253	NF-T6-1472	NF-T6-1116	NF-T6-1479	NF-T6-1322	NF-T6-1251	NF-T6-1356	NF-T6-1358	NF-T6-1234
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

Tabela 40. Najlepsze funkcje wytrenowane w scenariuszu T6 w porównaniu do funkcji oryginalnej AutoDock Vina. Walidacja w scenariuszu W8 (T6W8). Parametry statystyczne walidacji i trenowania oraz wagi równania funkcji.

L.p.	Kod funkcji	logAUC ¹ [%]	MAE ² (trenowanie)	Wagi funkcji ³						Wykres gęstości rozkładu logAUC
				w ₁	w ₂	w ₃	w ₄	w ₅	w ₆	
1.	NF-T6-1534	19.2 ± 8.1	2.3	-0.02819	-0.00542	0.55053	-0.02004	-0.27603	0.04854	<p style="text-align: center;">scenariusz T6W8</p> <p style="text-align: center;">funkcja</p> <p style="text-align: center;">czarowna linia – mediana, VINA – funkcja oryginalna</p>
2.	NF-T6-1324	19.2 ± 8.0	2.3	-0.02581	-0.00567	0.59569	-0.02218	-0.09391	0.04699	
3.	NF-T6-1200	19.2 ± 8.1	2.3	-0.02389	-0.00619	0.42315	-0.00420	-0.02926	0.04237	
4.	NF-T6-1202	19.2 ± 8.1	2.3	-0.02242	-0.00619	0.42315	-0.00420	-0.02926	0.04237	
5.	NF-T6-1357	19.1 ± 8.1	2.3	-0.03087	-0.00579	0.54429	-0.01747	-0.23276	0.05344	
6.	NF-T6-1472	19.1 ± 7.9	2.3	-0.02918	-0.00510	0.45951	-0.02499	-0.20444	0.05164	
7.	NF-T6-1470	19.1 ± 7.9	2.3	-0.02606	-0.00561	0.37740	-0.02101	-0.11166	0.05164	
8.	NF-T6-1306	19.1 ± 8.0	2.3	-0.02946	-0.00528	0.53707	-0.02439	-0.31235	0.05135	
9.	NF-T6-1173	19.0 ± 7.9	2.3	-0.02714	-0.00521	0.59758	-0.02750	-0.36310	0.04809	
10.	NF-T6-1145	19.0 ± 8.0	2.3	-0.01992	-0.00607	0.35640	-0.01623	-0.07149	0.04914	
601.	VINA	16.5 ± 7.2	2.3	-0.03558	-0.00516	0.84025	-0.03507	-0.58744	0.05846	
<p>Porównanie wartości logAUC dla poszczególnych funkcji wobec funkcji oryginalnej (VINA) N = 500 (500 losowanych podzbiorów, każdy po 7 związków aktywnych i 7*30 związków nieaktywnych)</p> <p>Jednokierunkowa ANOVA z testem wielokrotnych porównań Dunnetta, oznaczenia istotności: ns – p > 0.05 (brak istotnej statystycznie różnicy), * – p ≤ 0.05, ** – p ≤ 0.01, *** – p ≤ 0.001, p ≤ 0.001 – ****.</p>										
vs	NF-T6-1534	NF-T6-1324	NF-T6-1200	NF-T6-1202	NF-T6-1357	NF-T6-1472	NF-T6-1470	NF-T6-1306	NF-T6-1173	NF-T6-1145
VINA	****	****	****	****	****	****	****	****	****	****

¹ log AUC – pole pod krzywą charakterystyki operacyjnej odbiornika wykreślonej w skali semilogarytmicznej, średnia ± odchylenie standardowe; ² MAE – średni błąd bezwzględny (ang. *Mean Absolute Error*) predykcji entalpii swobodnej wiązania w zbiorze treningowym; ³ wagi funkcji oceniającej, por. równania 3–5.

V. Dyskusja

Szeroki udział receptora NK1 oraz SP w szeregu procesów fizjologicznych i patologicznych czyni je punktami centralnymi, zarówno jeśli chodzi o mechanizmy chorobowe, jak i potencjalne terapie. Na przestrzeni lat antagoniści NK1R były rozważane jako potencjalne terapie m.in. depresji, schizofrenii, uzależnień, czy leczeniu bólu [49,69]. Innym wątkiem badawczym było ich użycie do leczenia nudności i wymiotów związanych z chemioterapią. W tym zastosowaniu, użyteczność antagonistów NK1R została potwierdzona klinicznie [68] i obecnie w medycynie ludzkiej oraz weterynaryjnej stosuje się sześć zatwierdzonych antagonistów NK1R.

Bogata biologia NK1R powoduje jednak, że badania nad ligandami tego receptora trwają dalej. Wśród nowszych trendów dotyczących potencjalnego wykorzystania celowania w NK1R jako strategii terapeutycznej można wymienić tworzenie wielocelowych związków przeciwbólowych działających także przez NK1R [145,147–150,153–158], antagonizm endosomalnego sygnałowania przez NK1R [43,161–164], selektywną cytotoksyczność antagonistów NK1R [175,177,178], czy też wykorzystanie ligandów NK1R jako wektorów dla celowanego dostarczania elementu terapeutycznego (w tym radionuklidów) [179–188,195–200].

Możliwość wykorzystania NK1R jako punktu uchwytu dla coraz nowszych terapii wymaga poszukiwania nowych ligandów tego receptora lub takich modyfikacji ligandów już opisanych, aby umożliwić konkretne specyficzne zastosowanie (np. tworzenie koniugatów o odpowiednich właściwościach fizykochemicznych *etc.*).

Powyższe tło stanowi główną motywację dla badań opisanych w rozprawie doktorskiej. Miały one charakter dwutorowy. Po pierwsze, dotyczyły syntezy nowych związków, pochodnych fenyloalaniny i tryptofanu jako potencjalnych ligandów NK1R. Po drugie, dotyczyły usprawnienia funkcji oceniających dokowania pod kątem zastosowania ich w dokowaniu do receptora NK1.

W kontekście poszukiwania nowych ligandów NK1R, zaprojektowano kilka serii związków, które a) zawierały kluczowe dla powinowactwa do NK1R elementy zawieszane na nowych, niestosowanych wcześniej w antagonistach NK1R, szkieletach (serie W/F-71–W/F-74; Tabela 4 i Rysunek 17), b) stanowiły rozbudowanie znanego antagonisty NK1R L-732,138 (serie W-80–W-84; Tabela 5) lub c) były jego modyfikacją zawierającą „nowe” ugrupowanie pentafluorosulfanyłowe (SF₅) zastępujące grupę CF₃ (seria W-90;

Tabela 6). Zaprojektowane związki otrzymałam w ramach kilkietapowych, nieskomplikowanych syntez, wychodząc z tanich i komercyjnie dostępnych N^{α} -Boc lub N^{α} -Cbz pochodnych fenyloalaniny i tryptofanu, stosując znane techniki syntezy organicznej oraz syntezy peptydów w roztworze. Pozwoliło to na uzyskanie wielu zróżnicowanych analogów w stosunkowo krótkim czasie i przy względnie niskim nakładzie finansowym.

Ponieważ priorytetem moich prac było szybkie uzyskanie szerokiej biblioteki związków, nie prowadziłam optymalizacji poszczególnych etapów syntezy. Wydaje się jednak, że przy dalszej rozbudowie zestawu analogów warto byłoby podjąć się takiej optymalizacji w celu poprawy wydajności i ekonomii syntez. Przykładowo, można spróbować wytwarzać hydrazydy N^{α} -Cbz-aminokwasów oraz kwasów przez hydrazynolizę estrów metylowych odpowiednich aminokwasów/kwasów [300]. Użycie tego podejścia zamiast sprzęgania aktywowanego kwasu/aminokwasu z hydrazyną pozwoliłoby uniknąć tworzenia N,N' -dicykloheksylomocznika, którego usunięcie bywa kłopotliwe [301], a śladowe ilości mogą „przechodzić” nawet do produktów następnych etapów. Innym polem do optymalizacji mógłby być dobór odczynnika sprzęgającego i/lub dodatku 1,2-dinukleofila dla reakcji sprzęgania [302,303]. W przypadku pochodnych tryptofanu, w trakcie acydolizy grupy Boc za pomocą TFA, kilka razy obserwowałam tworzenie produktów ubocznych o masach +56u (lub wielokrotności). Prawdopodobnie są to produkty mono- i poli-*tert*-butylowania na pierścieniu indolowym przez uwolnione w trakcie acydolizy karbokationy *tert*-butylowe [304,305]. W niektórych syntezach te produkty uboczne stanowiły nawet ponad 20% procent materiału w mieszaninie reakcyjnej (według pola powierzchni pików z analiz HPLC) i były nieusuwalne inaczej niż chromatograficznie, co przyczyniało się do spadku wydajności acydolizy. W związku z tym optymalizacja mogłaby dotyczyć np. użycia odpowiednich „wymiataczy” karbokationów (ang. *scavengers*) [306], zastosowania zabezpieczonej pochodnej tryptofanu jako substratu, albo użycia innej metody deprotekcji [305].

Proces syntezy prowadzony był od umownego² C-końca, co oznacza, że kolejne elementy były dodawane w kierunku grupy N-końcowej cząsteczki. Takie podejście miało umożliwić 1) rozpoznanie zależności SAR na N-końcu, oraz 2) ewentualną „późną” modyfikację (rozbudowywanie) grupy N-końcowej. Z syntetycznego punktu widzenia

² O C- i N-końcu omawianych cząsteczek mówię tu przez analogię do terminologii stosowanej w syntezie peptydów, mając na myśli ugrupowania dołączana do odpowiednio grupy karboksylowej albo grupy aminowej szkieletowego aminokwasu (fenyloalaniny lub tryptofanu).

funkcjonalizacja przez grupę aminową nie jest skomplikowana (np. przez typowe sprzężanie albo aminowanie redukcyjne), a stwarza potencjał rozbudowania związku o kolejne fragmenty o określonej roli biologicznej. Daje zatem możliwość łatwego tworzenia radiofarmaceutyków, koniugatów, czy związków wielofunkcyjnych na bazie opracowanych związków. W przypadku moich związków z serii W-80–W-84, możliwość ta została wykorzystana przez dr. Pawła Halika z IChTJ do syntezy wektorów i radiofarmaceutyków [1], do czego wrócę jeszcze poniżej.

Dwa z otrzymanych związków, H-F-71 oraz Boc-F-74, wykryły się z frakcji zebranych podczas chromatografii preparatywnej, dając kryształy odpowiednie do rentgenowskiej analizy strukturalnej. Umożliwiło to rozwiązanie ich struktur w cieple stałym oraz pogłębione badania spektroskopowe, w tym z wykorzystaniem metod chiralooptycznych [4]. Prawdopodobnie, inne związki z tych serii w odpowiednich warunkach również mogłyby wykryć się i możliwe byłoby rozwiązanie ich struktur krystalograficznych. Może to stanowić ciekawy, poboczny wątek dla zastosowania moich związków jako przedmiotu badań chemii strukturalnej. Przykładowo, w strukturze H-F-71 [4] obserwuje się względnie rzadko opisywane oddziaływanie $N=C \cdots \pi$ pomiędzy ugrupowaniem hydrazonu a pierścieniami aromatycznymi [307], lub wewnątrzcząsteczkowe oddziaływanie typu dziury σ (ang. *σ -hole*) we fragmentach $C=N \cdots CH-NH_3$ [308]. Ze względu na relatywnie łatwą syntezę, możliwe jest przygotowanie i wykorzystanie dodatkowych analogów, zaprojektowanych pod kątem rozwiązywania zagadnień z zakresu chemii strukturalnej.

Zsyntezowane przeze mnie związki zostały poddane ewaluacji biologicznej (wykonanej przez dr Joannę Matalińską z ZN IMDiK PAN). Badania powinowactwa receptorowego wobec NK1R posłużyły analizie zależności struktura–aktywność w otrzymanych bibliotekach w kontekście receptora NK1.

Pomimo uwzględnienia w strukturze związków z serii W/F-70–W/F-74 kluczowych dla powinowactwa do receptora NK1 elementów, *N*-acylohydrazony (serie 71), *N,N'*-diacylohydrazydy (serie 72), *N*-acylo-*N'*-fenylohydrazydy (serie 73), *N*-acylo-*N'*-fenyloacetylohydrazydy (serie 74) w przeprowadzonym badaniu powinowactwa były nieaktywne lub wykazywały jedynie nieznaczną aktywność wobec receptora NK1. Struktura tych związków była projektowana w oparciu o uzyskane wcześniej w ZN IMDiK PAN dane biologiczne oraz w oparciu o wyniki symulacji *in silico*, m.in. dokowania molekularnego. W czasie przeprowadzania symulacji komputerowych nie była jednak rozwiązana żadna struktura krystalograficzna receptora NK1, co mogło prowadzić do błędnych założeń projektowych. Przy obecnym stanie wiedzy, gdy dostępnych jest dwanaście struktur

krystalograficznych NK1R (w tym sześć z antagonistą w miejscu wiążącym) [23–28], dalsze projektowanie i rozszerzanie bibliotek może w większym stopniu opierać się na dostępnych danych strukturalnych. Powinno to potencjalnie przełożyć się na wyższe powinowactwo nowych analogów.

Jednocześnie warto zauważyć, że brak aktywności (lub słaba aktywność) związków z serii 71–74 stanowi ciekawą daną pogłębiającą znajomość zależności SAR dla NK1R. Jest tak tym bardziej, że analogi z tych serii są bardzo podobne do wielu znanych antagonistów NK1R, w tym do związku L-732,138 (Ac-W-80), czyli macierzystego analogu dla serii W-80. Przykładowo, Ac-W-73 oraz Ac-W-80 strukturalnie różnią się tylko dwoma atomami w środkowej części molekuly (Ac-W-73: -NH-NH-, a Ac-W-80: -O-CH₂-), a jednocześnie powinowactwo Ac-W-73 jest o około 540 razy słabsze niż powinowactwo Ac-W-80.

W ramach przeprowadzonych badań wykazano, że rozbudowanie znanego antagonisty NK1R L-732,138 (Ac-W-80) o dodatkowe elementy od N-końca nie przynosi utraty powinowactwa do NK1R (seria związków W-80–W-84). Powoduje to, że możliwe jest łatwe tworzenie koniugatów i związków wielofunkcyjnych w oparciu o te struktury. Właściwość tę wykorzystał dr Paweł Halik z IChTJ, stosując pochodne z tych serii jako związki wyjściowe do tworzenia wektorów dla radiofarmaceutyków znakowanych ⁶⁸Ga/¹⁷⁷Lu [1]. Zarówno pochodne stanowiące związki wyjściowe dla konstrukcji koniugatów, jak i same koniugaty (nieznakowane oraz znakowane) wykazywały dobre powinowactwo do NK1R [1]. Wskazuje to na sukces tej strategii rozbudowywania związku macierzystego oraz sugeruje, że związki te mogłyby mieć potencjalnie zastosowanie w radiofarmacji. Niestety, w badaniach dr. Halika radioznakowane koniugaty nie wykazały stabilności metabolicznej *in vitro* w surowicy krwi. Prawdopodobnie wynika to z obecności wiązania estrowego. Rozwiązaniem tego problemu mogłoby być wprowadzenie w strukturze cząsteczek modyfikacji, takich jak zamiana izosteryczna [309], co jest interesującym aspektem do dalszych badań. Warto zauważyć, że w literaturze opisano radiokoniugaty oparte o strukturę takich antagonistów NK1R jak aprepitant [187,188], czy L-733,060 [185,186]. Wydaje się jednak, że ewentualne dodatkowe prace SAR lub prace dotyczące optymalizacji właściwości fizykochemicznych koniugatów łatwiej prowadzić wokół moich analogów (lub ich wystabilizowanych pochodnych), z uwagi na znacznie prostszą i tańszą syntezę wektora.

W niniejszej pracy doktorskiej zademonstrowałam możliwość wprowadzenia grupy pentafluorosulfanylowej (SF₅) do ligandów NK1R. Analogi z serii W-90 (wzory na Rysunku 19) są pierwszymi opisanymi w literaturze ligandami NK1R zawierającymi grupę SF₅. W ostatnich latach podstawnik pentafluorosulfanylowy cieszy się dużym zainteresowaniem w

wielu dziedzinach chemii, w tym chemii nowoczesnych materiałów, katalizie, chemii strukturalnej, agrochemii czy chemii medycznej [283]. Seria W-90 jako przykład udanego zastosowania grupy SF₅ w ligandach NK1R otwiera możliwość wykorzystania właściwości tej grupy zarówno do tworzenia związków aktywnych działających przez NK1R, jak też związków narzędziowych.

Potencjalnie ciekawym zastosowaniem dla moich związków W-90 jest ich użycie do obrazowania metodą rezonansu magnetycznego ¹⁹F MRI. Niedawno, Prinz i wsp. wykazali, że grupa SF₅ posiada lepsze właściwości jako grupa reporterowa (w tym daje lepszy stosunek sygnału do szumu) niż grupy CF₃ czy OCF₃ [310]. Obecnie w ZN IMDiK PAN trwają przygotowania do wykorzystania Ac-W-90-Br (i nowych analogów otrzymanych na bazie tej struktury) do takiego obrazowania *in vivo*.

Jeśli chodzi o dane biologiczne i strukturalne serii W-90, warto tutaj wspomnieć o kilku ważniejszych ustaleniach. Według modelowania grupa SF₅ w analogach W-90 naśladuje pozycyjnie w kompleksie wiążącym grupę CF₃ z aprepitantu [27]. SAR serii W-90 pokazuje, że istotny dla powinowactwa do NK1R jest rodzaj drugiego podstawnika w pierścieniu aromatycznym (rozmiar podstawnika). Związki z większym podstawnikiem były aktywniejsze. Na podstawie przeprowadzonego dokowania i obserwowanych zależności SAR wydaje się, że jeszcze większy podstawnik hydrofobowy mógłby bardziej poprawić powinowactwo receptorowe, co jest warte dalszych badań. Co więcej, zaobserwowano, że Ac-W-90-Cl (R₂ = Cl) i Ac-W-90-Br (R₂ = Br) mają znacząco wyższe powinowactwo niż Ac-W-90-F (R₂ = F). Wskazuje to na możliwość występowania wiązania halogenowego [311] (choć nie zaobserwowałam takiego oddziaływania w geometriach kompleksów z dokowania). W celu weryfikacji tej hipotezy, w dalszych pracach można byłoby zsyntezować i przebadać analog z grupą R₂ = I, którego znaczący wzrost aktywności wskazywałby na potwierdzenie tego założenia. Dodatkowo, dla pełnego zrozumienia wpływu grupy R₂ w dalszych badaniach warto byłoby zsyntezować i sprawdzić aktywność analogów zawierających jako R₂ = CF₃ oraz R₂ = SF₅. Na obecnym etapie przeszkodą do syntezy takich analogów był/jest brak komercyjnie dostępnych bromków benzytowych z R₁ = SF₅ i R₂ = CF₃ lub R₂ = SF₅.

SAR serii W-90 wskazuje też, że kluczowa dla aktywności jest obecność grupy acetylowej na N-końcu. Analogi z wolną grupą aminową wykazują słabsze powinowactwo niż ich N-acetylowane odpowiedniki. Obserwacje te są zgodne z danymi dla innych estrów benzytowych tryptofanu z literatury [139,140], lub aktywnościami serii W-80 (ale nie dotyczy to już „dłuższych” analogów z serii W-81–W-84). Zmniejszone powinowactwo

analogów z wolną grupą aminową mogą wyjaśniać także wyniki z dokowania molekularnego, które sugerują ograniczone lub brak specyficznych oddziaływań między białkiem a grupą aminową w tych cząsteczkach oraz hipotetyczny wyższy koszt energetyczny częściowej desolvatacji sprotonowanej grupy aminowej w porównaniu z grupą acetylową.

Wyniki dokowania molekularnego (struktury przewidywanych kompleksów wiążących) sugerują także możliwość rozbudowy związków z serii Ac-W-90 przez dobudowywanie fragmentów do N-końca, analogicznie jak w serii W-80. Otwiera to pole do tworzenia koniugatów lub związków wielofunkcyjnych.

Innym rodzajem badania *in vitro*, w którym testowana były związki serii W/F-71–W/F-74 było badanie ich wpływu na żywotność komórek (cytotoksyczność). Motywacją do tego badania był fakt, że ligandy NK1R mogą wykazywać właściwości przeciwnowotworowe [31,165], a we wczesnych fazach mojego projektu doktoranckiego, w ZN IMDiK PAN nie funkcjonował jeszcze skuteczny i powtarzalny protokół oznaczania powinowactwa do NK1R metodą radioizotopową. W związku z tym moje związki były poddawane badaniu przesiewowemu dotyczącego wpływu na metabolizm komórkowy. Pomiar przeprowadzano na komórkach nowotworowych czerniaka linii COLO679 i prawidłowych fibroblastach BJ. Wybór linii COLO679 spowodowany był tym, że w literaturze dostępne są dane o aktywności znanych ligandów NK1R wobec tej linii komórkowej [171], co byłoby dobrym punktem odniesienia do porównania otrzymanych w ZN IMDiK PAN wyników.

Związki serii W/F-71–74 wykazywały zróżnicowaną aktywność cytotoksyczną. Spośród nich dwa związki Z-dW-71 oraz H-dW-71 cechują się korzystnym profilem aktywności, wykazując silne działanie cytotoksyczne, a dodatkowo aktywność ta jest selektywna wobec komórek COLO679. Jednocześnie, w późniejszych badaniach receptorowych ustalono, że związki te nie wykazują znaczącego powinowactwa wobec NK1R, co sugerowało, że mechanizm działania cytotoksycznego tych związków jest prawdopodobnie inny niż przez receptor NK1. Wobec tego, w ramach niniejszej pracy podjęłam próbę opracowania hipotezy dotyczącej możliwego mechanizmu działania tych związków, korzystając z metod *in silico*. Na podstawie predykcji otrzymanych z dokowania molekularnego oraz przewidywań typu *target-fishing* [312] uzyskanych za pomocą serwera SwissTarget [289,290] można przypuszczać, że związek Z-dW-71 może oddziaływać przez receptor cholecystokininowy CCK2, który jest znany z udziału w procesach nowotworowych [291–293]. Hipoteza działania przez CCK2R jest ważną hipotezą roboczą, która umożliwi

racjonalne planowanie dalszych prac eksperymentalnych (w tym syntezę nowych analogów), co byłoby znacznie utrudnione, gdyby opierać się wyłącznie o dane komórkowe.

Należy jednak podkreślić, że przewidywania *in silico* nigdy nie mogą zastąpić badań eksperymentalnych. Wynika to przede wszystkim z ograniczeń metod komputerowych stosowanych w badaniach nad lekami. Przykładowo, modele stosowane do predykcji siły wiązania w dokowaniu opierają się na uproszczeniach teoretycznych [215], a także są trenowane na ograniczonych zbiorach danych eksperymentalnych, dodatkowo obciążonych błędami eksperymentalnymi oraz uzyskanych w heterogeny sposób [299]. Powoduje to, że uzyskane predykcje entalpii swobodnej wiązania są niedokładne. Dalej, *target-fishing* przez dokowanie jest ograniczony również przez dobór badanych receptorów i przepustowość symulacji.

Jeśli chodzi o serwery/systemy ekspertowe służące do *target-fishingu* w oparciu o strukturę ligandów, ich działanie opiera się na obliczaniu pewnej metryki podobieństwa cząsteczkowego między badanym związkiem a znanymi ligandami danego celu molekularnego [313]. Metody te przyjmują, że związki „podobne” działają podobnie. W oczywisty sposób oznacza to, że nie mogą one jako aktywne zidentyfikować tych związków, które są znacząco różne od znanych już ligandów. Dodatkowo, jakość predykcji jest silnie zależna od wielkości zbioru treningowego, względem którego dokonywane są porównania podobieństwa.

Wobec wymienionych ograniczeń metod *in silico*, konieczne są dalsze badania eksperymentalne w celu weryfikacji hipotezy oddziaływania z CCK2R, obejmujące m.in. badania powinowactwa do tego receptora. Co więcej, w celu szerszego określenia aktywności komórkowej dla serii tych związków warto byłoby przeprowadzić pomiary cytotoksyczności na różnych liniach komórkowych. Dalsze prace nad tymi związkami wymagać będą także oznaczenia właściwości fizykochemicznych i farmakokinetycznych.

Druga część moich badań dotyczyła aspektów metodologicznych, związanych ze skutecznością funkcji oceniających stosowanych w dokowaniu molekularnym. Dokowanie jest jedną z najważniejszych technik *in silico* stosowanych w poszukiwaniu nowych leków. Kluczowym elementem dokowania jest ocena wygenerowanych rozwiązań (*scoring*). Idealny *scoring* powinien dobrze szacować energię wiązania kompleksu (w skali bezwzględnej), identyfikować najkorzystniejszą energetycznie pozę liganda w miejscu wiążącym oraz poprawnie szeregować różne ligandy według ich siły działania. W rzeczywistości dokładność funkcji oceniających pozostawia wiele do życzenia [240], a skuteczność predykcji i zastosowań jest zróżnicowana i zależna m. in. od celu molekularnego, do którego

przeprowadza się dokowanie [238–241]. Niektórzy autorzy wskazują, że przynajmniej częściowo problem ten wynika z zakładanej uniwersalności funkcji oceniających, a rozwiązaniem pozwalającym na poprawę skuteczności *scoringu* i dokowania może być przygotowywanie funkcji oceniających specyficznych dla danego celu molekularnego [243].

W pracach prowadzonych w ZN IMDiK PAN wielokrotnie obserwowano, że domyślna funkcja oceniająca popularnego programu AutoDock Vina [206] daje niskiej jakości szacowania energii wiązania ligandów z receptorem NK1 [niepublikowane]. W moich badaniach, systematycznie sprawdzałam zdolność tego programu do rozróżniania związków wiążących się od niewiążących się (słabowiążących się) z NK1R na szerszym zbiorze danych eksperymentalnych, z wykorzystaniem dokowania do różnych struktur eksperymentalnych NK1R oraz w różnych wariantach testowych. W analizie tej korzystałam z danych eksperymentalnych o wiązalności do NK1R uzyskanych w ZN IMDiK PAN dla związków, które zsyntezowałam w trakcie mojego projektu doktoranckiego, a także z danych pochodzących z literatury.

Moje badanie potwierdziło wcześniejsze spostrzeżenia, i w „trudniejszych” scenariuszach walidacji Vina uzyskała wyniki „słabe” (według kryterium Jaiteha³; logAUC ~ 7–8) lub „niezłe”/„dobre” (logAUC ~ 15–18) w „łatwiejszych” przypadkach walidacyjnych. Wobec tego postanowiłam wykorzystać dane eksperymentalne o powinowactwie receptorowym zsyntezowanych przez mnie związków oraz dane literaturowe do wytrenowania funkcji oceniających specyficznych dla receptora NK1.

We wszystkich scenariuszach trenowania, udało się uzyskać systematyczną (i statystycznie istotną) poprawę wartości logAUC względem wartości uzyskanych dla oryginalnej funkcji AutoDock Vina. W niektórych przypadkach, w zależności od scenariusza trenowania i scenariusza walidacji, najlepsze nowe funkcje miały lepszą kategorię według kryterium Jaiteha niż funkcja wyjściowa. W wyniku trenowań, uzyskano w zasadzie całe rodziny funkcji, z których wiele modeli uzyskało bardzo podobne wyniki walidacji. W obecnie uzyskanych danych nie ma podstaw do wyróżnienia jednej lub kilku z uzyskanych funkcji.

Według mojej wiedzy, w literaturze nie opisano żadnej funkcji oceniającej specyficznej dla receptora NK1. Nie badano także systematycznie skuteczności skryningu

³ Jaiteh i wsp. zaproponowali opisową kategoryzację wyników skryningu według wartości logAUC [298]. Według kryterium Jaiteha, logAUC może być słabe < 10, niezłe (ang. *fair*) 10–15, dobre 15–20, bardzo dobre 20–25, lub doskonałe > 25.

wobec NK1R dla AutoDock Vina, czy jakiegokolwiek innego programu dokującego. Stąd też bezpośrednie porównanie wartości logAUC uzyskanych w badaniu skuteczności Viny czy podczas walidacji nowych funkcji nie jest możliwe. Można je jednak zestawić z przeciętnymi wartościami obserwowanymi w dużych badaniach skuteczności dokowania.

U Mysingera i wsp. [314] program DOCK, zastosowany do przesiewu wobec 40 celów molekularnych, na danych z bazy DUD [297], uzyskał średnie logAUC = 15.0. Jaiteh i wsp. [298] porównywali skuteczność programu DOCK3.6 w skryningu wobec kilkunastu receptorów GPCR (na danych DUD-E [255]) w zależności od zastosowanego szablonu dla modelu homologicznego oraz plastyczności miejsca wiążącego, i uzyskali średnie wartości logAUC w zakresie 10.0–20.0. W niedawnej pracy Steina i wsp. [315], gdzie testowano skuteczność SVBS (z wykorzystaniem programu DOCK i różnorodnych zestawów parametrów sterujących dokowaniem) wobec 43 różnych celów molekularnych na danych ze zbioru DUD-E [255], obserwowano zróżnicowane wartości logAUC od około 60.0 do -9.0, ze średnia ~ 17.0 i medianą ~ 14.0. Warto oddzielnie zwrócić uwagę, że przeciętne wartości logAUC dla zbliżonych strukturalnie do NK1R receptorów z rodziny GPCR wynosiły średnio 6.4 (receptor dopaminowy D4), 8.3 (receptor adrenergiczny β_2), 22.6 (receptor adenozynowy A_{2A}), czy 10.3 (receptor chemokinowy CXCR₄). W innej systematycznej analizie, Xu i wsp. porównywali skuteczność skryningu dla programów DOCK3.7 i AutoDock Vina [316], również na danych z DUD-E [255]. Zaobserwowana mediana logAUC dla programu DOCK3.7 wynosiła 16.2, zaś dla AutoDock Vina 16.8. Przesiew wobec receptorów GPCR miał przeciętnie niższe logAUC niż przesiew wobec kinaz, proteaz, innych enzymów, czy receptorów jądrowych.

Przedstawione w pracy badania nad funkcjami oceniającymi specyficznymi dla NK1R nie wyczerpują zakresu możliwych ulepszeń w podejściu treningowym oraz możliwych i wartych do przeprowadzenia testów i porównań. Zwrócę teraz uwagę na kilka ograniczeń moich badań oraz ewentualne kierunki dalszych prac.

Po pierwsze, moje funkcje oceniające przeznaczone są tylko do *rescoringu*, tzn. służą do oceny geometrii kompleksów wygenerowanych w dokowaniu w ramach etapu przeszukiwania. Oznacza to, że modelowanie z ich udziałem wymaga zastosowania dwóch funkcji – jednej do kierowania przeszukiwaniem (a często też rankingowania rozwiązań), a drugiej do re-oceny rozwiązania/rozwiązań. Powoduje to, że jeśli funkcja kierująca przeszukiwaniem „skieruje” algorytm w region przestrzeni rozwiązań oddalony od prawdziwej konformacji kompleksu, funkcja *rescorująca* „pracować” będzie na nieprawidłowej geometrii, co może prowadzić do nieprawidłowych wyników. W literaturze

znane są próby tworzenia funkcji TSSF, które stosowane są zarówno do etapu przeszukiwania i do etapu oceny [248,250,261,264]. Jednak w przypadku moich badań, zrezygnowano z próby tworzenia funkcji specyficznej dla NK1R, która mogłaby kierować przeszukiwaniem rozwiązań z powodu znacznie większej złożoności problemu optymalizacyjnego, oraz niewielkiej ilości eksperymentalnych danych strukturalnych dotyczących kompleksów NK1R z ligandami. W związku z tym, stosując nowe funkcje, należy pamiętać o tym potencjalnym źródle błędów.

Po drugie, w prezentowanym w pracy podejściu, *rescoring* jest stosowany tylko do pierwszej (najniższej energetycznie) pozy wygenerowanej dla danego liganda. Wiąże się to z problemem podobnym jak powyżej – jeśli funkcja rankingująca pozy wynikające z przeszukania wybiera jako „najlepszą” błędną pozę wiążącą, funkcja rescorująca „pracuje” na nieprawidłowej geometrii, co może prowadzić do nieprawidłowych wyników. Możliwe byłoby uwzględnianie w *rescoringu* większej ilości póz wiążących, jednak podobnie jak w poprzednim przypadku, zwiększałoby to złożoność procedury optymalizacji, oraz wymagałoby większej ilości eksperymentalnych danych strukturalnych.

Jeśli chodzi o walidację uzyskanych funkcji, pewnym ograniczeniem jest to, że nie przeprowadziłam testów skuteczności wobec zbioru molekuł-wabików (ang. *decoys*), co jest często praktyką w pracach dotyczących skuteczności metodologii dokowania [255,314]. Wabiki są cząsteczkami, które posiadają podobne właściwości fizykochemiczne do prawdziwych ligandów danego receptora, lecz jednocześnie zupełnie niepodobne struktury. Przyjmuje się, że wabiki są nieaktywne wobec określonego receptora, i stosuje się je do testowania określonej metody (wariantu dokowania) pod kątem skuteczności skryningu. Wadą tego podejścia jest jednak to, że w zasadzie nie wiadomo, czy wabiki są naprawdę nieaktywne, oraz to, że znacznie bardziej wymagające jest odróżnianie podobnych strukturalnie związków aktywnych i nieaktywnych. Stąd też w mojej pracy stosowałam zbiory mniejsze, ale złożone ze związków, dla których istnieją eksperymentalne dane o aktywności. Warto będzie jednak uzupełnić te porównania o walidacje z wykorzystaniem molekuł-wabików. Ostatnio Stein i wsp. zaproponowali także badanie skuteczności wyboru prawdziwych ligandów spośród bibliotek związków wysoce naładowanych dodatnio lub ujemnie (tzw. zbiór *Extrema*) albo związków reprezentujących przeciętne właściwości fizykochemiczne typowej biblioteki przesiewowej (tzw. zbiór *Goldilocks*) [315]. Te testy również wydają się istotne dla pełnego zrozumienia potencjału moich funkcji i będą przedmiotem przyszłych badań.

Innym ważnym kierunkiem będzie porównanie skuteczności moich funkcji do skuteczności funkcji oceniających programów innych niż AutoDock Vina. Wybór programu AutoDock Vina był podyktowany przede wszystkim praktycznym stosowaniem tego programu w ZN IMDiK PAN. Co więcej, AutoDock Vina jest programem stosunkowo łatwym w użytkowaniu i dostępnym dla osób bez większego doświadczenia w modelowaniu i programowaniu. Dalej, program ten jest powszechnie dostępny (darmowy) i szybki. Funkcja oceniająca AutoDock Vina ma dość prostą postać (6 wyrazów, 6 wag; por. Równanie 3–5 oraz Tabela 3) i w przeciwieństwie do popularnych ostatnio funkcji „uczonych maszynowo” (MLSF), w tym specyficznych dla celów molekularnych (ML-TSSF), jest bardziej zrozumiała fizycznie. Inną przesłanką decydującą o użyciu AutoDock Vina było to, że w literaturze można odnaleźć opisane próby opracowania funkcji rescorujących na podstawie tego programu [254,268].

Niemniej jednak, znanych i stosowanych jest wiele programów dokujących oraz wiele funkcji oceniających. Pełne zrozumienie jakości predykcji moich nowych funkcji wymagać będzie odniesienia się także do innych stosowanych modeli. Szczególnie ciekawe może być porównanie moich funkcji z funkcjami oceniającymi opartymi na uczeniu maszynowym (MLSF). W nowszej literaturze wielokrotnie wskazywano, że MLSF są znacznie skuteczniejsze niż tradycyjne funkcje oceniające [228,229]. Z drugiej strony, krytycy wskazują na wady MLSF (np. przeuczenie, brak sensu fizycznego, brak prawdziwej zdolności predykcyjnej *etc.*) [230–235]. Według mojej wiedzy, nie opublikowano jak dotąd żadnego zastosowania MLSF do receptora NK1. Stąd interesującym zagadnieniem może być porównanie klasycznych funkcji oceniających (innych niż AutoDock VINA), funkcji typu MLSF i moich funkcji specyficznych dla NK1R.

Najważniejszym jednak testem dla moich funkcji byłoby zastosowanie prospektywne, czyli wykorzystanie jednej lub kilku z nich do predykcji powinowactwa do NK1R dla zupełnie nowej grupy związków chemicznych, następnie ich synteza i eksperymentalna ocena wiązania, a potem obliczenie współczynnika sukcesu w „świecie rzeczywistym”.

VI. Wnioski

- Zaprojektowano i zsyntezowano bibliotekę nowych pochodnych fenyloalaniny i tryptofanu, jako potencjalnych ligandów NK1R.
- Z użyciem kilku nieskomplikowanych reakcji oraz dywergentnego podejścia otrzymano ze stosunkową dobrą wydajnością 94 związki, w tym 93 nowe pochodne.
- Zsyntezowane *N*-acylohydrazony (serie W/F-71), *N,N'*-diacylohydrazydy (serie W/F-72), *N*-acylo-*N'*-fenylohydrazydy (serie W/F-73), *N*-acylo-*N'*-fenyloacetylohydrazydy (serie W/F-74) były nieaktywne lub jedynie marginalnie aktywne wobec receptora NK1.
- Możliwe jest rozbudowanie (na sposób umożliwiający tworzenie koniugatów i związków wielofunkcyjnych) znanego antagonisty NK1R L-732,138 o dodatkowe elementy od N-końca bez utraty powinowactwa do NK1R (serie związków W-80–W-84). Częściowe wyniki zostały opublikowane w *International Journal of Molecular Sciences* [1].
- Możliwe jest zastąpienie w strukturze znanego antagonisty L-732,138 kluczowej dla powinowactwa grupy CF₃ inną grupą SF₅ bez znaczącego spadku powinowactwa (seria W-90); grupa SF₅ według modelowania naśladuje pozycyjnie grupę CF₃ w kompleksie wiążącym. Częściowe wyniki zostały opublikowane w *ChemMedChem* [2].
- W serii W-90, ważne znaczenie dla aktywności ma rodzaj drugiego podstawnika w pierścieniu aromatycznym (rozmiar podstawnika) oraz obecność *N*-acetylowej grupy na umownym N-końcu.
- Związki z serii W/F-71–W/F-74 przebadane pod kątem wpływu na żywotność komórek (cytotoksyczności) cechowały się zróżnicowaną aktywnością wobec komórek nowotworowych COLO679 i prawidłowych BJ.
- Szczególnym profilem aktywności, charakteryzującym się działaniem silnie cytotoksycznym, ale jednocześnie selektywnym wobec linii komórkowej COLO67 cechowały się związki Z-dW-71 oraz H-dW-71, które jednocześnie nie wykazują znaczącego powinowactwa do NK1R. Z powodu tych własności związki te zostały opatentowane (Patent PL244031B1) [3].
- Zastosowano metody *in silico* (przewidywania SwissTarget oraz dokowanie molekularne), by opracować hipotezę sugerującą, że związki te mogą oddziaływać z

receptorem CCK2R (konieczne są dalsze badania eksperymentalne w celu potwierdzenia tej hipotezy).

- Oryginalna funkcja oceniająca programu AutoDock Vina stosunkowo słabo radzi sobie w szacowaniu energii wiązania ligandów z NK1R i odróżnianiu związków aktywnych od nieaktywnych (niska moc przesiewania).
- Wykorzystano dane eksperymentalne nowosyntezowanych związków oraz literaturowe do wytrenowania funkcji oceniających specyficznych dla NK1R.
- Opracowano nowe funkcje oceniające (do *rescoringu*) specyficzne dla NK1R, które w testach walidacyjnych wykazywały systematyczną i statystycznie istotną poprawę skuteczności w porównaniu do oryginalnej funkcji oceniającej programu AutoDock Vina.
- Nowe funkcje mogą być przydatnym narzędziem *in silico* wspomagającym odkrywanie nowych ligandów NK1R.

VII. Materiały i metody

VII.1. Chemia: uwagi ogólne.

— Wszystkie rozpuszczalniki i odczynniki dostępne handlowo używano bez dodatkowego oczyszczania.

— Przebieg reakcji monitorowano za pomocą chromatografii cienkowarstwowej (TLC), z użyciem płytek aluminiowych pokrytych żelazem krzemionkowym (Silica Gel 60 F254, Merck), z detekcją UV przy długości fali $\lambda = 254$ nm. Płytki wywoływano za pomocą 1% roztworu ninhydryny w metanolu (MeOH).

— Wszystkie zsyntezowane związki analizowano (oznaczanie masy cząsteczkowej oraz zawartości produktu głównego) za pomocą wysokosprawnej chromatografii cieczowej w układzie faz odwróconych z detektorem diodowym i sprzężonej ze spektrometrią mas typu pojedynczy kwadrupol (LC-MS, ang. *Liquid Chromatography – Mass Spectrometry*) z wykorzystaniem metody jonizacji elektrorozpylania (ESI). Analizy wykonywano na zestawie firmy Shimadzu (Shimadzu LCMS-2020), z wykorzystaniem kolumny Jupiter® 4 μm , 250 x 4.6 mm, Proteo 90 Å (Phenomenex) oraz faz ruchomych A i B (faza A: 0.05 % kwas mrówkowy (FA) w wodzie miliQ (v/v), faza B: 0.05% FA w acetonitrylu (ACN) (v/v)), stosując gradient 3–97% fazy B w 31 minut, przy przepływie 1.2 ml/min, z detekcją przy długości fali $\lambda = 210$ nm, 254 nm lub 280 nm.

— Zsyntezowane związki próbowano wytrącać i krystalizować. Tam, gdzie nie dało to zadowalającej czystości, do celów analitycznych oczyszczano małe próbki przy użyciu preparatywnej wysokosprawnej chromatografii cieczowej w układzie faz odwróconych (RP-HPLC, ang. *Reverse-Phase High Performance Liquid Chromatography*). Za pomocą RP-HPLC oczyszczano też wszystkie związki zawierające wolną (odbezpieczoną) grupę aminową (trifluorooctany; oznaczenia kodowe zaczynające się od **H-**). Oczyszczanie zostało przeprowadzone na aparacie firmy Shimadzu (Shimadzu Prominence) z detektorem diodowym. Używano kolumny Jupiter® 10 μm , 250 x 21.2 mm, Proteo 90 Å (Phenomenex) oraz faz ruchomych A i B (faza A: 0.1 % kwas trifluorooctowy (TFA) w wodzie miliQ (v/v), faza B: 0.1% TFA w ACN (v/v)), z detekcją przy długości fali $\lambda = 210$ nm, 254 nm i 280 nm. Rozdział chromatograficzny prowadzono wykorzystując elucję gradientową. Każdy ze związków wymagał doboru odpowiedniej metody oczyszczania (odpowiedniego gradientu). Największy wpływ na dobór gradientu miała obecność grup R₁. Przykładowe stosowane

gradienty: dla $R_1 = H$: 30-60% B przez 20 min; dla $R_1 = Boc$: 60-90% B przez 20 min; dla $R_1 = Cbz$: 65-85% B przez 20 min; dla $R_1 = Ac$: 30-80% B przez 20 min.

— W celu potwierdzenia struktury chemicznej wybranych związków przeprowadzono analizy z zastosowaniem spektroskopii magnetycznego rezonansu jądrowego (NMR, ang. *Nuclear Magnetic Resonance*). Pomiary wykonano w Instytucie Chemii Organicznej PAN w Pracowni Magnetycznego Rezonansu Jądrowego. Widma NMR protonowe (1H) i węglowe (^{13}C) rejestrowane były na spektrometrach Bruker 400 MHz, Bruker 500 MHz lub Varian 600 MHz. Pomiary wszystkich związków przeprowadzono w deuterowanym dimetylosulfotlenku- d_6 (DMSO- d_6) w temperaturze otoczenia, stosując trimetylosilan jako wzorzec wewnętrzny. Przy opisie poszczególnych sygnałów w widmie stosowano skróty: AB – układ spinowy AB, ABX – układ spinowy ABX, bs – szeroki singlet (ang. *broad singlet*), s – singlet, d – dublet, dd – dublet dubletów, dt – dublet trypletów, t – tryplet, q – kwartet, m – multiplet. Przesunięcia chemiczne sygnałów podane są w ppm, a stałe sprzężenia J w hercach (Hz).

VII.2. Procedury syntetyczne

VII.2.1. Otrzymywanie hydrazydów (Procedura A)

Do roztworu aminokwasu z grupą α -aminową zabezpieczoną grupą ochronną (*tert*-butoksykarbonylową (Boc) lub benzyloksykarbonylową (Cbz)) lub pochodnej aromatycznego kwasu karboksylowego (1 ekw.) w dimetyloformamidzie (DMF) dodano *N*-hydroksysukcynoimid (HOSu) (1.1 ekw.) i całość ochłodzono do temperatury 0–5°C za pomocą łaźni lodowej. Następnie dodano *N,N'*-dicykloheksylokarbodiimid (DCC) (1 ekw.) i prowadzono reakcję przez godzinę w temperaturze 0–5°C i kolejno przez dwie godziny w temperaturze pokojowej. Po tym czasie do mieszaniny reakcyjnej dodano monohydrat hydrazyny ($NH_2NH_2 \cdot H_2O$) (10 ekw) i całość pozostawiono do mieszania w temperaturze pokojowej na 20–24 godziny. Postęp reakcji sprawdzano za pomocą chromatografii cienkowarstwowej (TLC). Wytrącony *N,N'*-dicykloheksylokarbodiimid (DCU) odsączono, a przesącz zateżono pod zmniejszonym ciśnieniem. Produkt wytrącono za pomocą zimnego 15% wodnego roztworu wodorowęglanu potasu ($KHCO_3$). Otrzymany osad odsączono, przemyto 5% $KHCO_3$ oraz wodą, pozostawiono do wysuszenia na powietrzu.

VII.2.2. Sprzęganie grupy karboksylowej i aminowej (Procedura B)

Do roztworu aminokwasu z grupą α -aminową zabezpieczoną grupą ochronną (1 ekw.) w DMF dodano HOSu (1.1 ekw) i całość ochłodzono do temperatury 0–5°C za pomocą łaźni lodowej. Dodano DCC (1 ekw.) i mieszaninę reakcyjną pozostawiono do mieszania przez godzinę w temperaturze 0–5°C i kolejno przez dwie godziny w temperaturze pokojowej (do momentu wytrącenia DCU). Po tym czasie dodano roztwór soli aminy (seria W-81 i W-82) lub hydrazynu (seria W/F-72 i W/F-74) lub 3,5-bis(trifluorometylo)fenylohydrazyny (seria W/F-73) (1.2 ekw.) w DMF z dodatkiem 1,1,3,3-tetrametyloguanidyny (TMG) (1.2 ekw.). Całość pozostawiono do mieszania w temperaturze pokojowej, do momentu przereagowania substratów. Przebieg reakcji sprawdzano za pomocą TLC. Wytrącony podczas reakcji DCU odsączono, a przesącz zatężono pod zmniejszonym ciśnieniem. Produkt wyodrębniano poprzez wytrącenie zimnym 15% roztworem KHCO_3 lub, jeśli produkt nie strącał się, poprzez ekstrakcję octanem etylu (EtOAc). Fazę organiczną przemyto trzykrotnie 5% wodnym roztworem KHCO_3 i trzykrotnie wodą, a następnie osuszono za pomocą bezwodnego siarczanu magnezu (bezw. MgSO_4). Środek suszący odsączono, przesącz zatężono pod zmniejszonym ciśnieniem i pozostawiono do wysuszenia na powietrzu.

VII.2.3. Sprzęganie grupy karboksylowej i aminowej z wykorzystaniem preformowanego estru aktywnego (Procedura C)

Do estru *N*-sukcynimidyłowego kwasu 6-((*tert*-butoksykarbonylo)amino)heksanowego (Boc-6-Ahx-OSu) (1 ekw.) rozpuszczonego w minimalnej ilości DMF dodano aminę (1.2 ekw.) oraz TMG (1.2 ekw.). Reakcję prowadzono przez 6 h w temperaturze pokojowej. Mieszaninę reakcyjną zatężono pod zmniejszonym ciśnieniem i ekstrahowano EtOAc. Fazę organiczną osuszono bezw. MgSO_4 , przesączono, zatężono pod zmniejszonym ciśnieniem i pozostawiono do wysuszenia na powietrzu.

VII.2.4. Otrzymywanie hydrazonów (Procedura D)

Do 3,5-bis(trifluorometylo)benzaldehydu (1 ekw.) rozpuszczonego w minimalnej ilości mieszaniny izopropanol/metanol (iPrOH/MeOH) (2:1, *v/v*) dodano hydrazynę aminokwasu z grupą α -aminową zabezpieczoną grupą ochronną (1 ekw.). Reakcję pozostawiono do mieszania w temperaturze pokojowej przez 20 godzin. Postęp reakcji sprawdzano za pomocą TLC. Wytrącony osad odsączono, przemyto iPrOH i pozostawiono do wysuszenia na powietrzu.

VII.2.5. Usuwanie grupy ochronnej Boc (Procedura E)

Grupę *tert*-butyloksykarbonylową zabezpieczającą grupę α -aminową usuwano za pomocą kwasu trifluorooctowego (TFA). Związek rozpuszczono w mieszaninie TFA/chlorek metylenu (DCM) (1:1) w temperaturze 0–5°C w łaźni lodowej. Całość mieszano przez godzinę w temperaturze pokojowej. Postęp reakcji sprawdzano za pomocą TLC. Mieszaninę reakcyjną zatężono pod zmniejszonym ciśnieniem i przemyto trzykrotnie eterem dietylowym (Et₂O), za każdym razem odpędzając. Otrzymany olej pozostawiono na dobę do suszenia w eksykatorze próżniowym nad środkiem suszącym dekatlenkiem tetrafosforu (P₄O₁₀). Następnie w celu wytrącenia produktu do oleju dodano zimny Et₂O. Uzyskany osad odsączono i pozostawiono do wysuszenia w eksykatorze próżniowym.

VII.2.6. Acetylowanie wolnej grupy aminowej (Procedura F)

Do pochodnej z wolną grupą aminową (1 ekw.) rozpuszczonej w minimalnej ilości DMF lub mieszaninie DMF/DCM dodano TMG (2 ekw.) i całość ochłodzono do temperatury 0–5°C za pomocą łaźni lodowej. Do mieszaniny reakcyjnej dodano bezwodnik octowy (Ac₂O) (2 ekw.) i pozostawiono do mieszania na 30 minut w temperaturze 0–5°C i kolejno na godzinę w temperaturze pokojowej. Postęp reakcji sprawdzano za pomocą TLC. Mieszaninę reakcyjną zatężono pod zmniejszonym ciśnieniem. Następnie pozostałość zakwaszono 10% wodnym roztworem kwasu cytrynowego, dodano wodę i ekstrahowano EtOAc. Fazę organiczną przemyto trzykrotnie 5% wodnym roztworem kwasu cytrynowego i trzykrotnie wodą, a następnie osuszono za pomocą bezw. MgSO₄. Środek suszący odsączono, przesącz zatężono pod zmniejszonym ciśnieniem i pozostawiono do wysuszenia na powietrzu.

VII.2.7. Alkilowanie bromkami benzyłowymi (Procedura G)

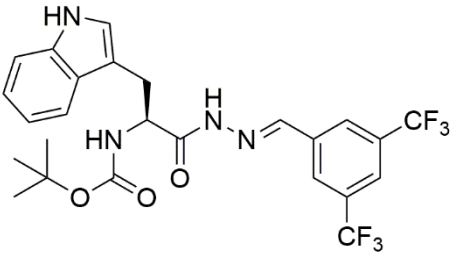
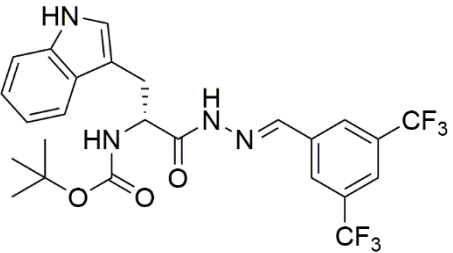
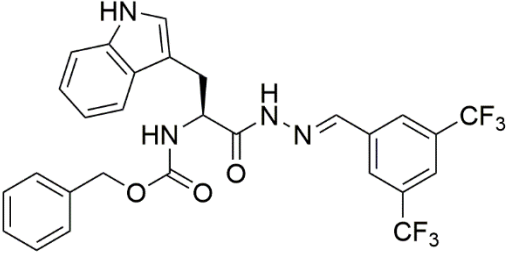
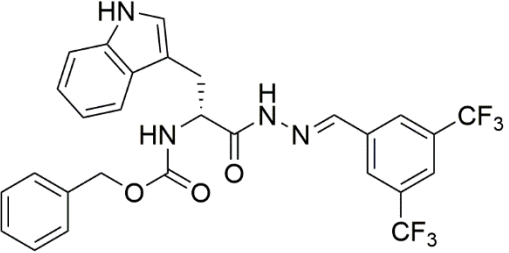
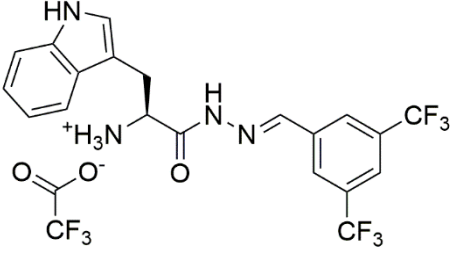
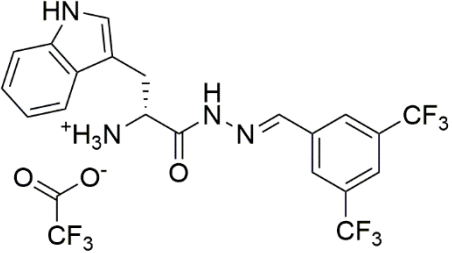
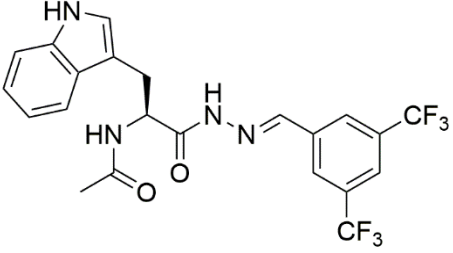
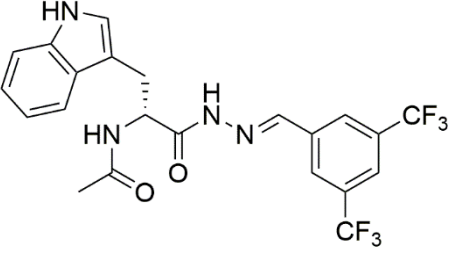
Do roztworu *N* ^{α} -(*tert*-butyloksykarbonylo)-L-tryptofanu (1 ekw.) w MeOH i wodzie dodano węglan cezu (Cs₂CO₃, 0.5 ekw.) rozpuszczony w wodzie i całość mieszano przez jedną godzinę w temperaturze pokojowej. Po tym czasie usunięto rozpuszczalnik pod zmniejszonym ciśnieniem, pozostałość przemyto dwukrotnie bezwodnym DMF, każdorazowo zatężając roztwór pod zmniejszonym ciśnieniem. Tak otrzymaną sól cezu zawieszono w DMF i dodano odpowiedni bromek benzyłowy podstawiony grupą SF₅ lub bromek 3,5-bis(trifluorometylo)benzyłowy (1.1 ekw.) w DMF. Całość mieszano przez 16–20 godzin w temperaturze pokojowej. Następnie mieszaninę reakcyjną zatężono pod zmniejszonym ciśnieniem, dodano wodę i ekstrahowano EtOAc. Fazę organiczną przemyto

trzykrotnie wodą, wysuszono bezwodnym MgSO_4 , przesączono i zateżono pod zmniejszonym ciśnieniem. Uzyskany olej suszono w eksykatorze próżniowym nad P_4O_{10} .

VII.3. Syntezy związków

VII.3.1. Seria W-71

Tabela 41. Wzory strukturalne związków z serii W-71.

<p>Boc-W-71</p> 	<p>Boc-dW-71</p> 
<p>Z-W-71</p> 	<p>Z-dW-71</p> 
<p>H-W-71</p> 	<p>H-dW-71</p> 
<p>Ac-W-71</p> 	<p>Ac-dW-71</p> 

Boc-L-Trp-NH-NH₂, Boc-L-Trp-NH-NH₂, (S)-(1-hydrazynylo-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 94%, ESI-MS m/z otrzymane m/z [M+H]⁺: 319.25, m/z obliczone [M+H]⁺: 319.18.

Boc-D-Trp-NH-NH₂, Boc-D-Trp-NH-NH₂, (R)-(1-hydrazynylo-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 93%, ESI-MS m/z otrzymane [M+H]⁺: 319.15, m/z obliczone [M+H]⁺: 319.18.

Z-L-Trp-NH-NH₂, Cbz-L-Trp-NH-NH₂, (S)-(1-hydrazynylo-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 87%, ESI-MS m/z otrzymane [M+H]⁺: 353.10, m/z obliczone [M+H]⁺: 353.16.

Z-D-Trp-NH-NH₂, Cbz-D-Trp-NH-NH₂, (R)-(1-hydrazynylo-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 68%, ESI-MS m/z otrzymane [M+H]⁺: 353.10, m/z obliczone [M+H]⁺: 353.16.

Boc-W-71, Boc-L-Trp-NH-N=CH-3,5(CF₃)₂Ph, (*S,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury D; kremowy osad, wydajność reakcji 94%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 543.10, m/z obliczone [M+H]⁺: 543.18; t_R = 19.94 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 12.09 i 11.78 (1H, 2s), 10.92 i 10.90 (1H, 2s), 8.48 i 8.41 (1H, 2s), 8.44 i 8.40 (2H, 2s), 8.24 i 8.22 (1H, 2s), 7.76-7.61 (1H, 2d, ³J_{HH} = 8.4, 7.8 Hz), 7.42 i 7.37 (1H, 2d, ³J_{HH} = 8.4, 7.2 Hz), 7.29 i 7.28 (1H, 2s), 7.22 i 7.16 (1H, 2d, ³J_{HH} = 8.4, 7.2 Hz), 7.16-7.12 i 7.11-7.06 (1H, 2m, nałożony), 7.11-7.06 i 6.95-6.90 (1H, 2m, nałożony), 5.30-5.24 i 4.45-4.38 (1H, 2m), 3.23-3.15 i 3.13-2.96 (2H, 2m), 1.43 i 1.40 (9H, 2s).

Boc-dW-71, Boc-D-Trp-NH-N=CH-3,5(CF₃)₂Ph, (*R,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury D; kremowy osad, wydajność reakcji 78%, czystość = 93% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 543.10, m/z obliczone [M+H]⁺: 543.18; t_R = 19,93 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 12.09 i 11.78 (1H, 2s), 10.92 i 10.90 (1H, 2s), 8.48 i 8.41 (1H, 2s), 8.44 i 8.40 (2H, 2s), 8.24 i 8.22 (1H, 2s), 7.76-7.61 (1H, 2d, ³J_{HH} = 8.4, 7.8

Hz), 7.42 i 7.37 (1H, 2d, $^3J_{\text{HH}} = 8.4, 7.2$ Hz), 7.29 i 7.28 (1H, 2s), 7.22 i 7.16 (1H, 2d, $^3J_{\text{HH}} = 8.4, 7.2$ Hz), 7.16-7.12 i 7.11-7.06 (1H, 2m, nałożony), 7.11-7.06 i 6.95-6.90 (1H, 2m, nałożony), 5.30-5.24 i 4.45-4.38 (1H, 2m), 3.23-3.15 i 3.13-2.96 (2H, 2m), 1.43 i 1.40 (9H, 2s).

Z-W-71, Cbz-L-Trp-NH-N=CH-3,5(CF₃)₂Ph, (*S,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury D; kremowy osad, wydajność reakcji 70%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 577.05, m/z obliczone [M+H]⁺: 577.17; t_R = 19.85 min;

¹H NMR (600 MHz, DMSO-*d*6) δ (ppm), 12.07 i 11.79 (1H, 2s), 10.84 i 10.82 (1H, 2s), 8.39 i 8.36 (1H, 2s), 8.36 i 8.31 (2H, 2s), 8.16 i 8.14 (1H, 2s), 7.73 i 7.68 (1H, 2d, $^3J_{\text{HH}} = 6.5, 6.5$ Hz), 7.66 i 7.54 (1H, 2d, $^3J_{\text{HH}} = 6.5, 6.5$ Hz), 7.37-7.10 (7H, nakładające się multiplety), 7.10-7.04 i 7.02-6.97 (1H, nakładające się multiplety), 7.02-6.97 i 6.86-6.80 (1H, nakładające się multiplety), 5.31-5.24 i 4.45-4.36 (1H, 2m), 5.03-4.84 (2H, m), 3.19-3.10 i 3.05-2.93 (2H, 2m).

Z-dW-71, Cbz-D-Trp-NH-N=CH-3,5(CF₃)₂Ph, (*R,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury D; kremowy osad, wydajność reakcji 89%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 577.10, m/z obliczone [M+H]⁺: 577.17; t_R = 19.85 min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 12.04 i 11.75 (1H, 2s), 10.83 i 10.81 (1H, 2s), 8.40 i 8.35 (1H, 2s), 8.35 i 8.31 (2H, 2s), 8.16 i 8.14 (1H, 2s), 7.71 i 7.68 (1H, 2d, $^3J_{\text{HH}} = 8.0$ i 8.0 Hz), 7.63 i 7.55 (1H, 2d, $^3J_{\text{HH}} = 8.0, 7.5$ Hz), 7.37-7.13 (7H, nakładające się multiplety), 7.11-7.04 i 7.04-6.96 (1H, nakładające się multiplety), 7.04-6.96 i 6.87-6.80 (1H, nakładające się multiplety), 5.33-5.24 i 4.47-4.36 (1H, 2m), 5.05-4.83 (2H, m), 3.21-3.10 i 3.06-2.93 (2H, 2m).

H-W-71, TFA*NH₂-L-Trp-NH-N=CH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*S,E*)-1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminowy; otrzymano wg procedury E; kremowy osad, wydajność reakcji 94%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 443.10, [M_{zasada}-H]⁻: 441.05, [M_{sól}-H]⁻: 555.00, m/z obliczone [M_{zasada}+H]⁺: 443.11, [M_{zasada}-H]⁻: 441.11, [M_{sól}-H]⁻: 555.11; t_R = 11.04 min;

¹H NMR (600 MHz, DMSO-*d*6) δ (ppm), 12.26 i 12.23 (1H, 2s), 11.03 i 10.97 (1H, 2s), 8.39 i 8.36 (1H, 2s), 8.31 i 8.21 (2H, 2s), 8.18 (3H, s), 8.14 (1H, s), 7.66 i 7.50 (1H, 2d, $^3J_{\text{HH}} = 6.5, 6.0$ Hz), 7.38 i 7.26 (1H, 2d, $^3J_{\text{HH}} = 7.0, 6.5$ Hz), 7.22 i 7.18 (1H, 2s, prawdopodobnie

nierozwiązane dublety), 7.10 i 7.03 (1H, 2d, $^3J_{\text{HH}} = 6.5, 6.0$ Hz), 7.00 i 6.89 (1H, 2t, $^3J_{\text{HH}} = 6.5, 6.5$ Hz), 5.04-4.97 i 4.11-4.06 (1H, 2bs, prawdopodobnie nierozwiązane multiplety), 3.36-3.29 (sygnał nakładający się z sygnałem od HDO) i 3.25-3.17 (2H, 2m).

H-dW-71, TFA*NH₂-D-Trp-NH-N=CH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R,E*)-1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminowy; otrzymano wg procedury E; kremowy osad, wydajność reakcji 86%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 443.05, [M_{zasada}-H]⁻: 441.00, [M_{sól}-H]⁻: 555.00, m/z obliczone [M_{zasada}+H]⁺: 443.13, [M_{zasada}-H]⁻: 441.11, [M_{sól}-H]⁻: 555.11; t_R = 11.00 min; ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 12.39 i 12.22 (1H, 2s), 11.04 i 11.03 (1H, 2s), 8.39 i 8.36 (1H, 2s), 8.29 i 8.19 (2H, 2s), 8.24 (3H, bs), 8.16 i 8.14 (1H, 2s), 7.67 i 7.51 (1H, 2t, $^3J_{\text{HH}} = 8.0, 8.0$ Hz), 7.38 i 7.26 (1H, 2d, $^3J_{\text{HH}} = 8.5, 8.5$ Hz), 7.23 i 7.20 (1H, 2d, $^3J_{\text{HH}} = 2.0, 2.0$ Hz), 7.13-7.06 i 7.05-7.00 (1H, 2m, nałożony), 6.99 i 6.89 (1H, 2t, $^3J_{\text{HH}} = 8.0, 8.0$ Hz), 5.05-4.98 i 4.16-4.10 (1H, 2m), 3.36-3.29 i 3.25-3.17 (2H, m).

Ac-W-71, Ac-NH-L-Trp-NH-N=CH-3,5(CF₃)₂Ph, (*S,E*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)acetamid; zsyntezowano wg procedury F; kremowy osad, wydajność reakcji 70%, czystość = 98% (HPLC-UV), m/z otrzymane [M+H]⁺: 485.05, m/z obliczone [M+H]⁺: 485.14; t_R = 16.90 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 12.12 i 11.81 (1H, 2s), 10.91 i 10.89 (1H, 2s), 8.45 i 8.44 (1H, 2s), 8.43 i 8.37 (2H, 2s), 8.31 i 8.22 (1H, d i s, $^3J_{\text{HH}} = 8.0$ Hz), 8.24 i 8.20 (1H, 2s), 7.74 i 7.62 (1H, 2d, $^3J_{\text{HH}} = 7.8, 7.8$ Hz), 7.41 i 7.35 (1H, 2d, $^3J_{\text{HH}} = 8.4, 8.4$ Hz), 7.26 i 7.25 (1H, 2s), 7.17-7.12 i 7.11-7.06 (1H, 2m, nałożony), 7.09-7.04 i 6.96-6.88 (1H, 2m, nałożony), 5.62-5.57 i 4.77-4.71 (1H, 2m), 3.26-3.20 i 3.10-3.01 (2H, 2m), 1.91 i 1.89 (3H, 2s).

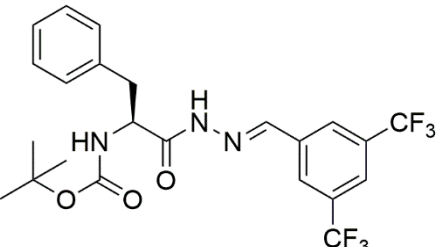
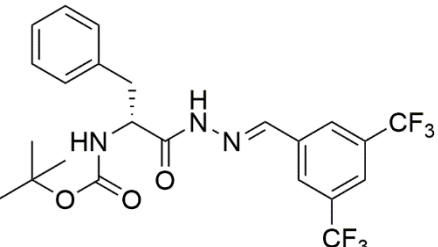
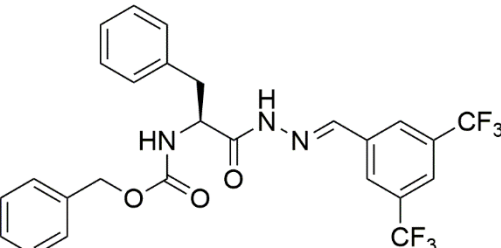
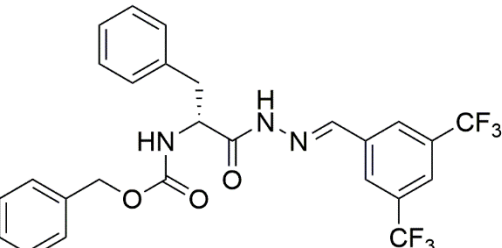
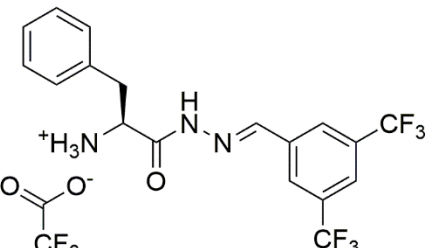
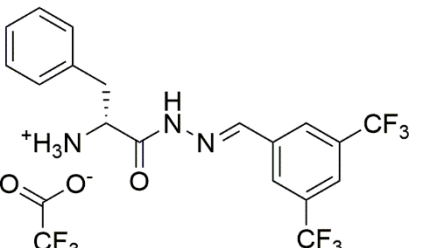
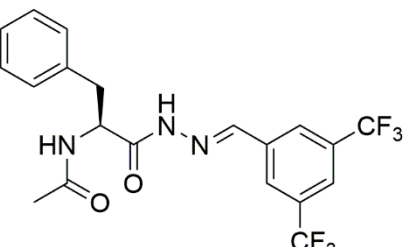
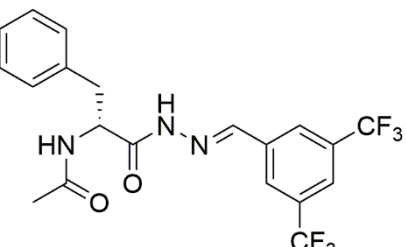
Ac-dW-71, Ac-NH-D-Trp-NH-N=CH-3,5(CF₃)₂Ph, (*R,E*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)acetamid; zsyntezowano wg procedury F; kremowy osad, wydajność reakcji 88%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 485.05, m/z obliczone [M+H]⁺: 485.14; t_R = 16.92 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 12.04 i 11.72 (1H, 2s), 10.82 i 10.80 (1H, 2s), 8.36 i 8.35 (1H, 2s), 8.34 i 8.28 (2H, 2s), 8.22 i 8.13 (1H, d i s, $^3J_{\text{HH}} = 8.0$ Hz), 8.15 i 8.12 (1H, 2s), 7.65 i 7.53 (1H, 2d, $^3J_{\text{HH}} = 7.8, 7.8$ Hz), 7.33 i 7.26 (1H, 2d, $^3J_{\text{HH}} = 7.8, 7.8$ Hz), 7.17 i 7.16 (1H, 2s), 7.10-7.03 i 7.02-6.95 (1H, 2m, nałożony), 7.00-6.95 i 6.86-6.80 (1H, 2m,

nałożony), 5.53-5.48 i 4.67-4.62 (1H, 2m), 3.17-3.10 i 3.01-2.92 (2H, 2m), 1.82 i 1.80 (3H, 2s).

VII.3.2. Seria F-71

Tabela 42. Wzory strukturalne związków z serii F-71.

<p>Boc-F-71</p> 	<p>Boc-dF-71</p> 
<p>Z-F-71</p> 	<p>Z-dF-71</p> 
<p>H-F-71</p> 	<p>H-dF-71</p> 
<p>Ac-F-71</p> 	<p>Ac-dF-71</p> 

Boc-L-Phe-NH-NH₂, Boc-L-Phe-NH-NH₂, (*S*)-(1-hydrazynylo-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 41%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 280.10, m/z obliczone [M+H]⁺: 280.34; t_R = 12.32 min.

Boc-D-Phe-NH-NH₂, Boc-D-Phe-NH-NH₂, (*R*)-(1-hydrazynylo-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 96%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 280.20, m/z obliczone [M+H]⁺: 280.34; t_R = 12.32 min.

Z-L-Phe-NH-NH₂, Cbz-L-Phe-NH-NH₂, (*S*)-(1-hydrazynylo-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 66%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 314.10, m/z obliczone [M+H]⁺: 314.15; t_R = 13.67 min.

Z-D-Phe-NH-NH₂, Cbz-D-Phe-NH-NH₂, (*R*)-(1-hydrazynylo-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury A; biały osad, wydajność reakcji 70%, czystość = 94 % (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 314.10, m/z obliczone [M+H]⁺: 314.15; t_R = 13.67 min.

Boc-F-71, Boc-L-Phe-NH-N=CH-3,5(CF₃)₂Ph, (*S,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury D; kremowy osad, wydajność reakcji 88%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 504.15, m/z obliczone [M+H]⁺: 504.17; t_R = 20.87 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 11.91 (bs), 8.49 i 8.45 (1H, 2s, nałożony), 8.49 i 8.27 (1H, 2s, nałożony), 8.25 i 8.24 (1H, 2s, nałożony), 7.46-7.26 (6H, m), 5.20-5.18 i 4.42-4.36 (1H, 2m), 3.14-3.08 i 3.02-2.89 (2H, 2m), 1.45 i 1.44 (9H, 2s).

Boc-dF-71, Boc-D-Phe-NH-N=CH-3,5(CF₃)₂Ph, (*R,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury D; kremowy osad, wydajność reakcji 79%, ESI-MS m/z otrzymane [M+H]⁺: 504.10, m/z obliczone [M+H]⁺: 504.17; t_R = 20.82 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 12.43-11.82 (bs), 8.50 i 8.45 (1H, 2s), 8.45 i 8.43 (2H, 2s), 8.25 i 8.22 (1H, 2s), 7.45-7.24 (6H, m), 5.16-5.09 i 4.39-4.33 (1H, 2m), 3.10-3.02 i 2.97-2.83 (2H, 2m), 1.41 i 1.39 (9H, 2s).

Z-F-71, Cbz-L-Phe-NH-N=CH-3,5(CF₃)₂Ph, (*S,E*)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury D, biały osad, wydajność reakcji 50%, czystość = 98%

(HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 538.10, m/z obliczone $[M+H]^+$: 538.16; t_R = 20.63 min;

1H NMR (500 MHz, DMSO- d_6) δ (ppm), 11.96 i 11.76 (1H, 2s), 8.33 i 8.31 (2H, 2s, nałożony), 8.11 i 8.10 (1H, 2s), 7.75 i 7.71 (1H, 2d, $^3J_{HH}$ = 8.0 Hz), 7.32-7.04 (10H, m, nakładające się multiplety), 5.11-5.05 i 4.36-4.28 (2H, 2s), 4.95-4.86 (2H, m), 3.03-2.91 i 2.85-2.74 (2H, 2m).

Z-dF-71, Cbz-D-Phe-NH-N=CH-3,5(CF₃)₂Ph, (R,E)-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury D, biały osad, wydajność reakcji 70%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 538.10, m/z obliczone $[M+H]^+$: 538.16; t_R = 20.63 min;

1H NMR (500 MHz, DMSO- d_6) δ (ppm), 12.19 i 11.43 (bs), 8.38 i 8.37 (1H, 2s, nałożony), 8.37 i 8.33 (2H, 2s, nałożony), 8.16 i 8.11 (1H, 2s, nałożony), 7.78 i 7.68 (1H, m, prawdopodobnie dwa nakładające się dublety), 7.39-7.07 (10H, m, nakładające się multiplety), 5.20-5.15 i 4.41-4.31 (2H, 2s), 5.02-4.88 (2H, m), 3.10-2.98 i 2.93-2.77 (2H, 2m).

H-F-71, TFA*NH₂-L-Phe-NH-N=CH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (S,E)-1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-aminowy; otrzymano wg procedury E; biały krystaliczny osad, wydajność reakcji 32%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M_{zasada}+H]^+$: 404.10, $[M_{zasada}-H]^-$: 402.05, $[M_{sól}-H]^-$: 516.00, m/z obliczone $[M_{zasada}+H]^+$: 404.12, $[M_{zasada}-H]^-$: 402.10, $[M_{sól}-H]^-$: 516.10; t_R = 10.92 min;

1H NMR (500 MHz, DMSO d_6) δ (ppm): 12.35 i 12.25 (1H, 2s), 8.42 (3H, s), 8.38 i 8.36 (1H, 2s), 8.31 i 8.16 (2H, 2s), 8.19 i 8.08 (1H, 2s), 7.40-7.09 (5H, m), 5.07 i 4.16 (2H, 2t, $^3J_{HH}$ = 7.0 Hz), 3.20-3.10 (2H, m);

^{13}C NMR (125 MHz, DMSO d_6) δ (ppm): 170.3 i 165.2, 157.9 (2q, $^2J_{CF}$ = 30.5 Hz), 145.2 i 141.6, 136.6 i 136.4, 135.1 i 134.6, 130.8 i 130.7 (2q, $^2J_{CF}$ = 32.5 Hz), 129.4 i 129.2, 128.5 i 128.2, 127.3 i 127.2, 127.0, 123.1 i 123.1 (2q, $^1J_{CF}$ = 272 Hz), 122.9, 117.3 (q, $^1J_{CF}$ = 299 Hz), 52.9 i 51.1, 36.8 i 36.5.

H-dF-71, TFA*NH₂-D-Phe-NH-N=CH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (R,E)-1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-aminowy; otrzymano wg procedury E; biały krystaliczny osad, wydajność reakcji 94%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M_{zasada}+H]^+$: 404.05, $[M_{zasada}-H]^-$:

402.05, $[M_{\text{sól-H}}]^-$: 516.00, m/z obliczone $[M_{\text{zasada+H}}]^+$: 404.12, $[M_{\text{zasada-H}}]^-$: 402.10, $[M_{\text{sól-H}}]^-$: 516.10; $t_R = 10.92$ min;

^1H NMR (500 MHz, DMSO *d*6) δ (ppm): 12.35 i 12.25 (1H, 2s), 8.42 (3H, s), 8.38 i 8.36 (1H, 2s), 8.31 i 8.16 (2H, 2s), 8.19 i 8.08 (1H, 2s), 7.40-7.09 (5H, m), 5.07 i 4.16 (2H, 2t, $^3J_{\text{HH}} = 7.0$ Hz), 3.20-3.10 (2H, m);

^{13}C NMR (125 MHz, DMSO *d*6) δ (ppm): 170.3 i 165.2, 157.9 (2q, $^2J_{\text{CF}} = 30.5$ Hz), 145.2 i 141.6, 136.6 i 136.4, 135.1 i 134.6, 130.8 i 130.7 (2q, $^2J_{\text{CF}} = 32.5$ Hz), 129.4 i 129.2, 128.5 i 128.2, 127.3 i 127.2, 127.0, 117.3 (q, $^1J_{\text{CF}} = 299$ Hz), 123.1 i 123.1 (2q, $^1J_{\text{CF}} = 272$ Hz), 122.9, 52.9 i 51.1, 36.8 i 36.6.

Ac-F-71, Ac-NH-L-Phe-NH-N=CH-3,5(CF₃)₂Ph, (*S,E*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid;
zsyntezowano wg procedury F; biały osad, wydajność reakcji 95%, czystość = 97% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 446.05, m/z obliczone $[M+H]^+$: 446.13; $t_R = 17.53$ min;

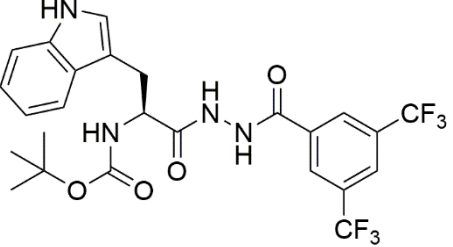
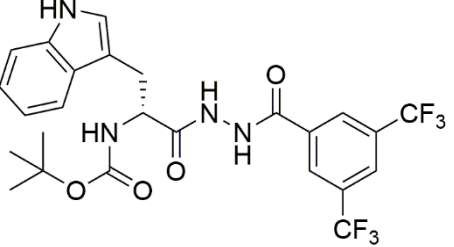
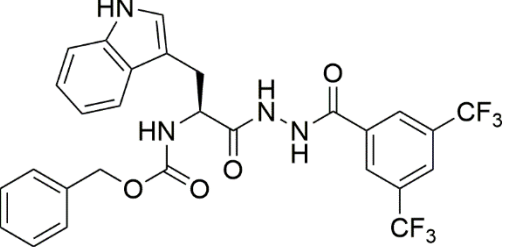
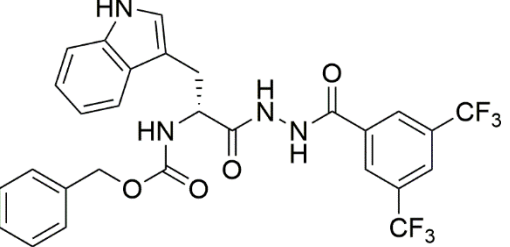
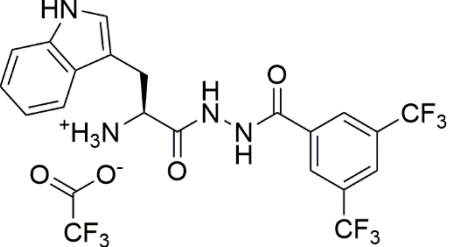
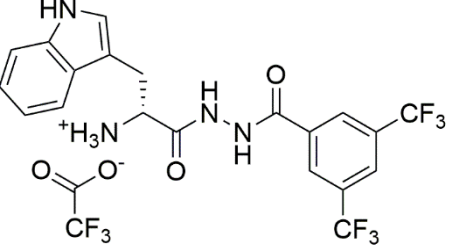
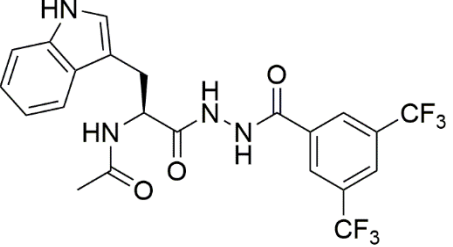
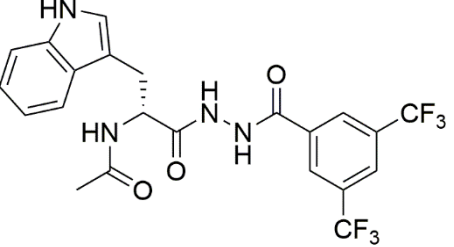
^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 11.98 i 11.77 (1H, 2s), 8.37 i 8.33 (1H, 2s), 8.32 (2H, s, nałożony), 8.29 i 8.11 (1H, d i s, $^3J_{\text{HH}} = 8.0$ Hz, nałożony), 8.12 i 8.11 (1H, 2s, nałożony), 7.29-7.24 (3H, m), 7.23-7.11 (2H, m), 5.38-5.26 i 4.62-4.52 (1H, 2m), 3.06-2.92 i 2.92-2.85 (2H, 2m), 1.76 (3H, s).

Ac-dF-71, Ac-NH-D-Phe-NH-N=CH-3,5(CF₃)₂Ph, (*R,E*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzylideno)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid;
zsyntezowano wg procedury F; biały osad, wydajność reakcji 51%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 446.05, m/z obliczone $[M+H]^+$: 446.13; $t_R = 17.58$ min;

^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 12.01 i 11.79 (1H, 2s), 8.39 i 8.36 (1H, 2s), 8.35 (2H, s, nałożony), 8.32 i 8.15 (1H, d i s, $^3J_{\text{HH}} = 8.0$ Hz, nałożony), 8.15 i 8.13 (1H, 2s, nałożony), 7.32-7.17 (5H, m), 5.38-5.30 i 4.64-4.55 (1H, 2m), 3.06-2.98 i 2.90-2.76 (2H, 2m), 1.80 (3H, s).

VII.3.3. Seria W-72

Tabela 43. Wzory strukturalne związków z serii W-72.

<p>Boc-W-72</p> 	<p>Boc-dW-72</p> 
<p>Z-W-72</p> 	<p>Z-dW-72</p> 
<p>H-W-72</p> 	<p>H-dW-72</p> 
<p>Ac-W-72</p> 	<p>Ac-dW-72</p> 

NH₂-NH-C(=O)-3,5(CF₃)₂Ph, 3,5-bis(trifluorometylo)benzohydrazyd; zsyntezowano wg procedury A; biały osad, wydajność reakcji 67%, ESI-MS m/z otrzymane [M+H]⁺: 273.05, m/z obliczone [M+H]⁺: 273.05; t_R = 10.76 min.

Boc-W-72, Boc-L-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, (S)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-

ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 84%, czystość = 97% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 559.10, *m/z* obliczone [M+H]⁺: 559.18; *t_R* = 18.56 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 11.02 (1H, s), 10.81 (1H, s), 10.43 (1H, s), 8.63 (1H, s), 8.55 (2H, s), 7.69 (1H, d, ²*J*_{HH} = 7.0 Hz), 7.32 (1H, d, ²*J*_{HH} = 6.5 Hz), 7.23 (1H, s), 7.05 (1H, dd, ²*J*_{HH} = 6.5, 6.0 Hz), 6.98 (1H, dd, ²*J*_{HH} = 6.5, 6.0 Hz), 6.90 (1H, d, ²*J*_{HH} = 7.0 Hz), 4.40-4.30 (1H, m), 3.21-3.12 i 3.01-2.92 (2H, 2m), 1.29 i 1.22 (9H, 2s).

Boc-dW-72, Boc-D-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, (R)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-

ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 86%, czystość > 97% (HPLC-UV), ESI-MS *m/z* otrzymane [M-H]⁻: 557.05, *m/z* obliczone [M-H]⁻: 557.16; *t_R* = 18.50 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 11.01 (1H, s), 10.81 (1H, s), 10.42 (1H, s), 8.54 (2H, s), 8.37 (1H, s), 7.68 (1H, d, ²*J*_{HH} = 6.5 Hz), 7.31 (1H, d, ²*J*_{HH} = 7.0 Hz), 7.05 (1H, dd, ²*J*_{HH} = 6.0, 6.0 Hz), 6.97 (1H, dd, ²*J*_{HH} = 6.5, 6.0 Hz), 6.90 (1H, d, ²*J*_{HH} = 7.0 Hz), 4.38-4.31 (1H, m), 3.19-3.13 i 3.00-2.93 (2H, 2m), 1.28 i 1.20 (9H, 2s).

Z-W-72, Cbz-L-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, (S)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-

ylo)karbaminian benzylu; zsyntezowano wg procedury B; kremowy osad, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 593.10, *m/z* obliczone [M+H]⁺: 593.16; *t_R* = 18.60 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 11.02 (1H, s), 10.84 (1H, s), 10.53 (1H, s), 8.57 (2H, s), 8.39 (1H, s), 7.73 (1H, d, ³*J*_{HH} = 9.6 Hz), 7.55 (1H, d, ³*J*_{HH} = 8.0 Hz), 7.40-7.21 (7H, nałożony), 7.08 (1H, d, ³*J*_{HH} = 7.5 Hz), 7.00 (1H, d, ³*J*_{HH} = 7.5 Hz), 5.01-4.87 (2H, m), 4.52-4.44 (1H, m), 3.26-3.14 i 3.06-2.95 (2H, 2m).

Z-dW-72, Cbz-D-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, (R)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-

ylo)karbaminian benzylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 73%, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 593.10, *m/z* obliczone [M+H]⁺: 593.16; *t_R* = 18.60 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 11.03 (1H, s), 10.84 (1H, s), 10.53 (1H, s), 8.57 (2H, s), 8.39 (1H, s), 7.73 (1H, d, ³*J*_{HH} = 5.0 Hz), 7.55 (1H, d, ³*J*_{HH} = 8.5 Hz), 7.39-7.20 (7H,

nałożony), 7.08 (1H, d, $^3J_{\text{HH}} = 7.5$ Hz), 7.00 (1H, d, $^3J_{\text{HH}} = 7.5$ Hz), 5.02-4.86 (2H, m), 4.53-4.44 (1H, m), 3.26-3.18 i 3.06-2.96 (2H, 2m).

H-W-72, TFA* NH_2 -L-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*S*)-1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminiowy; otrzymano wg procedury E; biały osad, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 459.05, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 457.05, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 571.00, m/z obliczone [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 459.13, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 457.12, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 571.11; $t_{\text{R}} = 10.23$ min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 11.31 (1H, s), 11.07 (1H, s), 11.04 (1H, s), 8.57 (2H, s), 8.42 (1H, s), 8.25 (3H, s), 7.78 (1H, d, $^3J_{\text{HH}} = 7.5$ Hz), 7.40 (1H, d, $^3J_{\text{HH}} = 8.5$ Hz), 7.33 (1H, d, $^3J_{\text{HH}} = 1.5$ Hz), 7.12 (1H, dd, $^3J_{\text{HH}} = 9.9$ Hz), 7.05 (1H, dd, $^3J_{\text{HH}} = 9.9$ Hz), 4.22-4.16 (1H, m), 3.42-3.34 i 3.22-3.15 (2H, 2m).

H-dW-72, TFA* NH_2 -D-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminiowy; otrzymano wg procedury E; biały osad, wydajność reakcji 69%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 459.05, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 457.05, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 571.00, m/z obliczone [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 459.13, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 457.12, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 571.11; $t_{\text{R}} = 10.21$ min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 11.31 (1H, s), 11.07 (1H, s), 11.04 (1H, s), 8.57 (2H, s), 8.42 (1H, s), 8.26 (3H, s), 7.78 (1H, d, $^3J_{\text{HH}} = 7.5$ Hz), 7.40 (1H, d, $^3J_{\text{HH}} = 8.0$ Hz), 7.33 (1H, d, $^3J_{\text{HH}} = 2.0$ Hz), 7.12 (1H, dd, $^3J_{\text{HH}} = 7.5, 7.5$ Hz), 7.05 (1H, dd, $^3J_{\text{HH}} = 7.5$ Hz), 4.22-4.16 (1H, m), 3.42-3.34 i 3.22-3.15 (2H, 2m).

Ac-W-72, Ac-NH-L-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*S*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-yl)acetamid; zsyntezowano wg procedury F; biały osad, wydajność reakcji 89%, czystość = 97% (HPLC-UV), ESI-MS m/z otrzymane [$\text{M}+\text{H}$]⁺: 501.05, m/z obliczone [$\text{M}+\text{H}$]⁺: 501.14; $t_{\text{R}} = 15.87$ min;

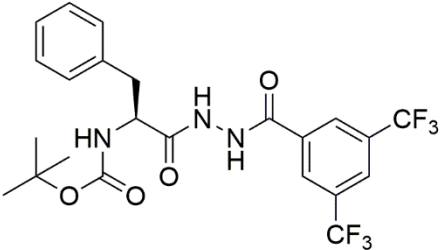
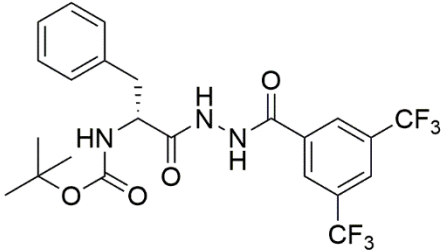
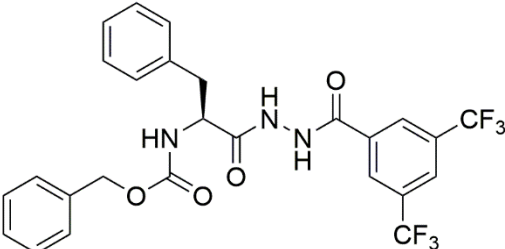
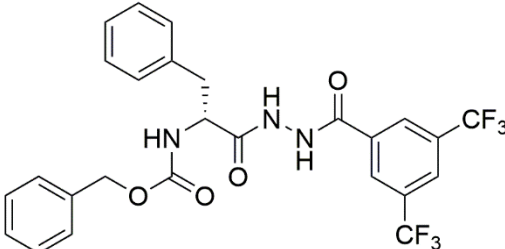
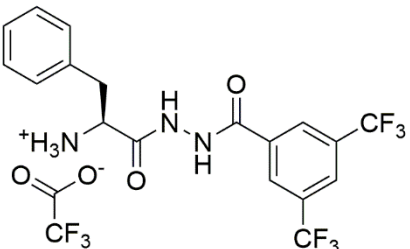
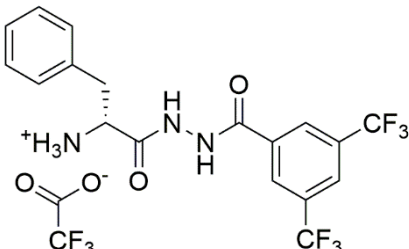
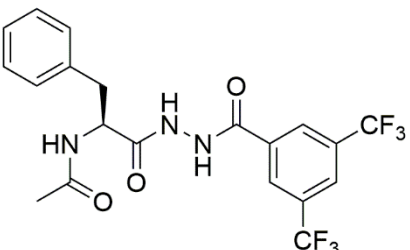
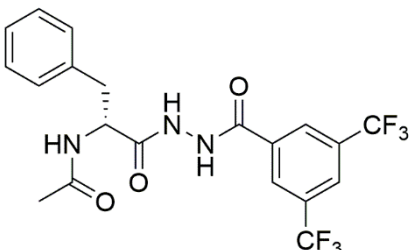
¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 10.99 (1H, s), 10.83 (1H, s), 10.48 (1H, s), 8.55 (2H, s), 8.40 (1H, s), 8.21 (1H, d, $^3J_{\text{HH}} = 7.8$ Hz), 7.67 (1H, d, $^3J_{\text{HH}} = 7.8$ Hz), 7.34 (1H, d, $^3J_{\text{HH}} = 7.8$ Hz), 7.23 (1H, s), 7.07 (1H, dd, $^3J_{\text{HH}} = 7.8, 7.5$ Hz), 7.00 (1H, dd, $^3J_{\text{HH}} = 7.5, 7.5$ Hz), 4.75-4.69 (1H, m), 3.24-3.18 i 3.00-2.94 (2H, 2m), 1.79 (3H, s).

Ac-dW-72, Ac-NH-D-Trp-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-yl)acetamid; zsyntezowano wg procedury F; biały osad, wydajność reakcji 68%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [$\text{M}+\text{H}$]⁺: 501.05, m/z obliczone [$\text{M}+\text{H}$]⁺: 501.14; $t_{\text{R}} = 15.81$ min;

^1H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 10.99 (1H, s), 10.83 (1H, s), 10.48 (1H, s), 8.55 (2H, s), 8.40 (1H, s), 8.21 (1H, d, $^3J_{\text{HH}} = 8.4$ Hz), 7.67 (1H, d, $^3J_{\text{HH}} = 7.8$ Hz), 7.34 (1H, d, $^3J_{\text{HH}} = 7.8$ Hz), 7.23 (1H, s), 7.07 (1H, dd, $^3J_{\text{HH}} = 7.8, 7.5$ Hz), 7.00 (1H, dd, $^3J_{\text{HH}} = 7.2, 7.2$ Hz), 4.75-4.69 (1H, m), 3.24-3.18 i 3.00-2.94 (2H, 2m), 1.79 (3H, s).

VII.3.4. Seria F-72

Tabela 44. Wzory strukturalne związków z serii F-72.

<p>Boc-F-72</p> 	<p>Boc-dF-72</p> 
<p>Z-F-72</p> 	<p>Z-dF-72</p> 
<p>H-F-72</p> 	<p>H-dF-72</p> 
<p>Ac-F-72</p> 	<p>Ac-dF-72</p> 

Boc-F-72, Boc-L-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, (S)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 70%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 518.05, m/z obliczone [M-H]⁻: 518.14; t_R = 18.95 min;

^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 11.05 (1H, s), 10.43 (1H, s), 8.57 (2H, s), 8.40 (1H, s), 7.37 (2H, d, $^3J_{\text{HH}} = 7.2$ Hz), 7.30 (2H, dd, $^3J_{\text{HH}} = 7.2, 7.2$ Hz), 7.21 (1H, t, $^3J_{\text{HH}} = 7.2$ Hz), 7.10 (1H, d, $^3J_{\text{HH}} = 9.0$ Hz), 4.37-4.31 (1H, m), 3.10-3.05 i 2.80-2.75 (2H, 2m), 1.31 i 1.27 (9H, 2s).

Boc-dF-72, Boc-D-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*R*)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 70%, czystość > 97% (HPLC-UV), ESI-MS *m/z* otrzymane [M-H]⁻: 518.05, *m/z* obliczone [M-H]⁻: 518.14; *t_R* = 18.95 min;

^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 11.10 (1H, bs), 10.43 (1H, bs), 8.60 (2H, s), 8.42 (1H, s), 7.48 (2H, d, $^3J_{\text{HH}} = 7.2$ Hz), 7.43 (2H, dd, $^3J_{\text{HH}} = 7.8, 7.2$ Hz), 7.34 (1H, t, $^3J_{\text{HH}} = 7.2$ Hz), 7.22 (1H, d, $^3J_{\text{HH}} = 9.0$ Hz), 4.48-4.43 (1H, m), 3.24-3.19 i 3.04-2.92 (2H, 2m), 1.43 i 1.39 (9H, 2s).

Z-F-72, Cbz-L-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*S*)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 73%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 554.05, *m/z* obliczone [M+H]⁺: 554.15; *t_R* = 19.01 min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.03 (1H, s), 10.49 (1H, s), 8.56 (2H, s), 8.39 (1H, s), 7.67 (1H, d, $^3J_{\text{HH}} = 7.5$ Hz), 7.40-7.10 (10H, m), 5.05-4.90 (2H, bs, prawdopodobnie dwa nierozwiązane dublety), 4.43 (1H, bs), 3.16-3.04 i 2.92-2.80 (2H, m).

Z-dF-72, Cbz-D-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*R*)-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 90%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 554.10, *m/z* obliczone [M+H]⁺: 554.15; *t_R* = 18.99 min;

^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 11.03 (1H, s), 10.50 (1H, s), 8.55 (2H, s), 8.40 (1H, s), 7.69 (1H, d, $^3J_{\text{HH}} = 7.8$ Hz), 7.40-7.10 (10H, m), 4.99-4.92 (2H, bs, prawdopodobnie dwa nierozwiązane dublety), 4.44-4.38 (1H, m), 3.14-3.08 i 2.88-2.81 (2H, m).

H-F-72, TFA*NH₂-L-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*S*)-1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-aminiowy; otrzymano wg procedury E; biały osad, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M_{zasada}+H]⁺: 420.05, [M_{zasada}-H]⁻: 418.00, [M_{sól}-H]⁻: 532.00, *m/z* obliczone [M_{zasada}+H]⁺: 420.11, [M_{zasada}-H]⁻: 418.09, [M_{sól}-H]⁻: 532.09; *t_R* = 10.20 min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.27 (1H, s), 10.94 (1H, s), 8.55 (2H, s), 8.41 (1H, s), 8.35 (3H, bs), 7.39-7.37 (4H, m), 7.34-7.30 (1H, m), 4.26-4.19 (1H, m), 3.28-3.22 i 3.11-3.04 (2H, 2m).

H-dF-72, TFA*NH₂-D-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-aminiowy; otrzymano wg procedury E; biały osad, wydajność reakcji 37%, czystość = 95% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 420.05, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 418.00, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 532.00, *m/z* obliczone [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 420.11, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 418.09, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 532.09; *t_R* = 10.24 min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.27 (1H, s), 10.94 (1H, s), 8.55 (2H, s), 8.50-8.28 (3H, bs), 8.41 (1H, s), 8.35 (3H, bs), 7.40-7.34 (4H, m), 7.34-7.30 (1H, m), 4.26-4.20 (1H, m), 3.28-3.23 i 3.11-3.04 (2H, 2m).

Ac-F-72, Ac-NH-L-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*S*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid; zsyntezowano wg procedury F; biały osad, wydajność reakcji 67%, czystość = 96% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}+\text{H}$]⁺: 462.10, *m/z* obliczone [$\text{M}+\text{H}$]⁺: 462.10; *t_R* = 16.06 min;

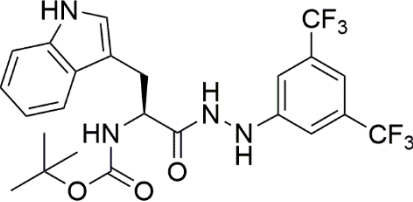
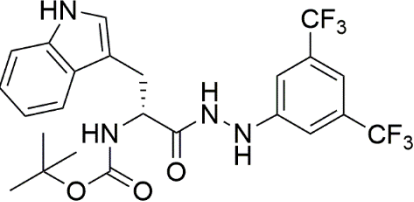
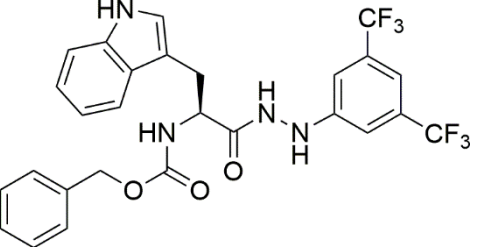
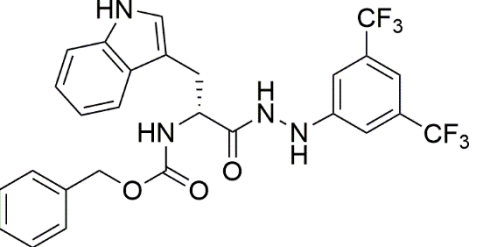
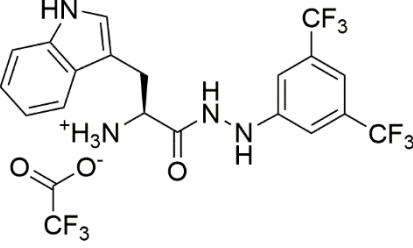
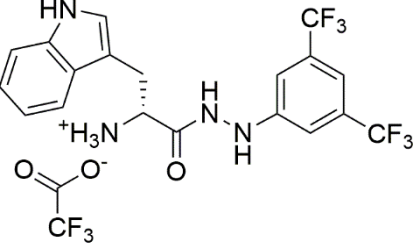
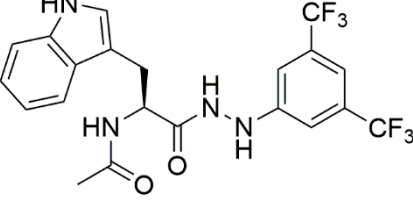
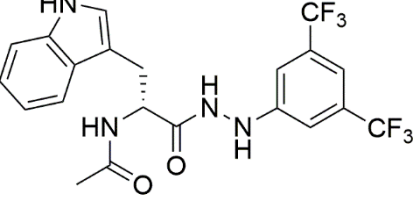
^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 10.99 (1H, s), 10.44 (1H, s), 8.54 (2H, s), 8.40 (1H, s), 8.26 (1H, d, ³*J_{HH}* = 8.4 Hz), 7.35-7.28 (4H, m), 7.24-7.19 (1H, m), 4.70-4.64 (1H, m), 3.12-3.06 i 2.85-2.78 (2H, 2m), 1.77 (3H, s).

Ac-dF-72, Ac-NH-D-Phe-NH-NH-C(=O)-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(3,5-bis(trifluorometylo)benzoilo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid; zsyntezowano wg procedury F; biały osad, wydajność reakcji 72%, czystość = 97% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}+\text{H}$]⁺: 462.05, *m/z* obliczone [$\text{M}+\text{H}$]⁺: 462.10; *t_R* = 16.05 min;

^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 10.99 (1H, s), 10.44 (1H, s), 8.54 (2H, s), 8.40 (1H, s), 8.26 (1H, d, ³*J_{HH}* = 8.4 Hz), 7.35-7.27 (4H, m), 7.24-7.19 (1H, m), 4.70-4.64 (1H, m), 3.12-3.06 i 2.85-2.78 (2H, 2m), 1.77 (3H, s).

VII.3.5. Seria W-73

Tabela 45. Wzory strukturalne związków z serii W-73.

<p>Boc-W-73</p> 	<p>Boc-dW-73</p> 
<p>Z-W-73</p> 	<p>Z-dW-73</p> 
<p>H-W-73</p> 	<p>H-dW-73</p> 
<p>Ac-W-73</p> 	<p>Ac-dW-73</p> 

Boc-W-73, Boc-L-Trp-NH-NH-3,5(CF₃)₂Ph, (S)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; żółto-pomarańczowy osad, czystość = 98 % (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 531.10, *m/z* obliczone [M+H]⁺: 531.18; *t_R* = 19.63 min.

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 10.82 (1H, s), 10.20 (1H, s), 8.66 (1H, s), 7.64 (1H, d, ²*J*_{HH} = 7.8 Hz), 7.33 (1H, d, ²*J*_{HH} = 7.8 Hz), 7.27 (2H, s), 7.24 (1H, s), 7.20 (1H, d, ²*J*_{HH} =

1.8 Hz), 7.08-7.04 (2H, prawdopodobnie nałożony dublet i dublet dubletów), 6.99 (1H, dd, $^2J_{\text{HH}} = 7.2, 7.2$ Hz), 4.30-4.20 (1H, m), 3.12-3.01 i 3.00-2.89 (2H, 2m), 1.31 i 1.19 (9H, 2s).

Boc-dW-73, Boc-D-Trp-NH-NH-3,5(CF₃)₂Ph, (R)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; żółto-pomarańczowy osad, wydajność reakcji 36%, czystość > 95% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 531.15, m/z obliczone [M+H]⁺: 531.18; t_R = 19.53 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 10.83 (1H, s), 10.21 (1H, s), 8.67 (1H, s), 7.64 (1H, d, $^2J_{\text{HH}} = 7.8$ Hz), 7.33 (1H, d, $^2J_{\text{HH}} = 8.4$ Hz), 7.28 (2H, s), 7.24 (1H, s), 7.21 (1H, s), 7.08-7.04 (2H, prawdopodobnie nałożony dublet i dublet dubletów), 6.99 (1H, dd, $^2J_{\text{HH}} = 7.2, 7.2$ Hz), 4.32-4.21 (1H, m), 3.13-3.02 i 2.99-2.86 (2H, 2m), 1.31 i 1.19 (9H, 2s).

Z-W-73, Cbz-L-Trp-NH-NH-3,5(CF₃)₂Ph, (S)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; żółto-brązowy osad, wydajność reakcji 70%, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 565.10, m/z obliczone [M+H]⁺: 565.17; t_R = 19.49 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.83 (1H, s), 10.30 (1H, s), 8.69 (1H, s), 7.68 (1H, d, $^2J_{\text{HH}} = 7.5$ Hz), 7.64 (1H, d, $^2J_{\text{HH}} = 7.5$ Hz), 7.37-7.17 (10H, m), 7.07 (1H, dd, $^2J_{\text{HH}} = 8.0, 6.5$ Hz), 6.99 (1H, dd, $^2J_{\text{HH}} = 7.5, 7.0$ Hz), 4.96 (2H, bs, prawdopodobnie nierozwiązany układ spinowy **AB**), 4.40-4.31 (1H, m), 3.14-3.07 i 3.02-2.93 (2H, 2m).

Z-dW-73, Cbz-D-Trp-NH-NH-3,5(CF₃)₂Ph, (R)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; żółto-brązowy osad, wydajność reakcji 34%, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 565.10, m/z obliczone [M+H]⁺: 565.17; t_R = 19.49 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.84 (1H, s), 10.29 (1H, s), 8.68 (1H, s), 7.66 (1H, d, $^2J_{\text{HH}} = 7.5$ Hz), 7.65 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.38-7.25 (8H, m), 7.20 (1H, s), 7.07 (1H, dd, $^2J_{\text{HH}} = 8.0, 7.0$ Hz), 6.99 (1H, dd, $^2J_{\text{HH}} = 8.0, 7.0$ Hz), 4.97 (2H, bs, prawdopodobnie nierozwiązany układ spinowy **AB**), 4.41-4.31 (1H, m), 3.15-3.08 i 3.02-2.94 (2H, 2m).

H-W-73, TFA*NH₂-L-Trp-NH-NH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (S)-1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminiowy; zsyntezowano wg procedury E; żółto-pomarańczowy osad, czystość = 95% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 431.05, [M_{zasada}-H]⁻: 429.00, [M_{sól}-H]⁻: 543.00, m/z obliczone [M_{zasada}+H]⁺: 431.13, [M_{zasada}-H]⁻: 429.11, [M_{sól}-H]⁻: 543.11; t_R = 10.64 min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.04 (1H, s), 10.57 (1H, s), 8.80 (1H, s), 8.24 (3H, bs), 7.70 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.38 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.32 (1H, s), 7.26 (2H, s), 7.22 (1H, s), 7.11 (1H, dd, $^2J_{\text{HH}} = 7.5, 7.0$ Hz), 7.04 (1H, dd, $^2J_{\text{HH}} = 7.5, 7.0$ Hz), 4.18-4.09 (1H, bs, prawdopodobnie nierozwiązany multiplet **ABX**), 3.35-3.29 i 3.18-3.08 (2H, 2m, nakładający się z sygnałem od HDO).

H-dW-73, TFA*NH₂-D-Trp-NH-NH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminiowy; zsyntezowano wg procedury E; żółto-pomarańczowy osad, wydajność reakcji 77%, czystość = 95% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 431.05, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 429.05, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 543.00, *m/z* obliczone [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 431.13, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 429.11, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 543.11; *t_R* = 10.64 min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.05 (1H, s), 11.05 (1H, s), 8.81 (1H, s), 8.28 (3H, bs), 7.70 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.38 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.32 (1H, s), 7.26 (2H, s), 7.22 (1H, s), 7.11 (1H, dd, $^2J_{\text{HH}} = 8.0, 7.0$ Hz), 7.03 (1H, dd, $^2J_{\text{HH}} = 7.5, 7.5$ Hz), 4.19-4.13 (1H, bs, prawdopodobnie nierozwiązany multiplet **ABX**), 3.33-3.30 i 3.18-3.10 (2H, 2m, nakładający się z sygnałem od HDO).

Ac-W-73, Ac-NH-L-Trp-NH-NH-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)acetamid; zsyntezowano wg procedury F; żółto-pomarańczowy osad, wydajność reakcji 80%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}+\text{H}$]⁺: 473.10, *m/z* obliczone [$\text{M}+\text{H}$]⁺: 473.13; *t_R* = 16.64 min;

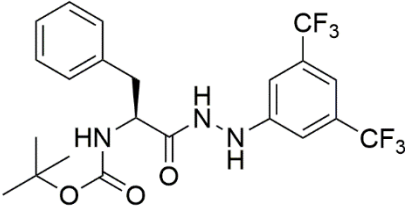
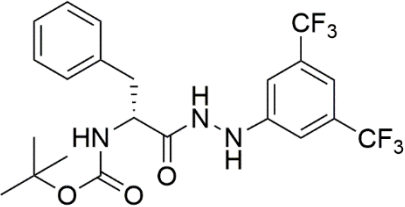
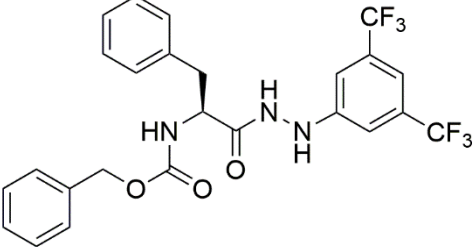
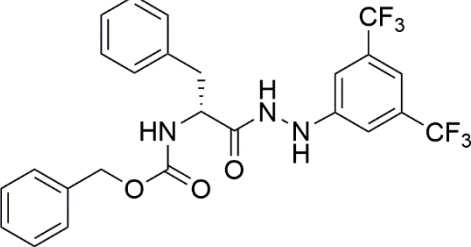
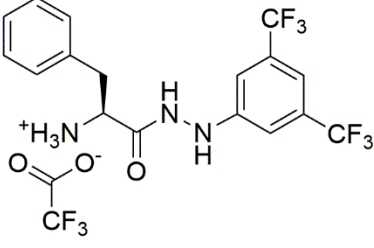
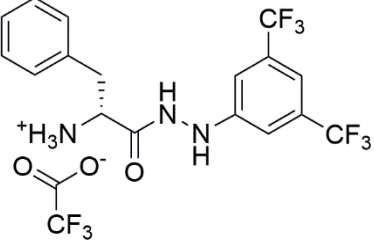
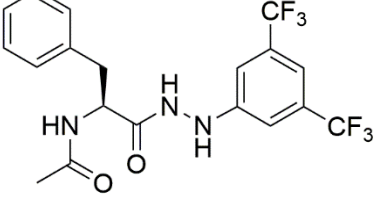
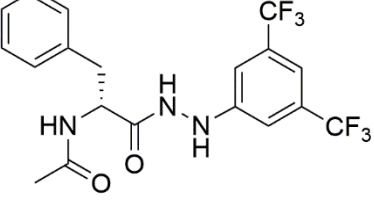
^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.83 (1H, s), 10.22 (1H, s), 8.65 (1H, s), 8.27 (1H, d, $^2J_{\text{HH}} = 7.5$ Hz), 7.60 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.33 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.25 (1H, s), 7.20 (2H, s), 7.18 (1H, s), 7.06 (1H, dd, $^2J_{\text{HH}} = 7.5, 7.5$ Hz), 6.98 (1H, dd, $^2J_{\text{HH}} = 8.0, 7.0$ Hz), 4.60 (1H, m), 3.15-3.09 i 2.98-2.91 (2H, 2m), 1.79 (3H, s).

Ac-dW-73, Ac-NH-D-Trp-NH-NH-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)acetamid; zsyntezowano wg procedury F; żółto-pomarańczowy osad, wydajność reakcji 85%, czystość > 98 % (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}+\text{H}$]⁺: 473.10, *m/z* obliczone [$\text{M}+\text{H}$]⁺: 473.13; *t_R* = 16.64 min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.82 (1H, s), 10.21 (1H, s), 8.64 (1H, s), 8.26 (1H, d, $^2J_{\text{HH}} = 7.5$ Hz), 7.61 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.33 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.26 (1H, s), 7.21 (2H, s), 7.18 (1H, s), 7.07 (1H, dd, $^2J_{\text{HH}} = 8.0, 7.5$ Hz), 6.99 (1H, dd, $^2J_{\text{HH}} = 7.5, 7.5$ Hz), 4.61-4.55 (1H, m), 3.17-3.09 i 2.99-2.90 (1H, 2m), 1.80 (3H, s).

VII.3.6. Seria F-73

Tabela 46. Wzory strukturalne związków z serii F-73.

<p>Boc-F-73</p> 	<p>Boc-dF-73</p> 
<p>Z-F-73</p> 	<p>Z-dF-73</p> 
<p>H-F-73</p> 	<p>H-dF-73</p> 
<p>Ac-F-73</p> 	<p>Ac-dF-73</p> 

Boc-F-73, Boc-L-Phe-NH-NH-3,5(CF₃)₂Ph, (S)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 70%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 490.10, m/z obliczone [M-H]⁻: 490.15 ; t_R = 20.08 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.11 (1H, s), 8.63 (1H, s), 7.37-7.13 (9H, m), 4.24-4.17 (1H, m), 2.98-2.90 i 2.89-2.78 (2H, 2m), 1.33 (9H, s).

Boc-dF-73, Boc-D-Phe-NH-NH-3,5(CF₃)₂Ph, (*R*)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 65%, czystość > 95% (HPLC-UV), ESI-MS *m/z* otrzymane [M-H]⁻: 490.10, *m/z* obliczone [M-H]⁻: 490.15 ; t_R = 20.05 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.11 (1H, s), 8.63 (1H, s), 7.33-7.17 (9H, m), 4.24-4.17 (1H, m), 2.97-2.90 i 2.88-2.79 (2H, 2m), 1.33 (9H, s).

Z-F-73, Cbz-L-Phe-NH-NH-3,5(CF₃)₂Ph, (*S*)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 69%, czystość = 95% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 526.10, *m/z* obliczone [M+H]⁺: 526.16; t_R = 20.02 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.22 (1H, s), 8.67 (1H, s), 7.77 (1H, d, ²J_{HH} = 7.5 Hz), 7.35-7.18 (13H, m), 4.33-4.24 (1H, m), 3.03-2.93 i 2.90-2.79 (2H, 2m), 4.97 (2H, **AB**, ²J_{HH} = 11.5 Hz, Δ*v*_{AB}=14.2 Hz).

Z-dF-73, Cbz-D-Phe-NH-NH-3,5(CF₃)₂Ph, (*R*)-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 43%, czystość = 95% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 526.10, *m/z* obliczone [M+H]⁺: 526.16; t_R = 19.93 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.21 (1H, s), 8.67 (1H, s), 7.76 (1H, d, ²J_{HH} = 8.0 Hz), 7.38-7.08 (13H, m), 4.35-4.25 (1H, m), 3.04-2.93 i 2.91-2.81 (2H, 2m), 4.98 (2H, **AB**, ²J_{HH} = 12.5 Hz, Δ*v*_{AB}=14.3 Hz).

H-F-73, TFA*NH₂-L-Phe-NH-NH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*S*)-1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-aminowy; zsyntezowano wg procedury E; biały osad, wydajność reakcji 80%, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M_{zasada}+H]⁺: 392.05, [M_{zasada}-H]⁻: 390.05, [M_{sól}-H]⁻: 504.00, *m/z* obliczone [M_{zasada}+H]⁺: 392.12, [M_{zasada}-H]⁻: 390.10, [M_{sól}-H]⁻: 504.10; t_R = 10.60 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.47 (1H, s), 8.76 (1H, s), 8.37 (3H, bs), 7.38-7.32 (2H, m), 7.32-7.25 (4H, m), 7.19 (2H, s), 4.17-4.12 (1H, m), 3.16-3.10 i 3.09-3.02 (2H, 2m).

H-dF-73, TFA*NH₂-D-Phe-NH-NH-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-aminowy; zsyntezowano wg procedury E; biały osad, wydajność reakcji 78%, czystość = 98% (HPLC-UV), ESI-MS

m/z otrzymane $[M_{zasada+H}]^+$: 392.05, $[M_{zasada-H}]^-$: 390.05, $[M_{sól-H}]^-$: 504.00, m/z obliczone $[M_{zasada+H}]^+$: 392.12, $[M_{zasada-H}]^-$: 390.10, $[M_{sól-H}]^-$: 504.10; $t_R = 10.60$ min;

1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.41 (1H, s), 8.73 (1H, s), 8.29 (3H, bs), 7.35-7.31 (2H, m), 7.30-7.24 (4H, m), 7.17 (2H, s), 4.14-4.08 (1H, m), 3.15-3.08 i 3.07-3.00 (2H, 2m).

Ac-F-73, Ac-NH-L-Phe-NH-NH-3,5(CF₃)₂Ph, (*S*)-*N*-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid;

zsyntezowano wg procedury F; biały osad, wydajność reakcji 87%, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 434.05, m/z obliczone $[M+H]^+$: 434.13; $t_R = 17.09$ min;

1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.12 (1H, s), 8.61 (1H, s), 8.28 (1H, d, $^2J_{HH} = 7.5$ Hz), 7.28-7.22 (5H, m), 7.21-7.17 (1H, m), 7.15 (2H, s), 4.54-4.46 (1H, m), 3.01-2.93 i 2.85-2.77 (2H, 2m), 1.76 (3H, s).

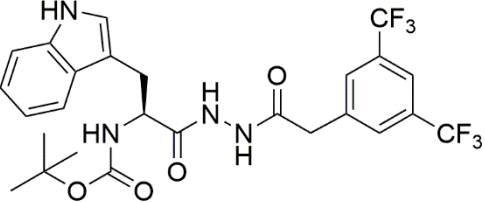
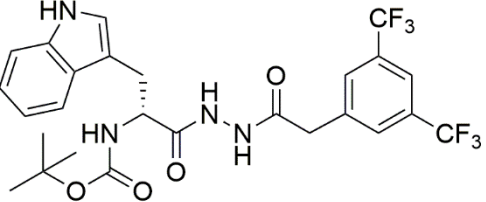
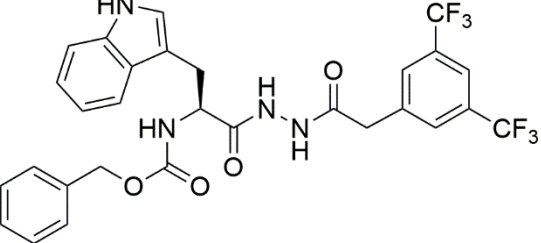
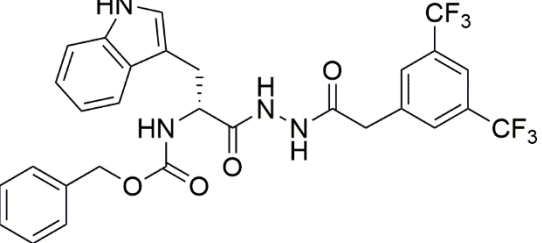
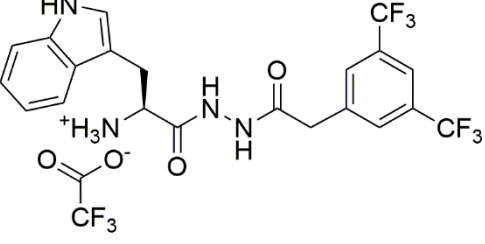
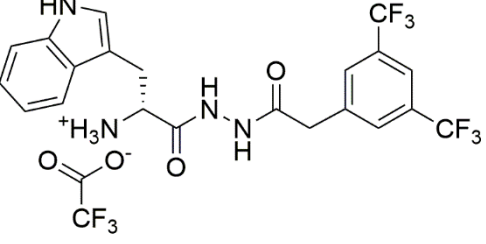
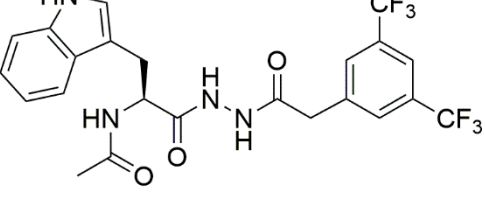
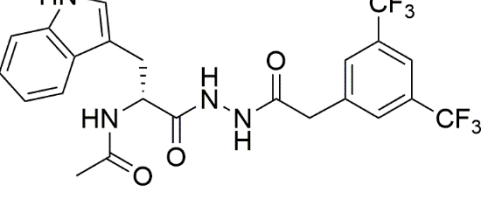
Ac-dF-73, Ac-NH-D-Phe-NH-NH-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(3,5-bis(trifluorometylo)fenylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid;

zsyntezowano wg procedury F; biały osad, wydajność reakcji 40%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 434.10, m/z obliczone $[M+H]^+$: 434.13; $t_R = 17.04$ min;

1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 10.14 (1H, s), 8.63 (1H, s), 8.31 (1H, d, $^2J_{HH} = 6.5$ Hz), 7.31-7.22 (5H, m), 7.22-7.17 (1H, m), 7.16 (2H, s), 4.54-4.47 (1H, m), 3.01-2.93 i 2.85-2.78 (2H, 2m), 1.77 (3H, s).

VII.3.7. Seria W-74

Tabela 47. Wzory strukturalne związków z serii W-74.

<p>Boc-W-74</p> 	<p>Boc-dW-74</p> 
<p>Z-W-74</p> 	<p>Z-dW-74</p> 
<p>H-W-74</p> 	<p>H-dW-74</p> 
<p>Ac-W-74</p> 	<p>Ac-dW-74</p> 

NH₂-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, 2-(3,5-bis(trifluorometylo)fenylo)acetohydrazyd;
 zsyntezowano wg procedury A; biały osad, wydajność reakcji 83%, ESI-MS m/z otrzymane
 [M+H]⁺: 287.10, m/z obliczone [M+H]⁺: 287.18; t_R = 14.02 min;
¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 9.27 (1H, s), 7.93 (3H, s, prawdopodobnie dwa
 nakładające się singlety), 4.26 (2H, s), 3.59 (2H, s).

Boc-W-74, Boc-L-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (S)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 76%, czystość > 97% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 571.10, m/z obliczone [M-H]⁻: 571.17; t_R = 18.45 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.78 (1H, s), 10.12 (bs), 8.02 (1H, s), 7.99 (1H, s), 7.63-7.55 (1H, m), 7.33-7.25 (1H, m), 7.14 (1H, s), 7.06-6.98 (1H, m), 6.97-6.91 (1H, m), 6.78 (1H, m), 4.30-4.09 (1H, m), 3.71 (2H, s), 3.12-3.02 i 2.92-2.84 (2H, 2m), 1.27 (9H, s).

Boc-dW-74, Boc-D-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (R)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; kremowy osad, wydajność reakcji 70%, czystość > 97% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 571.05, m/z obliczone [M-H]⁻: 571.17; t_R = 18.46 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.78 (1H, s), 10.03 (bs), 8.02 (2H, s), 7.99 (1H, s), 7.67-7.54 (1H, m), 7.33-7.23 (1H, m), 7.15 (1H, s), 7.05-6.99 (1H, m), 6.96-6.90 (1H, m), 6.79 (1H, m), 4.32-4.06 (1H, m), 3.72 (2H, s), 3.12-3.03 i 2.92-2.83 (2H, 2m), 1.27 (9H, s).

Z-W-74, Cbz-L-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (S)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 78%, czystość = 97% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 607.10, m/z obliczone [M+H]⁺: 607.18; t_R = 18.55 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.79 (1H, s), 10.26 (1H, s), 10.23 (1H, s), 8.03 (2H, s), 8.00 (1H, s), 7.66 (1H, d, ²J_{HH} = 8.0 Hz), 7.50-6.92 (10H, m), 4.98-4.83 (2H, m), 4.43-4.30 (1H, m), 3.77 (2H, s), 3.15-3.03 i 2.97-2.84 (2H, 2m).

Z-dW-74, Cbz-D-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, ((R)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 62%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 607.00, m/z obliczone [M+H]⁺: 607.18; t_R = 18.55 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.79 (1H, s), 10.24 (2H, bs), 8.02 (2H, s), 7.99 (1H, s), 7.65 (1H, d, ²J_{HH} = 8.5 Hz), 7.50-6.91 (10H, m), 4.97-4.86 (2H, m), 4.40-4.30 (1H, m), 3.76 (2H, s), 3.13-3.04 i 2.96-2.86 (2H, 2m).

H-W-74, TFA*NH₂-L-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (S)-1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-

aminowy; zsyntezowano wg procedury E; biały osad, wydajność reakcji 82%, czystość > 98 % (HPLC-UV), ESI-MS m/z otrzymane $[M_{zasada+H}]^+$: 473.05, $[M_{zasada-H}]^-$: 471.05, $[M_{sól-H}]^-$: 585.00, m/z obliczone $[M_{zasada+H}]^+$: 473.14, $[M_{zasada-H}]^-$: 471.12, $[M_{sól-H}]^-$: 585.12; t_R = 10.45 min;

1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.01 (1H, s), 10.76 (1H, s), 10.57 (1H, s), 8.15 (3H, s), 8.03 (2H, s), 7.01 (1H, s), 7.67 (1H, d, $^2J_{HH}$ = 8.0 Hz), 7.36 (1H, d, $^2J_{HH}$ = 8.0 Hz), 7.25 (1H, d, $^2J_{HH}$ = 1.5 Hz), 7.09 (1H, dd, $^2J_{HH}$ = 8.0, 7.5 Hz), 7.00 (1H, dd, $^2J_{HH}$ = 7.5, 7.5 Hz), 4.07-3.98 (1H, bs, prawdopodobnie nierozwiązany multiplet), 3.83 (2H, s), 3.31-3.24 i 3.13-3.05 (2H, 2m).

H-dW-74, TFA*NH₂-D-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-aminowy; zsyntezowano wg procedury E; biały osad, wydajność reakcji 83%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M_{zasada+H}]^+$: 473.10, $[M_{zasada-H}]^-$: 471.00, $[M_{sól-H}]^-$: 585.00, m/z obliczone $[M_{zasada+H}]^+$: 473.14, $[M_{zasada-H}]^-$: 471.12, $[M_{sól-H}]^-$: 585.12; t_R = 10.45 min;

1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.99 (1H, s), 10.73 (1H, s), 10.54 (1H, s), 8.13 (3H, s), 8.01 (2H, s), 7.98 (1H, s), 7.67 (1H, d, $^2J_{HH}$ = 8.0 Hz), 7.33 (1H, d, $^2J_{HH}$ = 8.0 Hz), 7.22 (1H, d, $^2J_{HH}$ = 1.5 Hz), 7.06 (1H, dd, $^2J_{HH}$ = 7.5, 7.5 Hz), 6.97 (1H, dd, $^2J_{HH}$ = 8.0, 7.0 Hz), 4.06 (1H, bs, prawdopodobnie nierozwiązany multiplet), 3.80 (2H, s), 3.30-3.20 i 3.12-3.02 (2H, 2m).

Ac-W-74, Ac-NH-L-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*S*)-*N*-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-ylo)acetamid; zsyntezowano wg procedury F; biały osad, wydajność reakcji 87%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 515.05, m/z obliczone $[M+H]^+$: 515.15; t_R = 15.88 min;

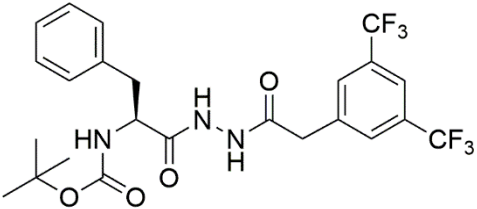
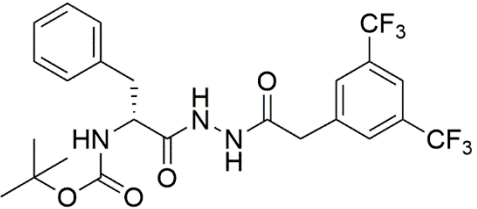
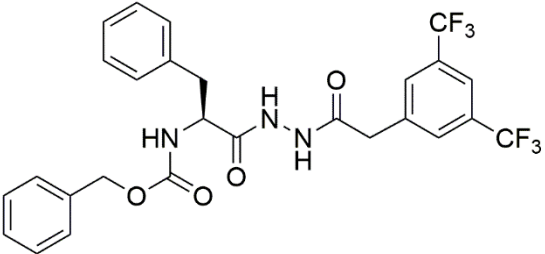
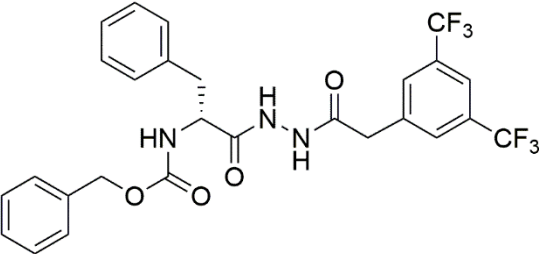
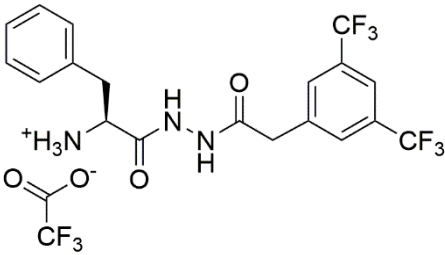
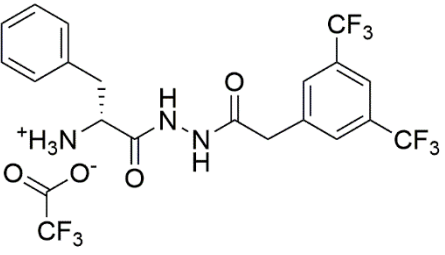
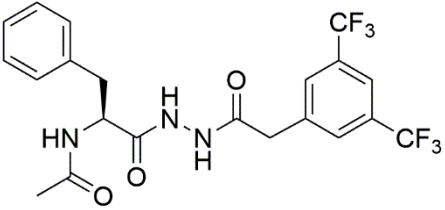
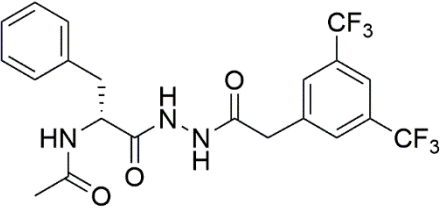
1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.77 (1H, s), 10.16 (2H, bs), 8.10 (1H, d, $^2J_{HH}$ = 9.5 Hz), 8.01 (2H, 1s), 7.99 (1H, s), 7.60 (1H, d, $^2J_{HH}$ = 9.0 Hz), 7.29 (1H, d, $^2J_{HH}$ = 7.5 Hz), 7.14 (1H, s), 7.03 (1H, dd, $^2J_{HH}$ = 7.5, 7.0 Hz), 6.94 (1H, dd, $^2J_{HH}$ = 8.0, 7.5 Hz), 4.64-4.47 (1H, m), 3.74 (2H, s), 3.14-3.06 (2H, 2m), 1.73 (3H, s).

Ac-dW-74, Ac-NH-D-Trp-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-3-(1H-indol-3-ilo)-1-oksopropan-2-ylo)acetamid; zsyntezowano wg procedury F; biały osad, wydajność reakcji 66%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 515.05, m/z obliczone $[M+H]^+$: 515.15; t_R = 15.89 min;

^1H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.77 (1H, s), 10.15 (2H, bs), 8.10 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 8.01 (2H, 1s), 7.99 (1H, s), 7.60 (1H, d, $^2J_{\text{HH}} = 8.5$ Hz), 7.29 (1H, d, $^2J_{\text{HH}} = 8.0$ Hz), 7.15 (1H, s), 7.03 (1H, dd, $^2J_{\text{HH}} = 9.0, 6.5$ Hz), 6.95 (1H, dd, $^2J_{\text{HH}} = 8.0, 7.5$ Hz), 4.61-4.54 (1H, m), 3.75 (3H, s), 3.13-3.05 i 2.93-2.84 (2H, 2m), 1.73 (3H, s).

VII.3.8. Seria F-74

Tabela 48. Wzory strukturalne związków z serii F-74.

<p>Boc-F-74</p> 	<p>Boc-dF-74</p> 
<p>Z-F-74</p> 	<p>Z-dF-74</p> 
<p>H-F-74</p> 	<p>H-dF-74</p> 
<p>Ac-F-74</p> 	<p>Ac-dF-74</p> 

Boc-F-74, Boc-L-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (S)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 53%, ESI-MS m/z otrzymane [M-H]⁻: 532.10, m/z obliczone [M-H]⁻: 532.16 ; t_R = 18.85 min;
¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.25 (1H, s), 10.09 (1H, s), 7.99 (2H, s), 7.96 (1H, s), 7.30-7.10 (5H, m), 6.93 (1H, d, ²J_{HH} = 8.5 Hz), 4.22-4.10 (1H, m), 3.73 (1H, s), 2.98-2.85 i 2.77-2.64 (2H, 2m), 1.25 i 1.18 (9H, 2s).

Boc-dF-74, Boc-D-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*R*)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian *tert*-butylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 69%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M-H]⁻: 532.10, *m/z* obliczone [M-H]⁻: 532.16 ; t_R = 18.79 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.28 (1H, s), 10.11 (1H, s), 8.02 (1H, s), 7.99 (1H, s), 7.35-7.12 (5H, m), 6.96 (1H, d, ²J_{HH} = 9.0 Hz), 4.25-4.14 (1H, m), 3.76 (1H, s), 3.01-2.89 i 2.79-2.66 (2H, 2m), 1.28 i 1.20 (9H, 2s).

Z-F-74, Cbz-L-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*S*)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 57%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 568.10, *m/z* obliczone [M+H]⁺: 568.17; t_R = 18.99 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.27 (1H, s), 10.19 (1H, s), 8.02 (2H, s), 7.99 (1H, s), 7.58 (1H, d, ²J_{HH} = 8.0 Hz), 7.37-7.14 (10H, m), 4.92 (2H, s), 4.38-4.25 (1H, m), 3.76 (2H, s), 3.08-2.94 i 2.82-2.69 (2H, 2m).

Z-dF-74, Cbz-D-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*R*)-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)karbaminian benzylu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 30%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 568.10, *m/z* obliczone [M+H]⁺: 568.17; t_R = 19.01 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.26 (1H, s), 10.19 (1H, s), 8.02 (2H, s), 7.99 (1H, s), 7.58 (1H, d, ²J_{HH} = 8.5 Hz), 7.35-7.10 (10H, m), 4.92 (2H, s), 4.39-4.25 (1H, m), 3.76 (2H, s), 3.09-2.94 i 2.83-2.70 (2H, 2m).

H-F-74, TFA*NH₂-L-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*S*)-1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-aminiowy zsyntezowano wg procedury E; biały osad, czystość > 98 % (HPLC-UV), ESI-MS *m/z* otrzymane [M_{zasada}+H]⁺: 434.05, [M_{zasada}-H]⁻: 432.00, [M_{sól}-H]⁻: 545.95, *m/z* obliczone [M_{zasada}+H]⁺: 434.13, [M_{zasada}-H]⁻: 432.11, [M_{sól}-H]⁻: 546.11; t_R = 10.25 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.63 (1H, s), 10.55 (1H, s), 8.23 (1H, s), 8.01 (2H, s), 8.00 (1H, s), 7.36-7.21 (5H, m), 4.10-3.98 (1H, m), 3.80 (2H, s), 3.16-3.07 i 3.00-2.91 (2H, 2m).

H-dF-74, TFA*NH₂-D-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-aminiowy;

zsyntezowano wg procedury E; biały osad, wydajność reakcji 80%, czystość = 97% (HPLC-UV), ESI-MS m/z otrzymane $[M_{zasada+H}]^+$: 434.10, $[M_{zasada-H}]^-$: 432.05, $[M_{sól-H}]^-$: 546.05, m/z obliczone $[M_{zasada+H}]^+$: 434.13, $[M_{zasada-H}]^-$: 432.11, $[M_{sól-H}]^-$: 546.11; $t_R = 10.35$ min; 1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.64 (1H, s), 10.54 (1H, s), 8.25 (1H, s), 8.02 (2H, s), 8.00 (1H, s), 7.37-7.20 (5H, m), 4.11-3.98 (1H, m), 3.81 (2H, s), 3.17-3.08 i 3.02-2.91 (2H, 2m).

Ac-F-74, Ac-NH-L-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*S*)-*N*-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid;

zsyntezowano wg procedury F; biały osad, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 476.10, m/z obliczone $[M+H]^+$: 476.14; $t_R = 16.08$ min;

1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.18 (1H, s), 10.10 (1H, s), 8.12 (1H, d, $^2J_{HH} = 9.0$ Hz), 7.99 (2H, s), 7.96 (1H, s), 7.27-7.12 (5H, m), 4.56-4.48 (1H, m), 3.73 (2H, s), 3.01-2.90 i 2.75-2.65 (2H, 2m), 1.67 (3H, s).

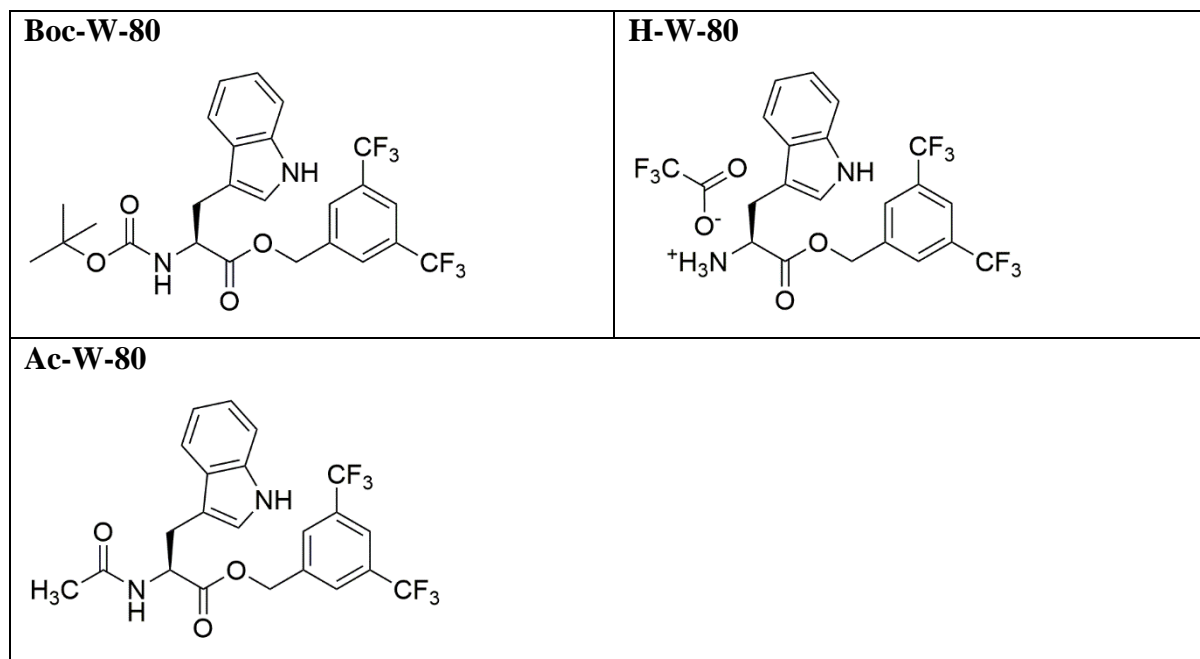
Ac-dF-74, Ac-NH-D-Phe-NH-NH-C(=O)-CH₂-3,5(CF₃)₂Ph, (*R*)-*N*-(1-(2-(2-(3,5-bis(trifluorometylo)fenylo)acetylo)hydrazynylo)-1-okso-3-fenylopropan-2-ylo)acetamid;

zsyntezowano wg procedury F; biały osad, wydajność reakcji 40%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 476.05, m/z obliczone $[M+H]^+$: 476.14; $t_R = 16.05$ min;

1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.18 (1H, s), 10.10 (1H, s), 8.12 (1H, d, $^2J_{HH} = 9.0$ Hz), 7.99 (2H, s), 7.96 (1H, s), 7.27-7.12 (5H, m), 4.56-4.48 (1H, m), 3.73 (2H, s), 3.01-2.90 i 2.75-2.65 (2H, 2m), 1.67 (3H, s).

VII.3.9. Seria W-80

Tabela 49. Wzory strukturalne związków z serii W-80.



Boc-W-80, Boc-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (*tert*-butyloksykarbonylo)-L-tryptofanu; zsyntezowano wg procedury G z zamianą rozpuszczalnika na acetonitryl (ACN); biały osad, wydajność reakcji 72%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 531.10, m/z obliczone [M+H]⁺: 531.17; t_R = 21.46 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm) 10.89 i 10.82 (1H, s), 8.05 i 8.03 (1H, s), 7.97 i 7.95 (2H, s), 7.45 (1H, d, *J* = 7.8 Hz), 7.36 (1H, m), 7.30 (1H, m), 7.14 (1H, bs), 7.03 (1H, m), 6.94 (1H, m), 5.28 (1H, d, *J* = 13.5 Hz), 5.15 (1H, d, *J* = 13.5 Hz), 4.23 (1H, m), 3.12 (1H, m), 3.03 (1H, m), 1.28 i 1.10 (9H, 2s). Niektóre sygnały wykazują podwojenie, które można przypisać obecności dwóch konformerów, prawdopodobnie związanych z zahamowaną rotacją wiązania C–N w ugrupowaniu karbaminianowym. Podobne obserwacje opisano dla aminokwasów zabezpieczonych grupą Boc, np. Boc-L-Phe [317].

¹³C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 172.4, 155.9, 140.0, 136.5, 130.7 (q, ²*J*_{CF} = 32.5 Hz), 128.7 (szeroki, prawdopodobnie nierozwiązany kwartet o ³*J*_{CF} ~ 4 Hz), 127.4, 124.3, 123.7 (q, ¹*J*_{CF} = 270.8 Hz), 122.0 (szeroki, prawdopodobnie nierozwiązany kwartet o ³*J*_{CF} ~ 4 Hz), 121.4, 118.9, 118.3, 111.9, 110.0, 78.8, 64.8, 55.5, 28.4, 27.1.

H-W-80, TFA* $\text{NH}_2\text{-L-Trp-O-CH}_2\text{-3,5(CF}_3)_2\text{Ph}$, 2,2,2-trifluorooctan (*S*)-1-((3,5-bis(trifluorometylo)benzylo)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminiowy; otrzymano wg procedury E; biały proszek, wydajność reakcji 81%, czystość > 98% (HPLC-UV) ESI-MS *m/z* otrzymane [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 431.05, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 429.00, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 542.95, *m/z* obliczone [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 431.12, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 429.10, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 543.10; $t_{\text{R}} = 11.38$ min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.03 (1H, s), 8.52 (3H, s), 8.09 (1H, s), 7.99 (2H, s), 7.48 (1H, m), 7.35 (1H, m), 7.22 (1H, d, $J = 2.2$ Hz), 7.07 (1H, m), 6.97 (1H, m), 5.37 (1H, d, $J = 13.2$ Hz), 5.24 (1H, d, $J = 13.2$ Hz), 4.37 (1H, m), 3.30-3.26 (2H, m);

¹³C NMR (125 MHz, DMSO-*d*6) δ (ppm), 169.7, 158.5 (słaby, prawdopodobnie kwartet o $^2J_{\text{CF}} \sim 30$ Hz, anion TFA), 138.8, 136.7, 130.8 (q, $^2J_{\text{CF}} = 32.5$ Hz), 129.4 (szeroki, prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 127.3, 125.3, 123.7 (q, $^1J_{\text{CF}} = 271$ Hz), 122.6 (szeroki, prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz, pokrywający się z sygnałem węgla grup CF_3), 121.6, 119.1, 118.3, 117.8 (słaby, prawdopodobnie kwartet o $^1J_{\text{CF}} \sim 300$ Hz, anion TFA), 112.7, 106.8, 66.0, 53.3, 26.8.

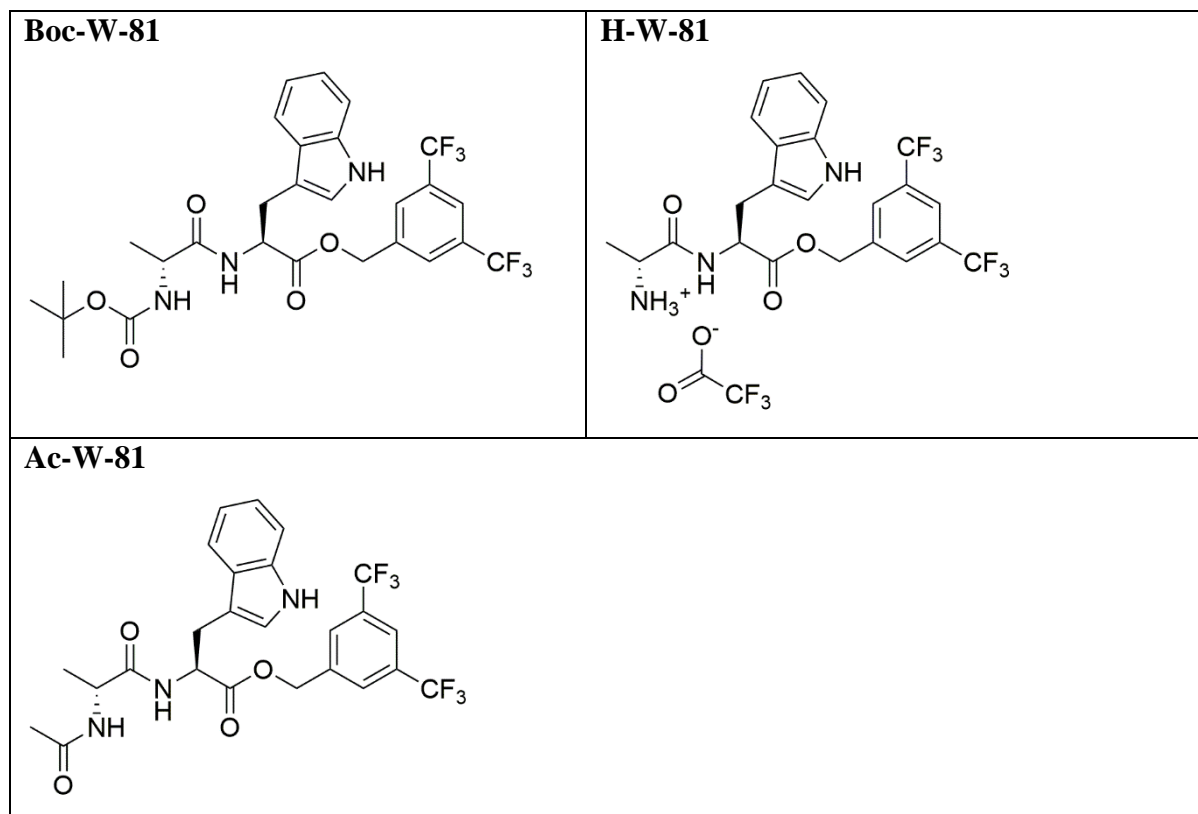
Ac-W-80, Ac-NH-L-Trp-O-CH₂-3,5(CF₃)₂Ph, (inna nazwa: L-732,138, opisany w literaturze: [139,140]); ester 3,5-bis(trifluorometylo)benzylowy *N*-acetylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 98%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}+\text{H}$]⁺: 473.10, *m/z* obliczone [$\text{M}+\text{H}$]⁺: 473.13; $t_{\text{R}} = 18.61$ min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.84 (1H, s), 8.41 (1H, d, $J = 6.9$ Hz), 8.05 (1H, s), 7.96 (2H, s), 7.47 (1H, m), 7.32 (1H, m), 7.16 (1H, d, $J = 2.2$ Hz), 7.06 (1H, m), 6.97 (1H, m), 5.28 (1H, d, $J = 13.5$ Hz), 5.18 (1H, d, $J = 13.5$ Hz), 4.53 (1H, m), 3.18 (1H, m), 3.07 (1H, m), 1.82 (3H, s);

¹³C NMR (125 MHz, DMSO-*d*6) δ (ppm), 172.4, 170.0, 140.0, 136.6, 130.7 (q, $^2J_{\text{CF}} = 32.5$ Hz), 128.6 (szeroki, prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 127.4, 124.1, 123.7 (q, $^1J_{\text{CF}} = 270.8$ Hz), 121.4, 122.0 (szeroki, prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 118.9, 118.3, 111.9, 109.8, 64.7, 54.0, 27.3, 22.6.

VII.3.10. Seria W-81

Tabela 50. Wzory strukturalne związków z serii W-81.



Boc-W-81, Boc-D-Ala-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (*tert*-butyloksykarbonylo)-D-alanylo-L-tryptofanu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 38%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 602.15, m/z obliczone [M+H]⁺: 602.21; t_R = 20.28 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.85 (1H, s), 8.22 (1H, d, *J* = 7.5 Hz), 8.05 (1H, s), 7.99 (2H, s), 7.46 (1H, m), 7.31 (1H, m), 7.13 (1H, d, *J* = 2.2 Hz), 7.05 (1H, m), 6.96 (1H, m), 6.72 (1H, d, *J* = 7.3 Hz), 5.29 (1H, d, *J* = 13.6 Hz), 5.20 (1H, d, *J* = 13.6 Hz), 4.56 (1H, m), 4.00 (1H, m), 3.21 (1H, m), 3.10 (1H, m), 1.34 (9H, s), 1.16 i 1.06 (3H, 2d, *J* = 6.9 Hz);

¹³C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 173.4, 172.1, 155.3, 139.8, 136.5, 130.6 (q, ²J_{CF} = 32.5 Hz), 128.9 (szeroki, prawdopodobnie nierozwiązany kwartet o ³J_{CF} ~ 4 Hz), 127.4, 124.7, 123.7 (q, ¹J_{CF} = 271.6 Hz), 121.4, 122.1 (szeroki, prawdopodobnie nierozwiązany kwartet o ³J_{CF} ~ 4 Hz), 118.8, 118.3, 111.9, 109.6, 78.5, 65.0, 53.6, 49.9, 28.6, 27.4, 16.7.

H-W-81, TFA*NH₂-D-Ala-L-Trp-O-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluoroocetan (*R*)-1-(((*S*)-1-((3,5-bis(trifluorometylo)benzyl)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-yl)amino)-1-oksopropano-2-aminiowy; otrzymano wg procedury E; biały proszek, wydajność reakcji

48%, czystość > 95% (HPLC-UV), ESI-MS m/z otrzymane $[M_{zasada+H}]^+$: 502.10, $[M_{zasada-H}]^-$: 500.05, $[M_{sól-H}]^-$: 614.00, m/z obliczone $[M_{zasada+H}]^+$: 502.16, $[M_{zasada-H}]^-$: 500.14, $[M_{sól-H}]^-$: 614.14; $t_R = 11.40$ min;

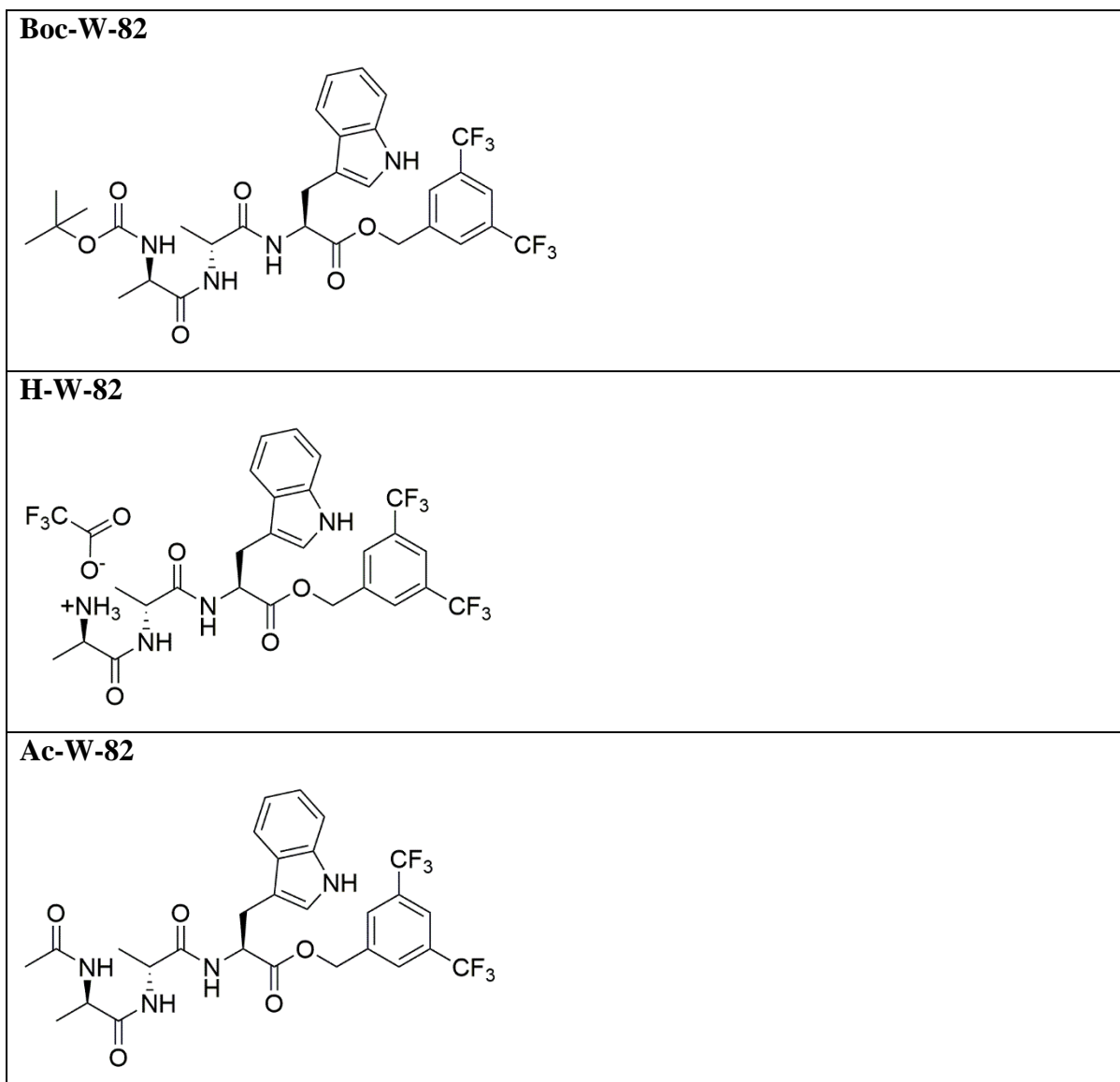
1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 10.84 (1H, s), 8.84 (1H, d, $J = 7.8$ Hz), 8.07 (1H, s), 7.99 (2H, s), 7.88 (3H, bs), 7.46 (1H, d, $J = 7.8$ Hz), 7.29 (1H, d, $J = 8.4$ Hz), 7.11 (1H, d, $J = 2.1$ Hz), 7.02 (1H, m), 6.93 (1H, m), 5.29 (1H, d, $J = 13.2$ Hz), 5.22 (1H, d, $J = 13.2$ Hz), 4.70 (1H, m), 3.77 (1H, m), 3.25 (1H, m), 3.08 (1H, m), 1.12 (3H, d, $J = 7.0$ Hz).

Ac-W-81, Ac-NH-D-Ala-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy *N*-acetylo-D-alanylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 61%, czystość > 97% (HPLC-UV), ESI-MS m/z otrzymane $[M+H]^+$: 544.15, m/z obliczone $[M+H]^+$: 544.17; $t_R = 17.72$ min;

1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 10.84 (1H, s), 8.29 (1H, d, $J = 7.5$ Hz), 8.03 (1H, s), 7.96 (2H, s), 7.90 (1H, d, $J = 7.5$ Hz), 7.44 (1H, d, $J = 7.8$ Hz), 7.28 (1H, m), 7.11 (1H, d, $J = 2.1$ Hz), 7.02 (1H, m), 6.92 (1H, m), 5.25 (1H, d, $J = 13.2$ Hz), 5.16 (1H, d, $J = 13.2$ Hz), 4.49 (1H, m), 4.26 (1H, m), 3.18 (1H, m), 3.06 (1H, m), 1.76 (3H, s), 1.01 (3H, d, $J = 7.1$ Hz).

VII.3.11. Seria W-82

Tabela 51. Wzory strukturalne związków z serii W-82.



Boc-W-82, Boc-D-Ala-D-Ala-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (*tert*-butyloksykarbonylo)-D-alanylo-D-alanylo-L-tryptofanu; zsyntezowano wg procedury B; biały osad, wydajność reakcji 38%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 673.15, *m/z* obliczone [M+H]⁺: 673.25; *t_R* = 19.60 min;

¹H NMR (600 MHz, DMSO-*d*₆) δ (ppm), 10.83 (1H, s), 8.32 (1H, d, *J* = 7.6 Hz), 8.03 (1H, s), 7.95 (2H, s), 7.70 (1H, d, *J* = 7.3 Hz), 7.43 (1H, m), 7.28 (1H, m), 7.12 (1H, d, *J* = 2.1 Hz), 7.02 (1H, m), 6.92 (1H, m), 6.47 (1H, bs), 5.25 (1H, d, *J* = 13.3 Hz), 5.15 (1H, d, *J* =

13.3 Hz), 4.50 (1H, m), 4.25 (1H, m), 3.90 (1H, m), 3.19 (1H, m), 3.05 (1H, m), 1.32 i 1.27 (9H, 2bs), 1.08 (3H, d, $J=7.1$ Hz), 1.03 (3H, d, $J = 7.1$ Hz);

^{13}C NMR (125 MHz, DMSO-*d*6) δ (ppm), 172.7, 172.6, 172.0, 155.6, 139.8, 136.5, 130.7 (q, $^2J_{\text{CF}} = 32.5$ Hz), 128.9 (szeroki, prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 127.4, 123.7 (q, $^2J_{\text{CF}} = 271.3$ Hz), 122.2 (szeroki, prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 121.4, 124.3, 118.8, 118.3, 111.9, 109.6, 78.6, 65.0, 53.6, 50.2, 48.3, 28.6, 27.3, 18.8, 18.3.

H-W-82, TFA*NH₂-D-Ala-D-Ala-L-Trp-O-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*R*)-1-(((*R*)-1-(((*S*)-1-((3,5-bis(trifluorometylo)benzylo)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)amino)-1-oksopropan-2-ylo)amino)-1-oksopropan-2-aminiowy; otrzymano wg procedury E; biały osad, wydajność reakcji 32%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 573.15, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 571.05, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 685.00, *m/z* obliczone [$\text{M}_{\text{zasada}}+\text{H}$]⁺: 573.20, [$\text{M}_{\text{zasada}}-\text{H}$]⁻: 571.18, [$\text{M}_{\text{sól}}-\text{H}$]⁻: 685.18; $t_{\text{R}} = 11.62$ min;

^1H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.85 (1H, s), 8.51 (1H, d, $J = 7.5$ Hz), 8.44 (1H, d, $J = 7.5$ Hz), 8.07 (1H, s), 7.99 (2H, s), 7.89 (3H, bs), 7.48 (1H, m), 7.32 (1H, m), 7.15 (1H, d, $J = 2.2$ Hz), 7.05 (1H, m), 6.96 (1H, m), 5.30 (1H, d, $J = 13.3$ Hz), 5.21 (1H, d, $J = 13.3$ Hz), 4.58 (1H, m), 4.39 (1H, m), 3.81 (1H, m), 3.24 (1H, m), 3.09 (1H, m), 1.26 (3H, d, $J = 7.0$ Hz), 1.09 (3H, d, $J = 7.1$ Hz);

^{13}C NMR (125 MHz, DMSO-*d*6) δ (ppm), 172.2, 172.0, 169.5, 158.3 (słaby, prawdopodobnie kwartet o $^2J_{\text{CF}} \sim 30$ Hz, anion TFA), 139.8, 136.5, 130.7 (q, $^2J_{\text{CF}} = 32.5$ Hz), 128.9 (prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 127.4, 124.3, 123.7 (q, $^2J_{\text{CF}} \sim 272$ Hz), 122.2 (prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4\text{Hz}$), 121.4, 118.8, 118.3, 111.9, 109.6, 65.0, 53.6, 48.5, 48.4, 27.5, 18.9, 17.6. Nie zaobserwowano sygnału pochodzącego od grupy CF₃ anionu trifluorooctanowego (oczekiwana wartość około 117.8 jako kwartet o $^1J_{\text{CF}} \sim 300$ Hz), co związane jest prawdopodobnie z dość krótkim czasem akwizycji widma dla małej próbki.

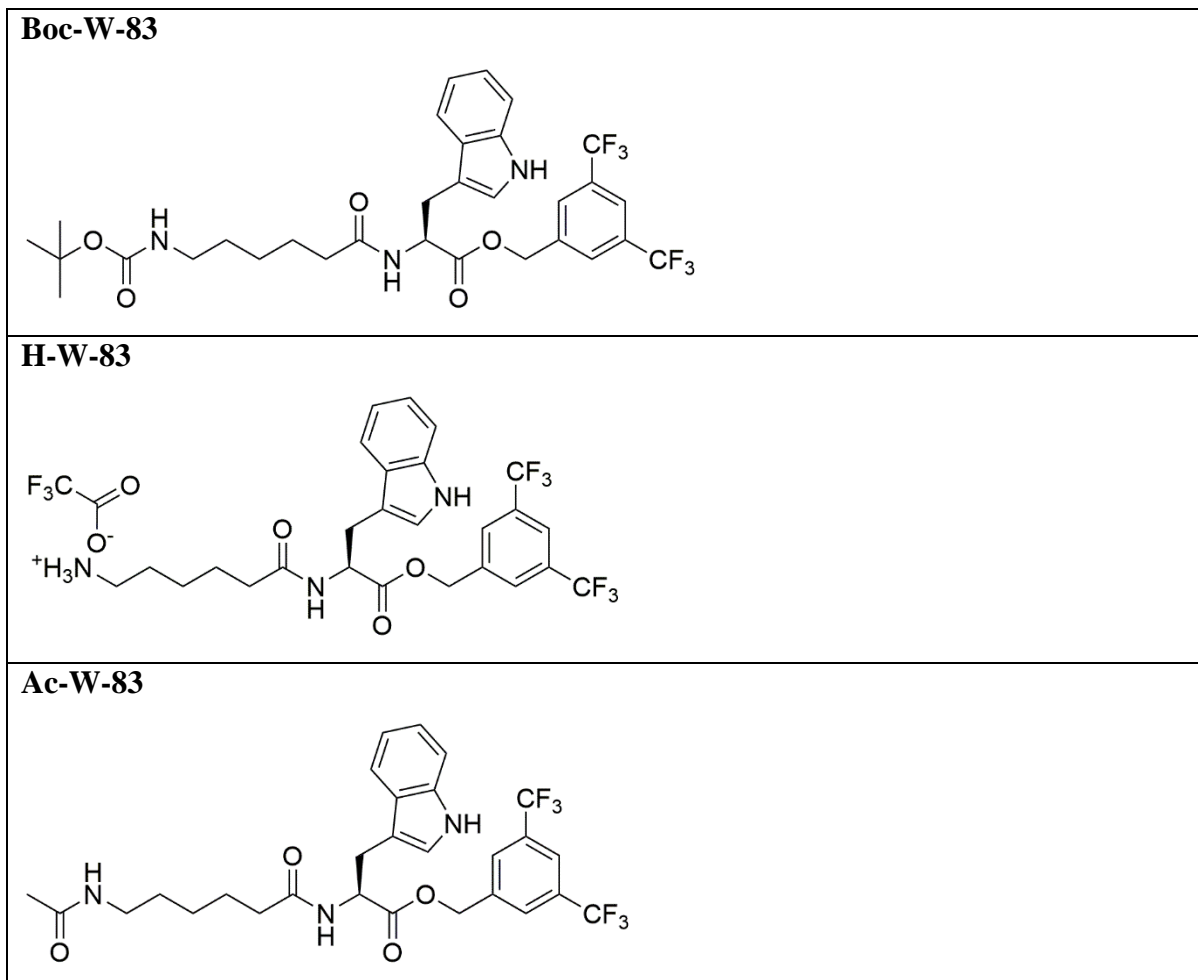
Ac-W-82, Ac-NH-D-Ala-D-Ala-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy *N*-acetylo-D-alanylo-D-alanylo-L-tryptofanu; zsyntezowano według procedury F; biały osad, wydajność reakcji 62%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [$\text{M}+\text{H}$]⁺: 615.20, *m/z* obliczone [$\text{M}+\text{H}$]⁺: 615.21; $t_{\text{R}} = 17.07$ min;

^1H NMR (600 MHz, DMSO-*d*6) δ (ppm), 10.83 (1H, s), 8.25 (1H, d, $J = 7.4$ Hz), 8.03 (1H, s), 7.98 (1H, d, $J = 7.3$ Hz), 7.96 (2H, s), 7.86 (1H, d, $J = 7.8$ Hz), 7.43 (1H, m), 7.28 (1H, m), 7.13 (1H, d, $J = 2.1$ Hz), 7.02 (1H, m), 6.92 (1H, m), 5.23 (1H, d, $J = 13.3$ Hz), 5.16 (1H,

d, $J = 13.3$ Hz), 4.50 (1H, m), 4.24 (1H, m), 4.19 (1H, m), 3.18 (1H, m), 3.06 (1H, m), 1.79 (3H, s), 1.09 (3H, d, $J = 7.2$ Hz), 1.05 (3H, d, $J = 7.2$ Hz).

VII.3.12. Seria W-83

Tabela 52. Wzory strukturalne związków z serii W-83.



Boc-W-83, Boc-Ahx-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (6-((*tert*-butyloksykarbonylo)amino)heksanoilo)-L-tryptofanu; zsyntezowano wg procedury C; biały osad, wydajność reakcji 57%, czystość > 97% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 644.15, m/z obliczone [M+H]⁺: 644.26; t_R = 20.57 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.83 (1H, s), 8.32 (1H, d, *J* = 7.4 Hz), 8.04 (1H, s), 7.95 (2H, s), 7.47 (1H, m), 7.32 (1H, m), 7.15 (1H, d, *J* = 2.2 Hz), 7.05 (1H, m), 6.96 (1H, m), 6.69 (1H, bs), 5.28 (1H, d, *J* = 13.5 Hz), 5.17 (1H, d, *J* = 13.5 Hz), 4.53 (1H, m), 3.18 (1H, m), 3.07 (1H, m), 2.82 (2H, m), 2.06 (2H, t, *J* = 7.4 Hz), 1.40 (2H, m), 1.36 (9H, s), 1.29 (2H, m), 1.13 (2H, m);

¹³C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 172.9, 172.4, 156.8, 140.0, 136.5, 130.7 (q, ²*J*_{CF} ~32.5 Hz), 128.7 (szeroki, prawdopodobnie nierozwiązany kwartet o ³*J*_{CF} ~ 4 Hz), 127.5, 124.1, 123.7 (q, ¹*J*_{CF} ~ 272 Hz), 122.1 (szeroki, prawdopodobnie nierozwiązany kwartet o

$^3J_{CF} \sim 4$ Hz), 121.4, 118.8, 118.3, 111.9, 109.9, 77.7, 64.7, 53.9, 35.3, 29.7, 28.7, 27.3, 26.3, 25.3. Sygnał C_ϵ Ahx pokrywa się z sygnałem DMSO- d_6 przy przesunięciu około 40 ppm.

H-W-83, TFA*NH₂-Ahx-L-Trp-O-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (S)-6-((1-((3,5-bis(trifluorometylo)benzylo)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-ylo)amino)-6-oksoheksano-1-aminowy; otrzymano wg procedury E; kremowy osad, wydajność reakcji 65%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane [$M_{zasada+H}$]⁺: 544.15, [$M_{zasada-H}$]⁻: 542.10, [$M_{sól-H}$]⁻: 656.05, m/z obliczone [$M_{zasada+H}$]⁺: 544.21, [$M_{zasada-H}$]⁻: 542.19, [$M_{sól-H}$]⁻: 656.19; t_R = 11.67 min;

¹H NMR (500 MHz, DMSO- d_6) δ (ppm), 10.84 (1H, s), 8.35 (1H, d, J = 7.1 Hz), 8.05 (1H, s), 7.95 (2H, s), 7.59 (3H, bs), 7.48 (1H, m), 7.32 (1H, m), 7.15 (1H, d, J = 2.0 Hz), 7.06 (1H, m), 6.97 (1H, m), 5.28 (1H, d, J = 13.5 Hz), 5.17 (1H, d, J = 13.5 Hz), 4.54 (1H, m), 3.19 (1H, m), 3.07 (1H, m), 2.70 (2H, m), 2.09 (2H, t, J = 7.6 Hz), 1.50-1.38 (4H, m), 1.20 (2H, m).

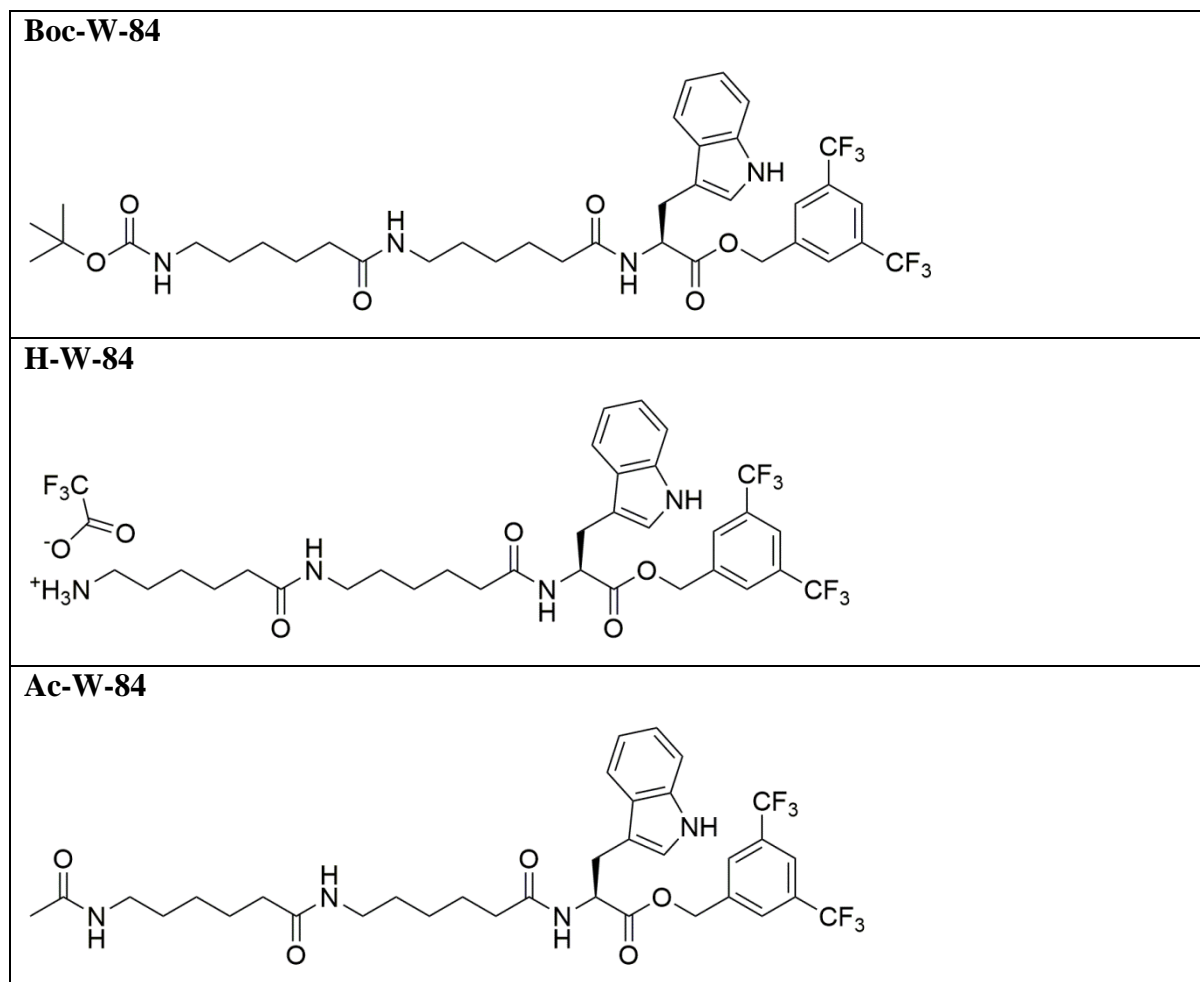
Ac-W-83, Ac-NH-Ahx-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (6-acetamidoheksanoilo)-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 63%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane [$M+H$]⁺: 586.15, m/z obliczone [$M+H$]⁺: 586.22; t_R = 17.58 min;

¹H NMR (600 MHz, DMSO- d_6) δ (ppm), 10.82 (1H, s), 8.31 (1H, d, J = 7.0 Hz), 8.02 (1H, s), 7.93 (2H, s), 7.71 (1H, bs), 7.44 (1H, m), 7.29 (1H, m), 7.12 (1H, d, J = 2.0 Hz), 7.05 (1H, m), 6.96 (1H, m), 5.25 (1H, d, J = 13.6 Hz), 5.14 (1H, d, J = 13.6 Hz), 4.50 (1H, m), 3.16 (1H, m), 3.04 (1H, m), 2.91 (2H, m), 2.04 (2H, t, J = 7.3 Hz), 1.74 (3H, s), 1.38 (2H, m), 1.27 (2H, m), 1.11 (2H, m);

¹³C NMR (125 MHz, DMSO- d_6) δ (ppm), 172.9, 172.4, 169.3, 140.0, 136.5, 130.7 (q, $^2J_{CF} \sim 32.5$ Hz), 128.7 (prawdopodobnie nierozwiązany kwartet o $^3J_{CF} \sim 4$ Hz), 127.5, 124.1, 123.7 (q, $^1J_{CF} \sim 272$ Hz), 122.1 (prawdopodobnie nierozwiązany kwartet o $^3J_{CF} \sim 4$ Hz), 121.4, 118.8, 118.3, 111.9, 109.9, 64.7, 53.8, 35.3, 29.4, 27.3, 26.5, 25.3, 23.0. Sygnał od C_ϵ Ahx pokrywa się z sygnałem DMSO- d_6 przy przesunięciu około 40 ppm.

VII.3.13. Seria W-84

Tabela 53. Wzory strukturalne związków z serii W-84.



Boc-W-84, Boc-Ahx-Ahx-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (6-(6-((*tert*-butyloksykarbonylo)amino)heksanoamido)heksanoilo)-L-tryptofanu; zsyntezowano wg procedury C, biały osad, wydajność reakcji 55%, czystość > 98% (HPLC-UV), ESI-MS *m/z* otrzymane [M+H]⁺: 757.30, *m/z* obliczone [M+H]⁺: 757.34; *t_R* = 19.61 min; ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.83 (1H, s), 8.32 (1H, d, *J* = 7.1 Hz), 8.04 (1H, s), 7.96 (2H, s), 7.66 (1H, t, *J* = 5.5 Hz), 7.48 (1H, m), 7.21 (1H, m), 7.15 (1H, d, *J* = 2.1 Hz), 7.05 (1H, m), 6.96 (1H, m), 6.71 (1H, t, *J* = 5.1 Hz), 5.28 (1H, d, *J* = 13.5 Hz), 5.17 (1H, d, *J* = 13.5 Hz), 4.53 (1H, m), 3.19 (1H, m), 3.08 (1H, m), 2.95 (2H, m), 2.87 (2H, m), 2.07 (2H, t, *J* = 7.6 Hz), 2.01 (2H, t, *J* = 7.5 Hz), 1.43 (4H, m), 1.36 (9H, s), 1.33 (4H, m), 1.16 (4H, m);

^{13}C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 172.9, 172.4, 172.3, 156.0, 140.0, 136.5, 130.7 (q, $^1J_{\text{CF}} \sim 32.5$ Hz), 128.7 (prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 127.5, 124.1, 123.7 (q, $^1J_{\text{CF}} \sim 272$ Hz), 122.1 (prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 121.4, 118.8, 118.3, 111.9, 109.9, 77.7, 64.7, 55.8, 38.7, 35.8, 35.3, 29.7, 29.4, 28.7, 27.3, 26.5, 25.5, 25.3. Sygnały od C_ε obu reszt Ahx pokrywają się z sygnałem DMSO-*d*₆ przy przesunięciu około 40 ppm.

H-W-84, TFA*NH₂-Ahx-Ahx-L-Trp-O-CH₂-3,5(CF₃)₂Ph, 2,2,2-trifluorooctan (*S*)-6-((6-((1-((3,5-bis(trifluorometylo)benzylo)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-yl)amino)-6-oksoheksylo)amino)-6-oksoheksano-1-aminiowy; otrzymano wg procedury E; kremowy osad, wydajność reakcji 57%, czystość > 97% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 657.20, [M_{zasada}-H]⁻: 655.10, [M_{sól}-H]⁻: 769.10, m/z obliczone [M_{zasada}+H]⁺: 657.29, [M_{zasada}-H]⁻: 655.27, [M_{sól}-H]⁻: 769.26; t_R = 11.73 min;

^1H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.84 (1H, s), 8.32 (1H, d, $J = 7.2$ Hz), 8.04 (1H, s), 7.96 (2H, s), 7.69 (1H, t, $J = 5.5$ Hz), 7.60 (3H, bs), 7.48 (1H, m), 7.32 (1H, m), 7.15 (1H, d, $J = 2.0$ Hz), 7.05 (1H, m), 6.96 (1H, m), 5.28 (1H, d, $J = 13.7$ Hz), 5.17 (1H, d, $J = 13.7$ Hz), 4.53 (1H, m), 3.19 (1H, m), 3.08 (1H, m), 2.95 (2H, m), 2.76 (2H, m), 2.10-2.00 (4H, m), 1.49 (4H, m), 1.42 (2H, m), 1.34-1.22 (4H, m), 1.15 (2H, s).

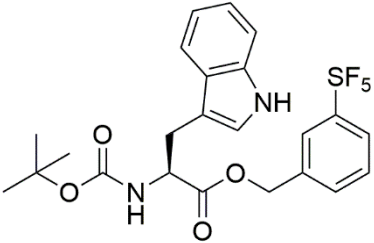
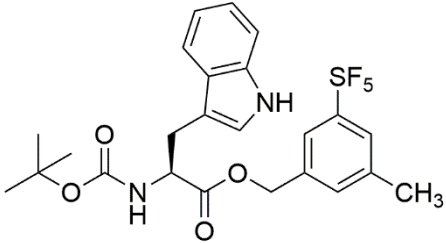
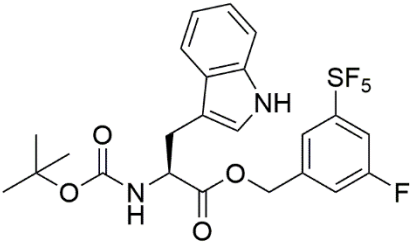
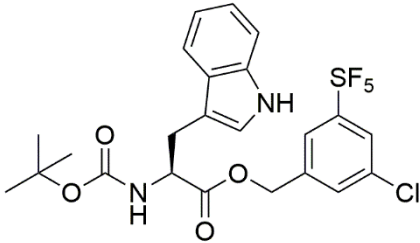
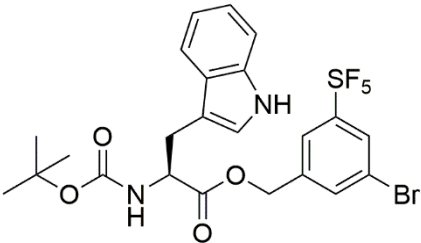
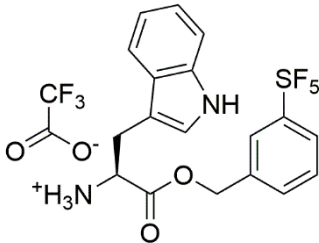
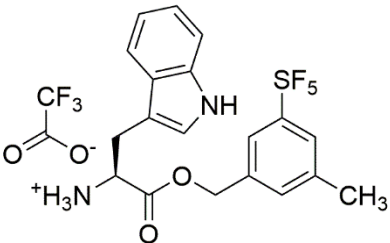
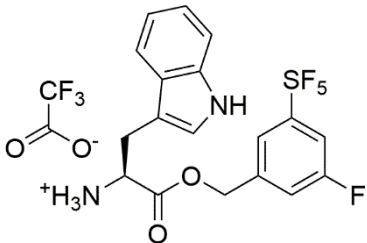
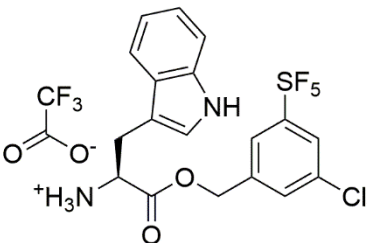
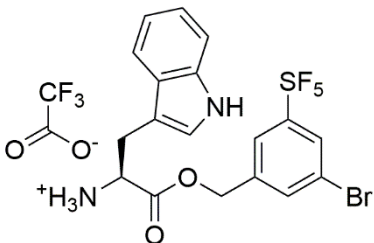
Ac-W-84, Ac-NH-Ahx-Ahx-L-Trp-O-CH₂-3,5(CF₃)₂Ph, ester 3,5-bis(trifluorometylo)benzylowy (6-(6-acetamidoheksanoamido)heksanoilo)-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność 68%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 699.25, m/z obliczone [M+H]⁺: 699.30; t_R = 17.01 min;

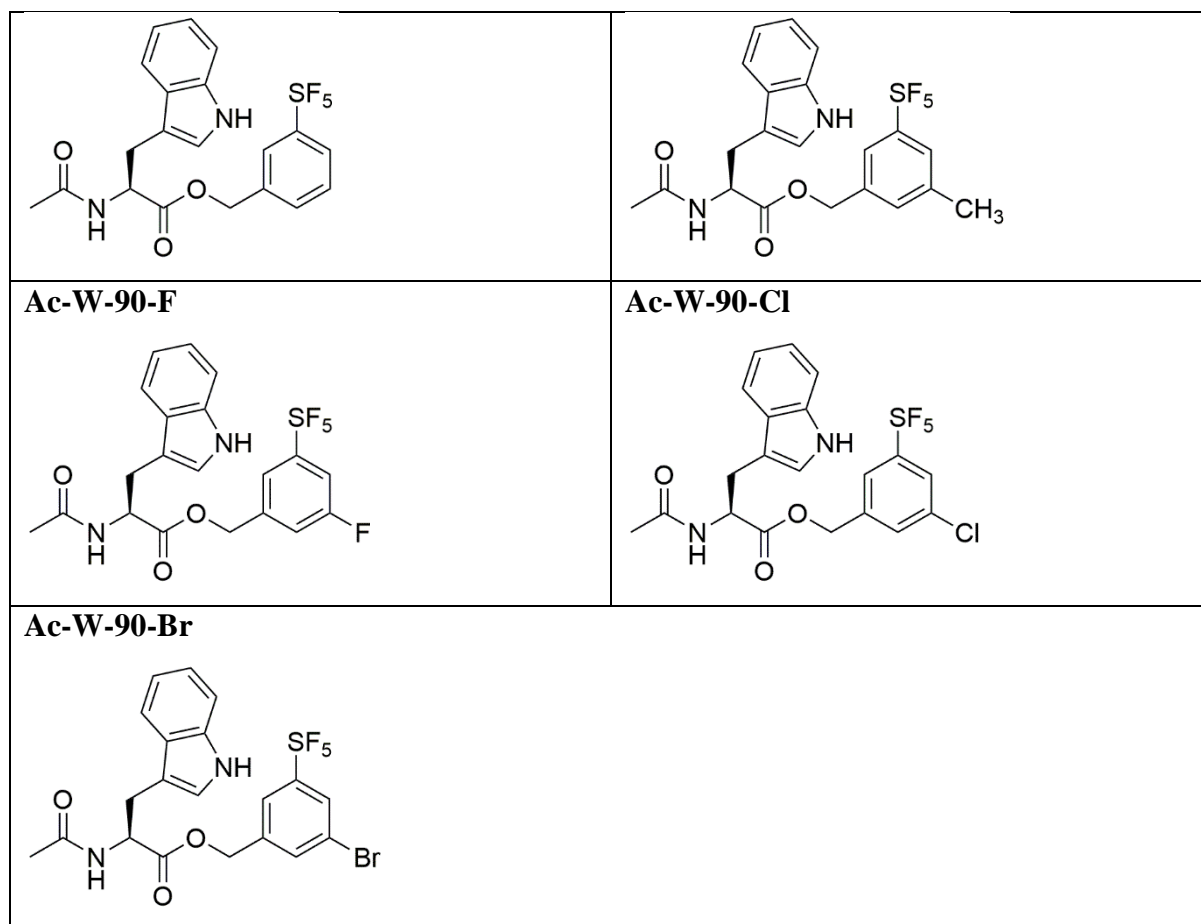
^1H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.84 (1H, s), 8.32 (1H, d, $J = 7.4$ Hz), 8.04 (1H, s), 7.96 (2H, s), 7.75 (1H, bs), 7.66 (1H, t, $J = 5.6$ Hz), 7.47 (1H, m), 7.32 (1H, m), 7.15 (1H, d, $J = 2.3$ Hz), 7.05 (1H, m), 6.96 (1H, m), 5.28 (1H, d, $J = 13.5$ Hz), 5.17 (1H, d, $J = 13.5$ Hz), 4.53 (1H, m), 3.19 (1H, m), 3.07 (1H, m), 3.02-2.92 (4H, m), 2.07 (2H, t, $J = 7.4$ Hz), 2.01 (2H, t, $J = 7.4$ Hz), 1.77 (3H, s), 1.49-1.27 (8H, m), 1.23-1.10 (4H, m);

^{13}C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 172.9, 172.4, 172.2, 169.3, 140.0, 136.5, 130.7 (q, $^1J_{\text{CF}} \sim 32.5$ Hz), 128.7 (prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 127.5, 124.1, 124.0 (q, $^1J_{\text{CF}} \sim 272$ Hz), 122.1 (prawdopodobnie nierozwiązany kwartet o $^3J_{\text{CF}} \sim 4$ Hz), 121.4, 118.8, 118.3, 111.9, 109.9, 64.7, 53.8, 38.9, 38.7, 35.8, 35.3, 29.4, 27.3, 26.6, 26.5, 25.5, 25.3, 23.0. Sygnał od C_ε jednej z reszt Ahx pokrywa się z sygnałem DMSO-*d*₆ przy przesunięciu około 40 ppm.

VII.3.14. Seria W-90

Tabela 54. Wzory strukturalne związków z serii W-90.

<p>Boc-W-90-H</p> 	<p>Boc-W-90-Me</p> 
<p>Boc-W-90-F</p> 	<p>Boc-W-90-Cl</p> 
<p>Boc-W-90-Br</p> 	<p>H-W-90-H</p> 
<p>H-W-90-Me</p> 	<p>H-W-90-F</p> 
<p>H-W-90-Cl</p> 	<p>H-W-90-Br</p> 
<p>Ac-W-90-H</p>	<p>Ac-W-90-Me</p>



Boc-W-90-H, Boc-L-Trp-O-CH₂-3-(SF₅)-Ph, ester 3-(pentafluoro-λ⁶-sulfanylo)benzylowy (*tert*-butyloksykarbonylo)-L-tryptofanu; zsyntezowano wg procedury G; biały osad, wydajność reakcji 82%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 519.00, m/z obliczone [M-H]⁻: 519.14; t_R = 20.59 min.

Boc-W-90-Me, Boc-L-Trp-O-CH₂-3-CH₃-5-(SF₅)-Ph, ester 3-metylo-5-(pentafluoro-λ⁶-sulfanylo)benzylowy (*tert*-butyloksykarbonylo)-L-tryptofanu; zsyntezowano wg procedury G; biały osad, wydajność reakcji 90%, czystość = 90% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 553.00, m/z obliczone [M-H]⁻: 553.15; t_R = 21.37 min.

Boc-W-90-F, Boc-L-Trp-O-CH₂-3-F-5-(SF₅)-Ph, ester 3-fluoro-5-(pentafluoro-λ⁶-sulfanylo)benzylowy (*tert*-butyloksykarbonylo)-L-tryptofanu; zsyntezowano wg procedury G; biały osad, wydajność reakcji 88%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 537.00, m/z obliczone [M-H]⁻: 537.11; t_R = 21.08 min.

Boc-W-90-Cl, Boc-L-Trp-O-CH₂-3-Cl-5-(SF₅)-Ph, ester 3-chloro-5-(pentafluoro-λ⁶-sulfanylo)benzylowy (*tert*-butyloksykarbonylo)-L-tryptofanu; zsyntezowano wg procedury G; biały osad, wydajność reakcji 88%, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 553.00, m/z obliczone [M-H]⁻: 553.10; t_R = 21.98 min.

Boc-W-90-Br, Boc-L-Trp-O-CH₂-3-Br-5-(SF₅)-Ph, ester 3-bromo-5-(pentafluoro-λ⁶-sulfanylo)benzylowy (*tert*-butyloksykarbonylo)-L-tryptofanu; zsyntezowano wg procedury G; biały osad, wydajność reakcji 93%, czystość = 83% (HPLC-UV), ESI-MS m/z otrzymane [M-H]⁻: 598.85, m/z obliczone [M-H]⁻: 598.40; t_R = 21.82 min.

H-W-90-H, TFA*NH₂-L-Trp-O-CH₂-3-(SF₅)-Ph, 2,2,2-trifluorooctan (*S*)-3-(1*H*-indol-3-ilo)-1-okso-1-((3-(pentafluoro-λ⁶-sulfanylo)benzyl)oksy)propan-2-aminowy; otrzymano wg procedury E; kremowy osad, wydajność reakcji 77%, czystość > 98% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 421.00, [M_{zasada}-H]⁻: 419.00, [M_{sól}-H]⁻: 533.00, m/z obliczone [M_{zasada}+H]⁺: 421.10, [M_{zasada}-H]⁻: 419.08, [M_{sól}-H]⁻: 533.07; t_R = 11.25 min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.09 (1H, s), 8.52 (3H, s), 7.88 (1H, s, nałożony), 7.88 (1H, prawdopodobnie d, nałożony), 7.57 (1H, d, ³J_{HH} = 8.3, 8.3 Hz), 7.49 (1H, d, ³J_{HH} = 8.0 Hz), 7.42 (1H, d, ³J_{HH} = 7.5 Hz), 7.39 (1H, d, ³J_{HH} = 8.0 Hz), 7.21 (1H, d, ³J_{HH} = 2.0 Hz), 7.10 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 7.00 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 5.23 (2H, **AB**, ³J_{AB} = 12.8 Hz, Δν_{AB} = 47.9 Hz), 4.40-4.34 (1H, m), 3.34-3.24 (2H, m).

H-W-90-Me, TFA*NH₂-L-Trp-O-CH₂-3-CH₃-5-(SF₅)-Ph, 2,2,2-trifluorooctan (*S*)-3-(1*H*-indol-3-ilo)-1-((3-metylo-5-(pentafluoro-λ⁶-sulfanylo)benzyl)oksy)-1-oksopropan-2-aminowy; otrzymano wg procedury E; biały osad, wydajność reakcji 72%, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 435.05, [M_{zasada}-H]⁻: 433.00, [M_{sól}-H]⁻: 546.95, m/z obliczone [M_{zasada}+H]⁺: 435.12, [M_{zasada}-H]⁻: 433.10, [M_{sól}-H]⁻: 547.09; t_R = 11.40 min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.06 (1H, s), 8.44 (3H, s), 7.70 (1H, s), 7.66 (1H, s), 7.48 (1H, d, ³J_{HH} = 8.0 Hz), 7.39 (1H, d, ³J_{HH} = 8.0 Hz), 7.20 (1H, s, nałożony), 7.20 (1H, s, nałożony), 7.10 (1H, dd, ³J_{HH} = 7.8, 7.8 Hz), 7.00 (1H, dd, ³J_{HH} = 7.3, 7.3 Hz), 5.16 (2H, **AB**, ³J_{AB} = 12.6 Hz, Δν_{AB} = 63.0 Hz), 4.38-4.34 (1H, m), 3.30-3.26 (2H, m), 2.37 (3H, s).

H-W-90-F, TFA*NH₂-L-Trp-O-CH₂-3-F-5-(SF₅)-Ph, 2,2,2-trifluorooctan (*S*)-1-((3-fluoro-5-(pentafluoro-λ⁶-sulfanylo)benzyl)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminowy; otrzymano wg procedury E; lekko pomarańczowy osad, wydajność reakcji 79%, czystość = 96% (HPLC-UV), ESI-MS m/z otrzymane [M_{zasada}+H]⁺: 439.00, [M_{zasada}-H]⁻: 437.00, [M_{sól}-H]⁻: 550.95, m/z obliczone [M_{zasada}+H]⁺: 439.09, [M_{zasada}-H]⁻: 437.07, [M_{sól}-H]⁻: 551.40; t_R = 11.54 min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.06 (1H, s), 8.49 (3H, s), 7.88 (1H, d, ³J_{HF} = 9.0 Hz), 7.77 (1H, s), 7.48 (1H, d, ³J_{HH} = 7.5 Hz), 7.37 (1H, d, ³J_{HH} = 7.5 Hz), 7.34 (1H, d, ³J_{HF} = 9.0 Hz), 7.22 (1H, d, ³J_{HH} = 2.0 Hz), 7.09 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 6.99 (1H, dd, ³J_{HH} =

7.5, 7.5 Hz), 5.24 (2H, **AB**, $^3J_{AB} = 13.3$ Hz, $\Delta\nu_{AB} = 54.1$ Hz), 4.45-4.34 (1H, m), 3.35-3.24 (2H, m).

H-W-90-Cl, TFA*NH₂-L-Trp-O-CH₂-3-Cl-5-(SF₅)-Ph, 2,2,2-trifluorooctan (*S*)-1-((3-chloro-5-(pentafluoro- λ^6 -sulfanylo)benzylo)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminowy; otrzymano wg procedury E; biały osad, wydajność reakcji 78%, czystość = 97% (HPLC-UV), ESI-MS *m/z* otrzymane [$M_{zasada+H}^+$]: 455.00, [$M_{zasada-H}^-$]: 452.95, [$M_{sól-H}^-$]: 566.90, *m/z* obliczone [$M_{zasada+H}^+$]: 455.06, [$M_{zasada-H}^-$]: 453.04, [$M_{sól-H}^-$]: 567.03; $t_R = 11.69$ min; ¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.06 (1H, s), 8.48 (3H, s), 8.02 (1H, s), 7.89 (1H, s), 7.58 (1H, s), 7.47 (1H, d, $^3J_{HH} = 8.0$ Hz), 7.37 (1H, d, $^3J_{HH} = 8.0$ Hz), 7.22 (1H, d, $^3J_{HH} = 1.5$ Hz), 7.09 (1H, dd, $^3J_{HH} = 7.8, 7.5$ Hz), 6.99 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 5.23 (2H, **AB**, $^3J_{AB} = 13.1$ Hz, $\Delta\nu_{AB} = 59.1$ Hz), 4.43-4.35 (1H, m), 3.32-3.26 (2H, m).

H-W-90-Br, TFA*NH₂-L-Trp-O-CH₂-3-Br-5-(SF₅)-Ph, 2,2,2-trifluorooctan (*S*)-1-((3-bromo-5-(pentafluoro- λ^6 -sulfanylo)benzylo)oksy)-3-(1*H*-indol-3-ilo)-1-oksopropan-2-aminowy; otrzymano wg procedury E; biały osad, wydajność reakcji 58%, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane [$M_{zasada+H}^+$]: 500.90, [$M_{zasada-H}^-$]: 498.85 *m/z* obliczone [$M_{zasada+H}^+$]: 500.22, [$M_{zasada-H}^-$]: 498.20; $t_R = 11.60$ min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 11.05 (1H, s), 8.42 (3H, s), 8.11 (1H, s), 7.93 (1H, s), 7.74 (1H, s), 7.47 (1H, d, $^3J_{HH} = 8.0$ Hz), 7.38 (1H, d, $^3J_{HH} = 8.0$ Hz), 7.21 (1H, d, $^3J_{HH} = 2.0$ Hz), 7.09 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 6.99 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 5.23 (2H, **AB**, $^3J_{AB} = 13.0$ Hz, $\Delta\nu_{AB} = 61.2$ Hz), 4.41-4.36 (1H, m), 3.31-3.25 (2H, m).

Ac-W-90-H, Ac-NH-L-Trp-O-CH₂-3-(SF₅)-Ph, ester 3-(pentafluoro- λ^6 -sulfanylo)benzylowy *N*-acetylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 67%, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane [$M+H^+$]: 463.05, [$M-H^-$]: 461.00, *m/z* obliczone [$M+H^+$]: 463.11, [$M-H^-$]: 461.09; $t_R = 17.60$ min;

¹H NMR (500 MHz, DMSO-*d*6) δ (ppm), 10.86 (1H, s), 8.39 (1H, d, $^3J_{HH} = 7.3$ Hz), 7.84 (d, $^3J_{HH} = 7.8$ Hz), 7.83 (1H, s), 7.56 (1H, dd, $^3J_{HH} = 7.8, 7.8$ Hz), 7.49 (1H, d, $^3J_{HH} = 7.8$ Hz), 7.49 (1H, d, $^3J_{HH} = 7.5$ Hz), 7.35 (1H, d, $^3J_{HH} = 8.5$ Hz), 7.15 (1H, d, $^3J_{HH} = 1.5$ Hz), 7.07 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 6.98 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 5.15 (2H, **AB**, $^3J_{AB} = 13.2$ Hz, $\Delta\nu_{AB} = 30.3$ Hz), 4.55 (1H, td, $^3J_{HH} = 7.9, 7.3, 6.7$ Hz), 3.12 (2H, **ABX**, $J_{AB} = 14.5$ Hz, $J_{AX} = 6.7$ Hz, $J_{BX} = 7.9$ Hz, $\nu_{AB} = 53.3$ Hz), 1.83 (3H, s);

¹³C NMR (125 MHz, DMSO-*d*6) δ (ppm), 171.9, 169.4, 152.8 (p, $^2J_{CF} = 15.0$ Hz), 137.8, 136.1, 131.3, 129.5, 127.0, 125.2 (p, $^3J_{CF} = 4.7$ Hz), 124.5 (p, $^3J_{CF} = 4.2$ Hz), 123.6, 121.0, 118.4, 117.9, 111.4, 109.4, 64.8, 53.3, 27.0, 22.2.

Ac-W-90-Me, Ac-NH-L-Trp-O-CH₂-3-CH₃-5-(SF₅)-Ph, ester 3-metylo-5-(pentafluoro-λ⁶-sulfanylo)benzylowy *N*-acetylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 57%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 477.05, [M-H]⁻: 475.00, m/z obliczone [M+H]⁺: 477.13, [M-H]⁻: 475.11; t_R = 18.32 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.86 (1H, s), 8.39 (1H, d, ³J_{HH} = 7.3 Hz), 7.67 (1H, s), 7.61 (1H, s), 7.48 (1H, d, ³J_{HH} = 8.0 Hz), 7.35 (1H, d, ³J_{HH} = 8.0 Hz), 7.23 (1H, s), 7.15 (1H, d, ³J_{HH} = 1.5 Hz), 7.07 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 6.98 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 5.09 (2H, **AB**, ³J_{AB} = 13.0 Hz, Δv_{AB} = 42.2 Hz), 4.54 (1H, td, ³J_{HH} = 8.3, 7.3, 6.2 Hz), 3.12 (2H, **ABX**, J_{AB} = 14.4 Hz, J_{AX} = 6.2 Hz, J_{BX} = 8.3 Hz, v_{AB} = 50.2 Hz), 2.37 (3H, s), 1.83 (3H, s);

¹³C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 175.3, 172.8, 156.2 (p, ²J_{CF} = 15.5 Hz), 143.0, 139.4, 135.2, 130.3, 128.7 (p, ³J_{CF} = 4.2 Hz), 127.0, 125.1 (p, ³J_{CF} = 4.3 Hz), 124.3, 121.7, 121.2, 114.8, 112.7, 68.2, 56.7, 30.3, 25.5, 24.1.

Ac-W-90-F, Ac-NH-L-Trp-O-CH₂-3-F-5-(SF₅)-Ph, ester 3-fluoro-5-(pentafluoro-λ⁶-sulfanylo)benzylowy *N*-acetylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 69%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 481.05, [M-H]⁻: 478.95, m/z obliczone [M+H]⁺: 481.10, [M-H]⁻: 479.08; t_R = 18.02 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.86 (1H, s), 8.41 (1H, d, ³J_{HH} = 7.2 Hz), 7.84 (1H, d, ³J_{HF} = 9.0 Hz), 7.72 (1H, s), 7.48 (1H, d, ³J_{HH} = 8.0 Hz), 7.34 (1H, d, ³J_{HH} = 8.0 Hz), 7.33 (1H, d, ³J_{HF} = 9.0 Hz), 7.16 (1H, d, ³J_{HH} = 2.0 Hz), 7.07 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 6.98 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 5.15 (2H, **AB**, ³J_{AB} = 13.6 Hz, Δv_{AB} = 47.7 Hz), 4.55 (1H, td, ³J_{HH} = 7.8, 7.2, 6.7 Hz), 3.13 (2H, **ABX**, J_{AB} = 14.5 Hz, J_{AX} = 6.7 Hz, J_{BX} = 7.8 Hz, v_{AB} = 48.7 Hz), 1.83 (3H, s);

¹³C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 171.8, 169.5, 161.1 (d, ¹J_{CF} = 247.0 Hz), 153.4-153.0 (słaby, prawdopodobnie nierozwiązany multiplet), 140.3 (p, ³J_{CF} = 3.7 Hz), 136.1, 127.0, 123.7, 121.0, 120.6 (p, ³J_{CF} = 4.0 Hz), 118.4, 118.3 (d, ²J_{CF} = 22.0 Hz), 117.8, 113.2 (d, ²J_{CF} = 26.8 Hz), 111.4, 109.3, 64.2, 53.4, 26.9, 22.1.

Ac-W-90-Cl, Ac-NH-L-Trp-O-CH₂-3-Cl-5-(SF₅)-Ph, ester 3-chloro-5-(pentafluoro-λ⁶-sulfanylo)benzylowy *N*-acetylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 44%, czystość = 98% (HPLC-UV), ESI-MS m/z otrzymane [M+H]⁺: 496.95, [M-H]⁻: 494.95, m/z obliczone [M+H]⁺: 497.07, [M-H]⁻: 495.05; t_R = 18.71 min;

¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.86 (1H, s), 8.41 (1H, d, ³J_{HH} = 7.1 Hz), 7.98 (1H, s), 7.83 (1H, s), 7.57 (1H, s), 7.48 (1H, d, ³J_{HH} = 7.5 Hz), 7.34 (1H, d, ³J_{HH} = 7.5 Hz), 7.17 (1H, d, ³J_{HH} = 1.5 Hz), 7.06 (1H, dd, ³J_{HH} = 7.5, 7.5 Hz), 6.98 (1H, dd, ³J_{HH} = 7.5, 7.5

Hz), 5.14 (2H, **AB**, $^3J_{AB} = 13.5$ Hz, $\Delta\nu_{AB} = 47.6$ Hz), 4.54 (1H, td, $^3J_{HH} = 7.8, 7.1, 6.7$ Hz), 3.13 (2H, **ABX**, $J_{AB} = 14.6$ Hz, $J_{AX} = 6.7$ Hz, $J_{BX} = 7.8$ Hz, $\nu_{AB} = 50.4$ Hz), 1.83 (3H, s); ^{13}C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 171.8, 169.5, 153.2 (p, $^2J_{CF} = 17.5$ Hz), 139.9, 136.0, 133.8, 131.3, 127.0, 125.1 (p, $^3J_{CF} = 4.3$ Hz), 123.6, 123.3 (p, $^3J_{CF} = 4.7$ Hz), 120.9, 118.4, 117.8, 111.5, 109.3, 64.1, 53.4, 26.9, 22.1.

Ac-W-90-Br, Ac-NH-L-Trp-O-CH₂-3-Br-5-(SF₅)-Ph, ester 3-bromo-5-(pentafluoro- λ^6 -sulfanylo)benzylowy *N*-acetylo-L-tryptofanu; zsyntezowano wg procedury F; biały osad, wydajność reakcji 51%, czystość = 98% (HPLC-UV), ESI-MS *m/z* otrzymane $[\text{M}+\text{H}]^+$: 542.90, $[\text{M}-\text{H}]^-$: 540.85, *m/z* obliczone $[\text{M}+\text{H}]^+$: 542.34, $[\text{M}-\text{H}]^-$: 540.32; $t_R = 18.83$ min;

^1H NMR (500 MHz, DMSO-*d*₆) δ (ppm), 10.86 (1H, s), 8.41 (1H, d, $^3J_{HH} = 7.4$ Hz), 8.08 (1H, s), 7.87 (1H, s), 7.73 (1H, s), 7.48 (1H, d, $^3J_{HH} = 8.0$ Hz), 7.34 (1H, d, $^3J_{HH} = 8.0$ Hz), 7.16 (1H, d, $^3J_{HH} = 1.5$ Hz), 7.07 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 6.98 (1H, dd, $^3J_{HH} = 7.5, 7.5$ Hz), 5.15 (2H, **AB**, $^3J_{AB} = 13.5$ Hz, $\Delta\nu_{AB} = 47.7$ Hz), 4.53 (1H, td, $^3J_{HH} = 8.3, 7.4, 6.1$ Hz), 3.13 (2H, **ABX**, $J_{AB} = 14.5$ Hz, $J_{AX} = 6.2$ Hz, $J_{BX} = 8.4$ Hz, $\nu_{AB} = 52.0$ Hz), 1.83 (3H, s); ^{13}C NMR (125 MHz, DMSO-*d*₆) δ (ppm), 171.8, 169.5, 153.3 (p, $^2J_{CF} = 15.3$ Hz), 140.1, 136.1, 134.2, 127.0, 125.1 (p, $^3J_{CF} = 4.3$ Hz), 123.7 (p, $^3J_{CF} = 4.4$ Hz), 123.6, 121.8, 120.9, 118.4, 117.8, 111.5, 109.3, 64.1, 53.4, 26.8, 22.1.

VII.4. Dokowanie do receptora NK1

Dokowanie molekularne do receptora NK1 przeprowadzono za pomocą programu AutoDock Vina [206]. Do dokowania użyto wszystkie zsyntezowane w trakcie projektu doktoranckiego związki (serie W/F-70–W/F-74, W-80–W-84 oraz W-90), związki otrzymane w ZN IMDiK PAN [318] oraz pobrane z bazy ChEMBL (opis Rodział VII.5.).

Struktury zsyntezowanych przeze mnie ligandów oraz struktury białek przygotowano w programie Biovia Discovery Studio [319]. Geometrie ligandów zminimalizowano za pomocą funkcji “clean geometry”. Ligandy przygotowano tak, aby wiązania hydrazonowe oraz estrowe miały płaską geometrię, a wiązanie estrowe było w konformacji *trans*. W przypadku ligandów z wolną grupą aminową, ustawiono stan protonowania +1. Związki posiadające na N-końcu grupę zabezpieczającą Cbz, Boc lub grupę acetylową dokowano zarówno w konformacjach *cis* i *trans* wokół wiązania amidowego. Dodatkowo, w programie AutoDock Tools [320] dla wszystkich ligandów ręcznie zahamowano rotację wokół wiązań estrowych, wiązania N-N w układzie hydrazonu oraz C-O w układzie karbaminianu. W

przypadku związków pobranych z bazy ChEMBL opis przygotowania ligandów do dokowania zamieszczono w Rozdziale VII.6. Do dokowania użyto struktury hNK1R pobrane z bazy PDB (6E59 [23], 6HLL [27], 6HLO [27], 7P00 [25], 7RMG [24]). Ze struktur receptorów usunięto krystalograficzne ligandy, dodano wodory i minimalizowano. Stany protonowania białek ustawiono zgodnie z oczekiwaniami przy pH = 7.4. Dla wszystkich struktur białek pułdo (box) dokowania dopasowano do pozycji wiązania krystalograficznego liganda i rozszerzono, uzyskując wymiary: 38.86 Å x 23.60 Å x 34.53 Å.

VII.5. Zestawy danych do oceny skuteczności i trenowania funkcji oceniających

Do oceny skuteczności i trenowania funkcji oceniających zostały użyte trzy zbiory danych.

Zbiór I to zestaw związków, które zsyntezowałam w trakcie projektu doktoranckiego (89 związków) i związków otrzymanych w ZN IMDiK PAN [318] (12 związków), których powinowactwo receptorowe do NK1R zostało oznaczone w ZN IMDiK PAN. Większość z tych związków jest nieaktywna lub słabo aktywna wobec NK1R. Zbiór I zawiera łącznie 101 związków, które są strukturalnie podobne do typowych antagonistów NK1R.

Zbiór II i III to zestawy związków z bazy ChEMBL [257], dla których dane o powinowactwie do receptora NK1 pochodzą z literatury naukowej. Z bazy ChEMBL pobrano związki, które zostały przebadane pod kątem powinowactwa do ludzkiego receptora NK1 (identyfikator: ChEMBL249) oraz do receptora NK1 kawii domowej (identyfikator: ChEMBL3942), ponieważ struktura obu receptorów jest bardzo podobna i brak jest większych międzygatunkowych różnic w wiązalności. Z pobranego zestawu usunięto pozycje, które nie zawierały danych o wiązalności receptorowej (IC_{50} lub K_i), o masie cząsteczkowej mniejszej niż 99 Da i większej niż 1000 Da oraz związki organometaliczne. Dane o aktywności konwertowano odpowiednio do pIC_{50} lub pK_i . Jeśli związki zawierały więcej niż jeden wpis o aktywności, dane te uśredniano. W przypadku gdy dostępne były dane tylko z przesiewu single-point, przypisywano pIC_{50} lub pK_i o 1 lub 2 jednostki niższe niż stężenie przesiewu (w zależności, czy podana była jakakolwiek inhibicja wiązania radioliganda).

Z tak opracowanego zestawu wyodrębniono Zbiór II i III. **Zbiór II** zawiera 2964 związki, które strukturalnie zbliżone są do znanego farmakoforu NK1. W większości są to ligandy aktywne wobec NK1R w zakresie średnionanomolowym i niskonanomolowym.

Zbiór III to związki, których powinowactwo do receptora NK1 było badane w dużym badaniu przesiewowym DRUGMATRIX [294,295]. Jest to zestaw bardzo zróżnicowany strukturalnie, zawiera m.in. zatwierdzone leki i znane substancje chemiczne. Składa się z 816 związków, nieaktywnych wobec NK1R.

VII.6. Przygotowanie ligandów z bazy ChEMBL do dokowania

Struktury związków pobranych z bazy ChEMBL przekonwertowano z formatu SMILES do formatu PDB za pomocą skryptów opartych o bibliotekę RDKit [321]. Zastosowano funkcję *EmbedMultipleConfs*, aby uzyskać wiele niskoenergetycznych konformacji pierścieni nasyconych, ponieważ algorytm przeszukujący AutoDock Vina nie przeszukuje przestrzeni konformacyjnej pierścieni. Za pomocą biblioteki Dimorphite-DL [322] wszystkim ligandom przypisano stany protonacyjne odpowiednie dla pH=7.4. Geometrie związków zoptymalizowano za pomocą mechaniki molekularnej polem siłowym MMFF94s [323]. Dodatkowo, w przypadku związków, dla których istniała taka potrzeba, wygenerowano oddzielnie konformery *cis* i *trans* wokół wiązania amidowego. Tak przygotowane struktury ligandów były używane do dokowania molekularnego.

VII.7. Parametry oceny skuteczności przesiewu wirtualnego

Oryginalna funkcja AutoDock Vina oraz nowoopracowane funkcje oceniające zostały przetestowane co do ich wydajności w przesiewie wirtualnym (odróżnianie „prawdziwych” silnych ligandów od związków o słabym wiązaniu do receptora NK1). Użyto powszechnie stosowane parametry oceny skuteczności metod przesiewu wirtualnego czyli logAUC oraz współczynnik wzbogacenia (EF, ang. *enrichment factor*).

logAUC – pole pod krzywą charakterystyki operacyjnej odbiornika ROC (ang. *receiver operating characteristic curve*) wykreślonej w skali semilogarytmicznej [297]. Krzywa ROC przedstawia zależność między czułością (ang. *True Positive Rate*) a specyficznością (ang.

True Negative Rate) modelu w rozpoznawaniu aktywnych ligandów spośród związków nieaktywnych.

Logarytmizacja ROC jest wykonywana w celu uwypuklenia skuteczności wczesnego rozpoznania (dla związków sklasyfikowanych w górnym procencie bazy). LogAUC jest obliczane poprzez całkowanie numeryczne metodą trapezów obszaru pod logarytmiczną krzywą ROC, zgodnie z Równaniem 8.

$$\log AUC_{\lambda} = \frac{\sum_i^{x \geq \lambda} \left(\frac{y_{i+1} - y_i}{\log_e 10} \right) + b_i (\log_{10} x_{i+1} - \log_{10} x_i)}{\log_{10} \left(\frac{1.0}{\lambda} \right)} \quad (8)$$

Gdzie b – punkt przecięcia z osią y dla danego odcinka trapezu, obliczany:

$$b_i = y_{i+1} - x_{i+1} \left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right) \quad (9)$$

Obszar logarytmiczny obliczany jest od ustalonego progu λ do 1.0 (typowo $\lambda=0.1$). Im wyższa wartość logAUC (bliska 1.0), tym model lepiej rozpoznaje aktywne ligandy spośród związków nieaktywnych.

W niniejszej pracy domyślnie podawane jest *dopasowane* logAUC czyli logAUC pomniejszone o 13.68%. Wartość ta jest wartością logAUC dla klasyfikacji opartej na losowaniu („rzut monetą”). Dzięki temu za pomocą *dopasowanego* logAUC ocenia się o ile „lepszy” jest dany model od wyboru na podstawie losowania.

Przy omawianiu wyników przesiewów zastosowano także opisową kategoryzację wyników według wartości logAUC zaproponowaną według Jaiteha i wsp. [298]. Według kryterium Jaiteha, logAUC może być słabe < 10 , niezłe (ang. *fair*) 10–15, dobre 15–20, bardzo dobre 20–25, lub doskonałe > 25 .

$EF_{x\%}$ - współczynnik wzbogacenia dla wybranego górnego $x\%$ bazy, wskaźnik oceny skuteczności metod przesiewu wirtualnego w identyfikowaniu aktywnych związków z dużych zbiorów danych. Matematycznie definiuje się go jako stosunek liczby aktywnych związków znalezionych w wybranym fragmencie (w górnym x procencie) posortowanego zbioru do liczby wszystkich związków w wybranym fragmencie rankingu, podzielony przez stosunek liczby aktywnych związków w całym zbiorze do liczby wszystkich związków w całym zbiorze (Równanie 10).

Współczynnik ten informuje o tym, czy metoda przesiewu preferencyjnie wybiera aktywne związki (gdy $EF_{x\%} > 1$), czy też związki nieaktywne (gdy $EF_{x\%} < 1$). Wartość $EF_{x\%}$ równa 1 sugeruje brak dodatkowego wzbogacenia w porównaniu do losowego wyboru. Im

wartość $EF_{x\%}$ jest bliższa maksymalnej możliwej wartości dla danego problemu (EF_{\max}), tym skuteczniejsza jest metoda przesiewu.

$$EF_{x\%} = \frac{\frac{\text{liczba aktywnych związków w górnym } x\% \text{ posortowanego zbioru}}{\text{liczba wszystkich związków w górnym } x\% \text{ posortowanego zbioru}}}{\frac{\text{liczba aktywnych związków w całym zbiorze}}{\text{liczba wszystkich związków w całym zbiorze}}} \quad (10)$$

$EF_{\text{frac}x\%}$ – znormalizowany współczynnik wzbogacenia, to jest $EF_{x\%}$ podzielony przez maksymalny możliwy EF (EF_{\max}) dla danego problemu (Równanie 11). Im wartość EF_{frac} jest bliższa 1, tym metoda przesiewu jest bardziej skuteczna (EF_{frac} równy 1 oznacza pełne możliwe wzbogacenie w porównaniu z maksymalnym teoretycznym wzbogaceniem).

$$EF_{\text{frac}x\%} = \frac{EF_{x\%}}{EF_{\max}} \quad (11)$$

VII.8. Trenowanie funkcji oceniających

Trenowanie nowych funkcji przeprowadzono za pomocą zestawu skryptów *VinaTuner* opracowanych w ZN IMDiK PAN przez dr. hab. Piotra Lipińskiego. Trenowanie i walidacja zostały przeprowadzone w kilku scenariuszach opisanych w dziale IV.5. oraz podsumowanych w Tabelach 16–27 i 29–40. Scenariusze różniły się zbiorami treningowymi i testowymi, oraz tym, jak traktowano wyniki

Dane eksperymentalne o powinowactwie (pAff) uzyskane z literatury lub z pomiarów w ZN IMDiK PAN zostały przekonwertowane na energię swobodną wiązania, dzięki czemu możliwe było porównywanie wartości otrzymanych ze *scoringu* wyrażonych jako energia z energią eksperymentalną.

Do trenowania wykorzystano algorytm genetyczny. Celem dostrajania było uzyskanie jak najniższej wartości średniego błędu bezwzględnego (MAE, ang. *Mean Absolute Error*) między predykcją entalpi swobodnej wiązania nowej funkcji a danymi eksperymentalnymi. Algorytm genetyczny pracował na populacjach liczących po 1000 osobników (losowych zestawów współczynników w_1 – w_6), w której najlepszych 50 osobników mogło się krzyżować (wymieniać parametrami w_1 – w_6). Dopuszczalne były losowe mutacje (losowa zmiana któregoś parametru w_1 – w_6 bez związku z krzyżowaniem). Trenowanie nowych funkcji trwały 25 pokoleń. Populacje startowe powstały przez losowanie małych perturbacji wag domyślnej funkcji AutoDock Vina. Dozwolony zakres zmienności dla wag funkcji populacji startowej

obejmował 6 różnych wariantów i mieścił się w zakresie $\pm 10, 20, 30, 40, 50$ lub 60% wartości funkcji domyślnej. Przyjęto, że oryginalna funkcja oceniająca AutoDock Vina zasadniczo odzwierciedla założenia fizyki wiązania i wymaga dostrojenia w małym zakresie, dlatego wykluczono drastyczne zmiany. Całą procedurę powtórzono 20 razy, uzyskując 120 przebiegów (20 przebiegów * 6 wariantów dopuszczalnej zmienności).

Ze zbioru otrzymanych nowych funkcji usunięto pozycje zdublowane oraz funkcje, dla których wartości $w_1 > 0, w_2 > 0, w_3 < 0, w_4 > 0, w_5 > 0, w_6 < 0$ (aby nie uzyskać funkcji sprzecznych z założeniami fizyki wiązania, np. promujących giętkość ligandów, albo zawady steryczne w pozach wiążących).

Dla każdego scenariusza trenowania, do walidacji wybrano po 5 najlepszych funkcji z każdego przebiegu (uszeregowanych według rosnących wartości MAE). Oznacza to, że dla każdego scenariusza, około 500–600 funkcji było przedmiotem walidacji na zbiorze testowym. W początkowych etapach badań rozważano wybór 100 górnych funkcji z całej populacji wszystkich stworzonych funkcji, ale otrzymano bardzo słabe wyniki w walidacji.

W ramach walidacji, wybrane funkcje zostały użyte na zbiorze walidacyjnym do *rescoringu* póź z dokowania. Ranking został posortowany według wartości energii (scoringu według nowej funkcji), a następnie obliczono logAUC (tak jak w przesiewie wirtualnym), które przyjęto jako miarę sukcesu. Nowe funkcje porównywano do oryginalnej funkcji AutoDock Vina za pomocą analizy ANOVA z testem wielokrotnych porównań Dunnetta.

VII.9. Predykcja celów molekularnych za pomocą metody SwissTarget

W celu przewidywania potencjalnych celów molekularnych, struktury związków z serii W-71 (w formacie SMILES) zadano na ogólnodostępny serwer SwissTargetPrediction [289,290]. Narzędzie to przewiduje potencjalne cele molekularne badanych związków na podstawie obliczenia ich podobieństwa do znanych ligandów różnych celów molekularnych. Otrzymane wyniki w postaci listy celów i prawdopodobieństwa wiązania się z danym celem, zostały poddane analizie. Dla większości badanych związków jako pierwszy potencjalny cel molekularny narzędzie SwissTarget przewidziało receptor NK1, jednakże na podstawie naszych danych eksperymentalnych odrzucono to białko. W wyniku dalszych analiz

odrzucono także białka związane z metabolizmem, przewodnictwem nerwowym, czy białka transporterowe.

VII.10. Dokowanie do receptora CCK2

Ligandy z serii W-71 zadokowano do receptora CCK2 za pomocą programu AutoDock Vina [206]. Struktury ligandów przygotowano w programie Biovia Discovery Studio [319]. Geometrie ligandów zminimalizowano za pomocą funkcji “clean geometry”. W szczególności zwrócono uwagę na uzyskanie płaskiej geometrii wiązania hydrazonowego oraz płaskiej geometrii wiązania estrowego i konformacji *trans*. W przypadku ligandów z wolną grupą aminową, ustawiono stan protonowania +1. Przygotowano różne konformery ligandów (ligandy dokowano w konformacjach *cis* i *trans* wokół wiązań amidowych). W programie AutoDock Tools [320] dodatkowo ręcznie zahamowano rotację wokół wiązania N-N w układzie hydrazonu oraz C-O w układzie karbaminianu. Do dokowania użyto modelu homologicznego receptora CCK2 (w konformacji nieaktywnej), pobranego z bazy danych GPCRdb [324]. Do struktury białka dodano wodory i minimalizowano. Stany protonowania białka ustawiono zgodnie z oczekiwanymi przy pH=7.4. Pudło (box) dokowania ustawiono ręcznie tak, aby obejmowało reszty Arg50, Arg208, Trp346 i rozszerzono, uzyskując wymiary: 31.96 Å x 34.38 Å x 40.76 Å. Wygenerowano po 9 pól i na podstawie inspekcji wizualnej wybrano najlepszą. Dla liganda Z-dW-71 otrzymane wyniki zwizualizowano w programie PyMOL [325].

VII.11. Badania przeprowadzone przez współpracowników

Rentgenograficzne pomiary strukturalne wykonał dr Krzysztof Łyczko z Instytutu Chemii i Techniki Jądrowej w Warszawie. Metodologia pomiaru została opisana w publikacji [4].

Oznaczenia powinowactwa receptorowego do receptora NK1 wykonała dr Joanna Matalińska z Zakładu Neuropeptydów IMDiK PAN. Metodologia oznaczeń została opisana w pracach [1] i [2]. W skrócie, było to oznaczenie powinowactwa przez pomiar kompetycyjnego wypierania radioliganda selektywnego wobec receptora NK1 ($[^3\text{H}]$ -Substancji P).

Oznaczenia wpływu na przeżywalność komórek (cytotoksyczność) wykonała dr Joanna Matalińska z Zakładu Neuropeptydów IMDiK PAN. Metodyka pomiaru opisana jest

w patencie [3]. W skrócie, za pomocą testu MTS oznaczano wpływ badanych związków na aktywność metaboliczną komórek czerniaka COLO679 i komórek prawidłowych fibroblastów BJ.

VIII. Literatura

1. Matalińska, J.; Kosińska, K.; Halik, P.K.; Koźmiński, P.; Lipiński, P.F.J.; Gniazdowska, E.; Misicka, A. Novel NK1R-Targeted ⁶⁸Ga-/¹⁷⁷Lu-Radioconjugates with Potential Application against Glioblastoma Multiforme: Preliminary Exploration of Structure–Activity Relationships. *Int. J. Mol. Sci.* **2022**, *23*, 1214.
2. Witoszka, K.; Matalińska, J.; Misicka, A.; Lipiński, P.F.J. Moving out of CF 3 -Land: Synthesis, Receptor Affinity, and in silico Studies of NK1 Receptor Ligands Containing a Pentafluorosulfanyl (SF 5) Group. *ChemMedChem* **2023**, *18*.
3. Lipiński, P.F.J.; Matalińska, J.; Kosińska, K.; Dyniewicz, J.; Misicka, A. Patent PL244031B1: Nowe pochodne peptydowe, zwłaszcza stosowania w terapii nowotworów 2023.
4. Rode, J.E.; Lyczko, K.; Kosińska, K.; Matalińska, J.; Dyniewicz, J.; Misicka, A.; Dobrowolski, J.C.; Lipiński, P.F.J. The solid state VCD of a novel N-acylhydrazone trifluoroacetate. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *269*, 120761.
5. Datar, P.; Srivastava, S.; Coutinho, E.; Govil, G. Substance P: Structure, Function, and Therapeutics. *Curr. Top. Med. Chem.* **2004**, *4*, 75–103.
6. Pennefather, J.N.; Lecci, A.; Candenas, M.L.; Patak, E.; Pinto, F.M.; Maggi, C.A. Tachykinins and tachykinin receptors: a growing family. *Life Sci.* **2004**, *74*, 1445–1463.
7. Bost, Kenneth, L. Tachykinin-mediated modulation of the immune response. *Front. Biosci.* **2004**, *9*, 3331.
8. Rameshwar, P. Substance P: A Regulatory Neuropeptide for Hematopoiesis and Immune Functions. *Clin. Immunol. Immunopathol.* **1997**, *85*, 129–133.
9. Huang, S.-C.; Korlipara, V.L. Neurokinin-1 receptor antagonists: a comprehensive patent survey. *Expert Opin. Ther. Pat.* **2010**, *20*, 1019–1045.
10. Gaddum, J.H.; Schild, H. Depressor substances in extracts of intestine. *J. Physiol.* **1934**, *83*, 1–14.
11. Harrison, S. Substance P. *Int. J. Biochem. Cell Biol.* **2001**, *33*, 555–576.
12. KIMURA, S.; OKADA, M.; SUGITA, Y.; KANAZAWA, I.; MUNEKATA, E. Novel neuropeptides, neurokinin .ALPHA. and .BETA., isolated from porcine spinal cord. *Proc. Japan Acad. Ser. B* **1983**, *59*, 101–104.
13. Kangawa, K.; Minamino, N.; Fukuda, A.; Matsuo, H. Neuromedin K: A novel mammalian tachykinin identified in porcine spinal cord. *Biochem. Biophys. Res. Commun.* **1983**, *114*, 533–540.
14. Onaga, T. Tachykinin: recent developments and novel roles in health and disease. *Biomol. Concepts* **2014**, *5*, 225–243.
15. Rodríguez, F.D.; Coveñas, R. The Neurokinin-1 Receptor: Structure Dynamics and Signaling. *Receptors* **2022**, *1*, 54–71.
16. Douglas, S.D.; Leeman, S.E. Neurokinin-1 receptor: functional significance in the immune system in reference to selected infections and inflammation. *Ann. N. Y. Acad. Sci.* **2011**, *1217*, 83–95.
17. Maggi, C.A. The mammalian tachykinin receptors. *Gen. Pharmacol. Vasc. Syst.* **1995**, *26*, 911–944.
18. Ptak, Krzysztof, Lewandowski, M.H.; Monteau, R. Neurokininy i ich receptory. *POLISH Soc. CELL Biol.* 2000, *27*, 273–285.

19. Garcia-Recio, S.; Gascón, P. Biological and Pharmacological Aspects of the NK1-Receptor. *Biomed Res. Int.* 2015, 2015, Article ID 495704.
20. García-Aranda, M.; Téllez, T.; McKenna, L.; Redondo, M. Neurokinin-1 Receptor (NK-1R) Antagonists as a New Strategy to Overcome Cancer Resistance. *Cancers (Basel)*. **2022**, *14*, 2255.
21. Kowalczyk, S.; Hetmann, A. Receptory sprzężone z białkami G, heterotrimeryczne białka G i białka efektorowe roślin. *Postepy Biochem.* 2008, *54*, 412–422.
22. Muñoz, M.; Coveñas, R. Substance P. In *Encyclopedia of Endocrine Diseases*; Elsevier, 2018; pp. 571–578.
23. Yin, J.; Chapman, K.; Clark, L.D.; Shao, Z.; Borek, D.; Xu, Q.; Wang, J.; Rosenbaum, D.M. Crystal structure of the human NK 1 tachykinin receptor. *Proc. Natl. Acad. Sci.* **2018**, *115*, 13264–13269.
24. Harris, J.A.; Faust, B.; Gondin, A.B.; Dämgen, M.A.; Suomivuori, C.-M.; Veldhuis, N.A.; Cheng, Y.; Dror, R.O.; Thal, D.M.; Manglik, A. Selective G protein signaling driven by substance P–neurokinin receptor dynamics. *Nat. Chem. Biol.* **2022**, *18*, 109–115.
25. Thom, C.; Ehrenmann, J.; Vacca, S.; Waltenspühl, Y.; Schöppe, J.; Medalia, O.; Plückthun, A. Structures of neurokinin 1 receptor in complex with G q and G s proteins reveal substance P binding mode and unique activation features. *Sci. Adv.* **2021**, *7*.
26. Chen, S.; Lu, M.; Liu, D.; Yang, L.; Yi, C.; Ma, L.; Zhang, H.; Liu, Q.; Frimurer, T.M.; Wang, M.-W.; et al. Human substance P receptor binding mode of the antagonist drug aprepitant by NMR and crystallography. *Nat. Commun.* **2019**, *10*, 638.
27. Schöppe, J.; Ehrenmann, J.; Klenk, C.; Rucktooa, P.; Schütz, M.; Doré, A.S.; Plückthun, A. Crystal structures of the human neurokinin 1 receptor in complex with clinically used antagonists. *Nat. Commun.* **2019**, *10*, 17.
28. Chen, M.; Schmid, M.F.; Chiu, W. Improving resolution and resolvability of single-particle cryoEM structures using Gaussian mixture models. *Nat. Methods* **2024**, *21*, 37–40.
29. Sun, W.; Yuan, Q.; Zhang, H.; Yang, F.; Ling, S.; Luo, Y.; Lv, P.; Eric Xu, H.; Tian, C.; Yin, W.; et al. Structural insights into the activation of neurokinin 2 receptor by neurokinin A. *Cell Discov.* **2022**, *8*, 72.
30. Neer, E.J. Heterotrimeric G proteins: Organizers of transmembrane signals. *Cell* **1995**, *80*, 249–257.
31. Munoz, M.; Rosso, M.; Covenas, R. A New Frontier in the Treatment of Cancer: NK-1 Receptor Antagonists. *Curr. Med. Chem.* **2010**, *17*, 504–516.
32. Fukuhara, S.; Shimizu, M.; Matsushima, H.; Mukai, H.; Munekata, E. Signaling pathways via NK1 receptors and their desensitization in an AR42J cell line. *Peptides* **1998**, *19*, 1349–1357.
33. Khawaja, A.M.; Rogers, D.F. Tachykinins: Receptor to effector. *Int. J. Biochem. Cell Biol.* 1996, *28*.
34. Xu, M.; Seas, A.; Kiyani, M.; Ji, K.S.Y.; Bell, H.N. A temporal examination of calcium signaling in cancer- from tumorigenesis, to immune evasion, and metastasis. *Cell Biosci.* **2018**, *8*, 25.
35. Nakajima, Y.; Tsuchida, K.; Negishi, M.; Ito, S.; Nakanishi, S. Direct linkage of three tachykinin receptors to stimulation of both phosphatidylinositol hydrolysis and cyclic AMP cascades in transfected Chinese hamster ovary cells. *J. Biol. Chem.* **1992**, *267*, 2437–2442.
36. Laniyonu, A.; Sliwinski-Lis, E.; Fleming, N. Different tachykinin receptor subtypes are coupled to the phosphoinositide or cyclic AMP signal transduction pathways in rat submandibular cells. *FEBS Lett.*

- 1988**, 240, 186–190.
37. Meshki, J.; Douglas, S.D.; Lai, J.-P.; Schwartz, L.; Kilpatrick, L.E.; Tuluc, F. Neurokinin 1 Receptor Mediates Membrane Blebbing in HEK293 Cells through a Rho/Rho-associated Coiled-coil Kinase-dependent Mechanism. *J. Biol. Chem.* **2009**, 284, 9280–9289.
 38. Garcia, M.; Sakamoto, K.; Shigekawa, M.; Nakanishi, S.; Ito, S. Multiple mechanisms of arachidonic acid release in Chinese hamster ovary cells transfected with cDNA of substance P receptor. *Biochem. Pharmacol.* **1994**, 48, 1735–1741.
 39. Steinhoff, M.S.; von Mentzer, B.; Geppetti, P.; Pothoulakis, C.; Bunnett, N.W. Tachykinins and Their Receptors: Contributions to Physiological Control and the Mechanisms of Disease. *Physiol. Rev.* **2014**, 94, 265–301.
 40. Della Rocca, G.J.; van Biesen, T.; Daaka, Y.; Luttrell, D.K.; Luttrell, L.M.; Lefkowitz, R.J. Ras-dependent Mitogen-activated Protein Kinase Activation by G Protein-coupled Receptors. *J. Biol. Chem.* **1997**, 272, 19125–19132.
 41. Zhang, X.-W.; Li, L.; Hu, W.-Q.; Hu, M.-N.; Tao, Y.; Hu, H.; Miao, X.-K.; Yang, W.-L.; Zhu, Q.; Mou, L.-Y. Neurokinin-1 receptor promotes non-small cell lung cancer progression through transactivation of EGFR. *Cell Death Dis.* **2022**, 13, 41.
 42. Murphy, J.E.; Roosterman, D.; Cottrell, G.S.; Padilla, B.E.; Feld, M.; Brand, E.; Cedron, W.J.; Bunnett, N.W.; Steinhoff, M. Protein phosphatase 2A mediates resensitization of the neurokinin 1 receptor. *Am. J. Physiol. Physiol.* **2011**, 301, C780–C791.
 43. Jensen, D.D.; Lieu, T.; Halls, M.L.; Veldhuis, N.A.; Imlach, W.L.; Mai, Q.N.; Poole, D.P.; Quach, T.; Aurelio, L.; Conner, J.; et al. Neurokinin 1 receptor signaling in endosomes mediates sustained nociception and is a viable therapeutic target for prolonged pain relief. *Sci. Transl. Med.* **2017**, 9.
 44. DeFea, K.A.; Vaughn, Z.D.; O'Bryan, E.M.; Nishijima, D.; Déry, O.; Bunnett, N.W. The proliferative and antiapoptotic effects of substance P are facilitated by formation of a β -arrestin-dependent scaffolding complex. *Proc. Natl. Acad. Sci.* **2000**, 97, 11086–11091.
 45. Fong, T.M.; Anderson, S.A.; Yu, H.; Huang, R.; Strader, C.D. Differential activation of intracellular effector by two isoforms of human neurokinin-1 receptor. *Mol. Pharmacol.* **1992**, 41.
 46. Lai, J.-P.; Lai, S.; Tuluc, F.; Tansky, M.F.; Kilpatrick, L.E.; Leeman, S.E.; Douglas, S.D. Differences in the length of the carboxyl terminus mediate functional properties of neurokinin-1 receptor. *Proc. Natl. Acad. Sci.* **2008**, 105, 12605–12610.
 47. Spitsin, S.; Pappa, V.; Douglas, S.D. Truncation of neurokinin-1 receptor—Negative regulation of substance P signaling. *J. Leukoc. Biol.* **2018**, 103, 1043–1051.
 48. Muñoz, M.; Coveñas, R. Neurokinin receptor antagonism: a patent review (2014-present). *Expert Opin. Ther. Pat.* **2020**, 30, 527–539.
 49. Muñoz, M.; Coveñas, R. Involvement of substance P and the NK-1 receptor in human pathology. *Amino Acids* **2014**, 46, 1727–1750.
 50. Muñoz, M.; Coveñas, R. Involvement of substance P and the NK-1 receptor in cancer progression. *Peptides* **2013**, 48, 1–9.
 51. Mistrova, E.; Kruzliak, P.; Chottova Dvorakova, M. Role of substance P in the cardiovascular system. *Neuropeptides* **2016**, 58, 41–51.

52. Lisowska Barbara Rola i znaczenie substancji P Role and significance of substance P. *Geriatrics* **2007**, *1*, 18–21.
53. Singh, R.B.; Naderi, A.; Cho, W.; Ortiz, G.; Musayeva, A.; Dohman, T.H.; Chen, Y.; Ferrari, G.; Dana, R. Modulating the tachykinin: Role of substance P and neurokinin receptor expression in ocular surface disorders. *Ocul. Surf.* **2022**, *25*, 142–153.
54. Mantyh, P.W. Neurobiology of substance P and the NK1 receptor. *J. Clin. Psychiatry* **2002**, 6–10.
55. Armstrong, D.M.; Pickel, V.M.; Joh, T.H.; Reis, D.J.; Miller, R.J. Immunocytochemical localization of catecholamine synthesizing enzymes and neuropeptides in area postrema and medial nucleus tractus solitarius of rat brain. *J. Comp. Neurol.* **1981**, *196*, 505–517.
56. Corrigan, F.; Vink, R.; Turner, R.J. Inflammation in acute CNS injury: a focus on the role of substance P. *Br. J. Pharmacol.* **2016**, *173*, 703–715.
57. McMahan, S.B.; Cafferty, W.B.J.; Marchand, F. Immune and glial cell factors as pain mediators and modulators. *Exp. Neurol.* **2005**, *192*, 444–462.
58. Kleczkowska, P.; Nowicka, K.; Bujalska-Zadrozny, M.; Hermans, E. Neurokinin-1 receptor-based bivalent drugs in pain management: The journey to nowhere? *Pharmacol. Ther.* **2019**, *196*, 44–58.
59. Ho, W.-Z.; Douglas, S.D. Substance P and neurokinin-1 receptor modulation of HIV. *J. Neuroimmunol.* **2004**, *157*, 48–55.
60. Redkiewicz, P. The Regenerative Potential of Substance P. *Int. J. Mol. Sci.* **2022**, *23*, 750.
61. Quartara, L.; Maggi, C.. The tachykinin NK1 receptor. Part II: Distribution and pathophysiological roles. *Neuropeptides* **1998**, *32*, 1–49.
62. Groneberg, D.; Harrison, S.; Thai Dinh, Q.; Geppetti, P.; Fischer, A. Tachykinins in the Respiratory Tract. *Curr. Drug Targets* **2006**, *7*, 1005–1010.
63. Esteban, F.; Ramos-García, P.; Muñoz, M.; González-Moles, M.Á. Substance P and Neurokinin 1 Receptor in Chronic Inflammation and Cancer of the Head and Neck: A Review of the Literature. *Int. J. Environ. Res. Public Health* **2021**, *19*, 375.
64. Gadais, C.; Ballet, S. The Neurokinins: Peptidomimetic Ligand Design and Therapeutic Applications. *Curr. Med. Chem.* **2020**, *27*, 1515–1561.
65. JONES, S.; GIBBINS, J. The neurokinin 1 receptor: a potential new target for anti-platelet therapy? *Curr. Opin. Pharmacol.* **2008**, *8*, 114–119.
66. Rosso, M.; Muñoz, M.; Berger, M. The Role of Neurokinin-1 Receptor in the Microenvironment of Inflammation and Cancer. *Sci. World J.* **2012**, *2012*, 1–21.
67. Munoz, M.; Covenas, R. Neurokinin-1 receptor: a new promising target in the treatment of cancer. *Discov. Med.* **2010**, *10*, 305–313.
68. Drugbank Neurokinin 1 receptor https://go.drugbank.com/bio_entities/BE0000384.
69. Hill, R. NK1 (substance P) receptor antagonists – why are they not analgesic in humans? *Trends Pharmacol. Sci.* **2000**, *21*, 244–246.
70. Kramer, M.S.; Cutler, N.; Feighner, J.; Shrivastava, R.; Carman, J.; Sramek, J.J.; Reines, S.A.; Liu, G.; Snavely, D.; Wyatt-Knowles, E.; et al. Distinct Mechanism for Antidepressant Activity by Blockade of Central Substance P Receptors. *Science (80-.)*. **1998**, *281*, 1640–1645.
71. Keller, M.; Montgomery, S.; Ball, W.; Morrison, M.; Snavely, D.; Liu, G.; Hargreaves, R.; Hietala, J.;

- Lines, C.; Beebe, K.; et al. Lack of Efficacy of the Substance P (Neurokinin1 Receptor) Antagonist Aprepitant in the Treatment of Major Depressive Disorder. *Biol. Psychiatry* **2006**, *59*, 216–223.
72. George, D.T.; Gilman, J.; Hersh, J.; Thorsell, A.; Herion, D.; Geyer, C.; Peng, X.; Kielbasa, W.; Rawlings, R.; Brandt, J.E.; et al. Neurokinin 1 Receptor Antagonism as a Possible Therapy for Alcoholism. *Science (80-.)*. **2008**, *319*, 1536–1539.
73. Coe, M.A.; Lofwall, M.R.; Vessels, V.; Nuzzo, P.A.; Walsh, S.L. Evaluation of tradipitant, a selective NK1 antagonist, on response to oxycodone in humans. *Psychopharmacology (Berl)*. **2021**, *238*, 1857–1866.
74. Yosipovitch, G.; Ständer, S.; Kerby, M.B.; Larrick, J.W.; Perlman, A.J.; Schnipper, E.F.; Zhang, X.; Tang, J.Y.; Luger, T.; Steinhoff, M. Serlopitant for the treatment of chronic pruritus: Results of a randomized, multicenter, placebo-controlled phase 2 clinical trial. *J. Am. Acad. Dermatol.* **2018**, *78*, 882-891.e10.
75. Torres, T.; Fernandes, I.; Selores, M.; Alves, R.; Lima, M. Aprepitant: Evidence of its effectiveness in patients with refractory pruritus continues. *J. Am. Acad. Dermatol.* **2012**, *66*, e14–e15.
76. Welsh, S.E.; Xiao, C.; Kaden, A.R.; Brzezynski, J.L.; Mohrman, M.A.; Wang, J.; Smieszek, S.P.; Przychodzen, B.; Ständer, S.; Polymeropoulos, C.; et al. Neurokinin-1 receptor antagonist tradipitant has mixed effects on itch in atopic dermatitis: results from EPIONE, a randomized clinical trial. *J. Eur. Acad. Dermatology Venereol.* **2021**, *35*.
77. Le Gall-Ianotto, C.; Verdet, R.; Nowak, E.; Le Roux, L.; Gasse, A.; Fiedler, A.; Carlhant-Kowalski, D.; Marcorelles, P.; Misery, L.; Ianotto, J.C. Rationale and design of the multicentric, double-blind, double-placebo, randomized trial APrepitant versus HYdroxyzine in association with cyto reductive treatments for patients with myeloproliferative neoplasia suffering from Persistent Aquagenic Pruritus. *Trials* **2021**, *22*, 938.
78. Pariser, D.M.; Bagel, J.; Lebowhl, M.; Yosipovitch, G.; Chien, E.; Spellman, M.C. Serlopitant for psoriatic pruritus: A phase 2 randomized, double-blind, placebo-controlled clinical trial. *J. Am. Acad. Dermatol.* **2020**, *82*, 1314–1320.
79. Smith, J.; Allman, D.; Badri, H.; Miller, R.; Morris, J.; Satia, I.; Wood, A.; K. Trower, M. The Neurokinin-1 Receptor Antagonist Orvepitant Is a Novel Antitussive Therapy for Chronic Refractory Cough. *Chest* **2020**, *157*, 111–118.
80. Efficacy and Safety Study of Orvepitant for Chronic Cough in Patients With Idiopathic Pulmonary Fibrosis (IPF-COMFORT) <https://classic.clinicaltrials.gov/ct2/show/study/NCT05185089>.
81. Mullard, A. FDA approves first-in-class NK3 receptor antagonist for hot flushes. *Nat. Rev. Drug Discov.* **2023**, *22*, 526–526.
82. Almeida, T.A.; Rojo, J.; Nieto, P.M.; Pinto, F.M.; Hernandez, M.; Martín, J.D.; Candenias, M.L. Tachykinins and Tachykinin Receptors: Structure and Activity Relationships. *Curr. Med. Chem.* **2004**, *11*, 2045–2081.
83. Leander, S.; Håkanson, R.; Rosell, S.; Folkers, K.; Sundler, F.; Tornqvist, K. A specific substance P antagonist blocks smooth muscle contractions induced by non-cholinergic, non-adrenergic nerve stimulation. *Nature* **1981**, *294*, 467–469.
84. Folkers, K.; Håkanson, R.; Hörig, J.; Jie-Cheng, X.; Leander, S. Biological evaluation of substance P

- antagonists. *Br. J. Pharmacol.* **1984**, *83*, 449–456.
85. Quartara, L.; Maggi, C.. The tachykinin NK1 receptor. Part I: Ligands and mechanisms of cellular activation. *Neuropeptides* **1997**, *31*, 537–563.
86. Folkers, K.; Feng, D.M.; Asano, N.; Håkanson, R.; Weisenfeld-Hallin, Z.; Leander, S. Spantide II, an effective tachykinin antagonist having high potency and negligible neurotoxicity. *Proc. Natl. Acad. Sci.* **1990**, *87*, 4833–4835.
87. Maggi, C.A.; Patacchini, R.; Dong Mei Feng; Folkers, K. Activity of spantide I and II at various tachykinin receptors and NK-2 tachykinin receptor subtypes. *Eur. J. Pharmacol.* **1991**, *199*, 127–129.
88. Lavielle, S.; Brunissen, A.; Carruette, A.; Garret, C.; Chassaing, G. Highly potent substance P antagonists substituted with β -phenyl- or β -benzyl-proline at position 10. *Eur. J. Pharmacol.* **1994**, *258*, 273–276.
89. Morimoto, H.; Murai, M.; Maeda, Y.; Hagiwara, D.; Miyake, H.; Matsuo, M.; Fujii, T. FR 113680: a novel tripeptide substance P antagonist with NK1 receptor selectivity. *Br. J. Pharmacol.* **1992**, *106*, 123–126.
90. Fujii, T.; Murai, M.; Morimoto, H.; Maeda, Y.; Yamaoka, M.; Hagiwara, D.; Miyake, H.; Ikari, N.; Matsuo, M. Pharmacological profile of a high affinity dipeptide NK1 receptor antagonist, FK888. *Br. J. Pharmacol.* **1992**, *107*, 785–789.
91. Wang, Z.-Y.; Tung, S.R.; Strichartz, G.R.; Håkanson, R. Investigation of the specificity of FK 888 as a tachykinin NK1 receptor antagonist. *Br. J. Pharmacol.* **1994**, *111*, 1342–1346.
92. Murai, M.; Maeda, Y.; Hagiwara, D.; Miyake, H.; Ikari, N.; Matsuo, M.; Fujii, T. Effects of an NK1 receptor antagonist, FK888, on constriction and plasma extravasation induced in guinea pig airway by neurokinins and capsaicin. *Eur. J. Pharmacol.* **1993**, *236*, 7–13.
93. HASHIMOTO, M.; HAYASHI, K.; MURAI, M.; FUJII, T.; NISHIKAWA, M.; KIYOTO, S.; OKUHARA, M.; KOHSAKA, M.; IMANAKA, H. WS9326A, a novel tachykinin antagonist isolated from *Streptomyces violaceusniger* No. 9326. II Biological and pharmacological properties of WS9326A and tetrahydro-WS9326A (FK224). *J. Antibiot. (Tokyo)*. **1992**, *45*, 1064–1070.
94. Qtsuka, M.; Yoshioka, K.; Yanagisawa, M.; Suzuki, H.; Zhao, F.-Y.; Guo, J.-Z.; Hosoki, R.; Kurihara, T. Use of NK 1 receptor antagonists in the exploration of physiological functions of substance P and neurokinin A. *Can. J. Physiol. Pharmacol.* **1995**, *73*, 903–907.
95. Appell, K.C.; Babb, B.E.; Goswami, R.; Hall, P.L.; Lawrence Kristine B.; Logan, M.E.; Przyklek-Elling, R.; Tomczuk, B.E.; Venepalli, B.R.; Yanni, J.M. Imidazo[4,5-b]quinoxaline cyanines as neurokinin antagonists. *J. Med. Chem.* **1991**, *34*, 1751–1753.
96. Venepalli, B R Aimone, L D Appell, K.C. et al. Synthesis and Androstano[3, substance P receptor binding activity of 2-b]pyrimido[1, 2-a]benzimidazoles. *J. Med. Chem.* **1992**, *35*, 374–378.
97. Garret, C.; Carruette, A.; Fardin, V.; Moussaoui, S.; Peyronel, J.F.; Blanchard, J.C.; Laduron, P.M. Pharmacological properties of a potent and selective nonpeptide substance P antagonist. *Proc. Natl. Acad. Sci.* **1991**, *88*, 10208–10212.
98. Tabart, M.; Peyronel, J.-F. Synthesis of RPR 100893, prototype of a new series of potent and selective non peptide NK1 antagonists : the triarylperhydroisoindolols. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 673–676.

99. Diener, H.-C. Rpr100893, A Substance-P Antagonist, is Not Effective in the Treatment of Migraine Attacks. *Cephalalgia* **2003**, *23*, 183–185.
100. Mutti, S.; Daubié, C.; Malpart, J.; Radisson, X. Practical enantiospecific synthesis of RPR 111905: A novel non-peptide substance P antagonist. *Tetrahedron Lett.* **1996**, *37*, 8743–8746.
101. Swain, C.J. 2 Neurokinin receptor Antagonists. In *Progress in medical chemistry*; 1998; Vol. 35, pp. 57–82.
102. Snider, R.M.; Constantine, J.W.; Lowe, J.A.; Longo, K.P.; Lebel, W.S.; Woody, H.A.; Drozda, S.E.; Desai, M.C.; Vinick, F.J.; Spencer, R.W.; et al. A Potent Nonpeptide Antagonist of the Substance P (NK 1) Receptor. *Science (80-.)*. **1991**, *251*, 435–437.
103. Lowe, J.A.; Drozda, S.E.; Snider, R.M.; Longo, K.P.; Zorn, S.H.; Morrone, J.; Jackson, E.R.; McLean, S.; Bryce, D.K. The discovery of (2S,3S)-cis-2-(diphenylmethyl)-N-[(2-methoxyphenyl)methyl]-1-azabicyclo[2.2.2]octan-3-amine as a novel, nonpeptide substance P antagonist. *J. Med. Chem.* **1992**, *35*, 2591–2600.
104. Schmidt, A.W.; McLean, S.; Heym, J. The substance P receptor antagonist CP-96,345 interacts with Ca²⁺ channels. *Eur. J. Pharmacol.* **1992**, *215*, 351–352.
105. Lowe, J.A.; Drozda, S.E.; McLean, S.; Crawford, R.T.; Bryce, D.K.; Bordner, J. N-alkyl quinuclidinium substance P antagonists. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1153–1156.
106. Evangelista, S. Ezlopitant. Pfizer. *Curr. Opin. Investigational Drugs* **2001**, *2*, 1441–1443.
107. Lewicki, J. Maropitant–nowy lek przeciwwymiotny z grupy antagonistów receptora neurokininowego typu-1. Część II. Toksyczność, działania niepożądane, zastosowanie kliniczne. *Życie Wet* **2008**, *83*, 594–601.
108. Desai, M.C.; Lefkowitz, S.L.; Thadeio, P.F.; Longo, K.P.; Snider, R.M. Discovery of a potent substance P antagonist: recognition of the key molecular determinant. *J. Med. Chem.* **1992**, *35*, 4911–4913.
109. McLean, S.; Snider, R.M.; Desai, M.C.; Rosen, T.; Bryce, D.K.; Longo, K.P.; Schmidt, A.W.; Heym, J. CP-99,994, a nonpeptide antagonist of the tachykinin NK1 receptor. *Regul. Pept.* **1993**, *46*, 329–331.
110. Dionne, R.A.; Max, M.B.; Gordon, S.M.; Parada, S.; Sang, C.; Gracely, R.H.; Sethna, N.F.; MacLean, D.B. The substance P receptor antagonist CP-99,994 reduces acute postoperative pain*. *Clin. Pharmacol. Ther.* **1998**, *64*, 562–568.
111. Ward, P.; Armour, D.R.; Bays, D.E.; Evans, B.; Giblin, G.M.P.; Heron, N.; Hubbard, T.; Liang, K.; Middlemiss, D. Discovery of an Orally Bioavailable NK1 Receptor Antagonist, (2S,3S)-(2-Methoxy-5-tetrazol-1-ylbenzyl)(2-phenylpiperidin-3-yl)amine (GR203040), with Potent Antiemetic Activity. *J. Med. Chem.* **1995**, *38*, 4985–4992.
112. Armour, D.R.; Chung, K.M.L.; Congreve, M.; Evans, B.; Guntrip, S.; Hubbard, T.; Kay, C.; Middlemiss, D.; Mordaunt, J.E.; Pegg, N.A.; et al. Tetrazole NK1 receptor antagonists: The identification of an exceptionally potent orally active antiemetic compound. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1015–1020.
113. Quartara, L.; Altamura, M. Tachykinin Receptors Antagonists: From Research to Clinic. *Curr. Drug Targets* **2006**, *7*, 975–992.
114. McLean, S.; Ganong, A.; Seymour, P.A.; Bryce, D.K.; Crawford, R.T.; Morrone, J.; Reynolds, L.S.; Schmidt, A.W.; Zorn, S.; Watson, J.; et al. Characterization of CP-122,721; a nonpeptide antagonist of

- the neurokinin NK1 receptor. *J. Pharmacol. Exp. Ther.* **1996**, *277*, 900–908.
115. Rosen, T.J.; Coffman, K.J.; McLean, S.; Crawford, R.T.; Bryce, D.K.; Gohda, Y.; Tsuchiya, M.; Nagahisa, A.; Nakane, M.; Lowe, J.A. Synthesis and structure-activity relationships of CP-122,721, a second-generation NK-1 receptor antagonist. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 281–284.
 116. Seward, E.M.; Swain, C.J.; Merchant, K.J.; Owen, S.N.; Sabin, V.; Cascieri, M.A.; Sadowski, S.; Strader, C.; Baker, R. Quinuclidine-based NK-1 antagonists I: 3-benzyloxy-1-azabicyclo[2.2.2]octanes. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1361–1366.
 117. Swain, C.J.; Seward, E.M.; Sabin, V.; Owen, S.; Baker, R.; Cascieri, M.A.; Sadowski, S.; Strader, C.; Ball, R.G. Quinuclidine based NK-1 antagonists 2: determination of the absolute stereochemical requirements. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1703–1706.
 118. Harrison, T.; Williams, B.J.; Swain, C.J.; Ball, R.G. Piperidine-ether based hNK1 antagonists 1: Determination of the relative and absolute stereochemical requirements. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2545–2550.
 119. Harrison, T.; Owens, A.P.; Williams, B.J.; Swain, C.J.; Baker, R.; Hutson, P.H.; Sadowski, S.; Cascieri, M.A. Piperidine-ether based hNK1 antagonists 2: Investigation of the effect of N-substitution. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 209–212.
 120. Ladduwahetty, T.; Baker, R.; Cascieri, M.A.; Chambers, M.S.; Haworth, K.; Keown, L.E.; MacIntyre, D.E.; Metzger, J.M.; Owen, S.; Rycroft, W.; et al. N -Heteroaryl-2-phenyl-3-(benzyloxy)piperidines: A Novel Class of Potent Orally Active Human NK 1 Antagonists. *J. Med. Chem.* **1996**, *39*, 2907–2914.
 121. Hale, J.J.; Mills, S.G.; MacCoss, M.; Shah, S.K.; Qi, H.; Mathre, D.J.; Cascieri, M.A.; Sadowski, S.; Strader, C.D.; MacIntyre, D.E.; et al. 2(S)-((3,5-Bis(trifluoromethyl)benzyl)oxy)-3(S)-phenyl-4-((3-oxo-1,2,4-triazol- 5-yl)methyl)morpholine (1): A Potent, Orally Active, Morpholine-Based Human Neurokinin-1 Receptor Antagonist. *J. Med. Chem.* **1996**, *39*, 1760–1762.
 122. Hale, J.J.; Mills, S.G.; MacCoss, M.; Finke, P.E.; Cascieri, M.A.; Sadowski, S.; Ber, E.; Chicchi, G.G.; Kurtz, M.; Metzger, J.; et al. Structural Optimization Affording 2-(R)-(1-(R)-3,5-Bis(trifluoromethyl)phenylethoxy)-3-(S)-(4-fluoro)phenyl-4- (3-oxo-1,2,4-triazol-5-yl)methylmorpholine, a Potent, Orally Active, Long-Acting Morpholine Acetal Human NK-1 Receptor Antagonist. *J. Med. Chem.* **1998**, *41*, 4607–4614.
 123. Humphrey, J. Medicinal Chemistry of Selective Neurokinin-1 Antagonists. *Curr. Top. Med. Chem.* **2003**, *3*, 1423–1435.
 124. Borsook, D.; Upadhyay, J.; Klimas, M.; Schwarz, A.J.; Coimbra, A.; Baumgartner, R.; George, E.; Potter, W.Z.; Large, T.; Bleakman, D.; et al. Decision-making using fMRI in clinical drug development: revisiting NK-1 receptor antagonists for pain. *Drug Discov. Today* **2012**, *17*, 964–973.
 125. Kramer, M.S.; Winokur, A.; Kelsey, J.; Preskorn, S.H.; Rothschild, A.J.; Snavelly, D.; Ghosh, K.; Ball, W.A.; Reines, S.A.; Munjack, D.; et al. Demonstration of the Efficacy and Safety of a Novel Substance P (NK1) Receptor Antagonist in Major Depression. *Neuropsychopharmacology* **2004**, *29*, 385–392.
 126. Muñoz, M.; Rosso, M. The NK-1 receptor antagonist aprepitant as a broad spectrum antitumor drug. *Invest. New Drugs* **2010**, *28*, 187–193.
 127. Muñoz, M.; Coveñas, R.; Esteban, F.; Redondo, M. The substance P/NK-1 receptor system: NK-1 receptor antagonists as anti-cancer drugs. *J. Biosci.* **2015**, *40*, 441–463.

128. Liu, X.; Zhu, Y.; Zheng, W.; Qian, T.; Wang, H.; Hou, X. Antagonism of NK-1R using aprepitant suppresses inflammatory response in rheumatoid arthritis fibroblast-like synoviocytes. *Artif. Cells, Nanomedicine, Biotechnol.* **2019**, *47*, 1628–1634.
129. Manak, M.M.; Moshkoff, D.A.; Nguyen, L.T.; Meshki, J.; Tebas, P.; Tuluc, F.; Douglas, S.D. Anti-HIV-1 activity of the neurokinin-1 receptor antagonist aprepitant and synergistic interactions with other antiretrovirals. *AIDS* **2010**, *24*, 2789–2796.
130. Navari, R.M. Fosaprepitant (MK-0517): a neurokinin-1 receptor antagonist for the prevention of chemotherapy-induced nausea and vomiting. *Expert Opin. Investig. Drugs* **2007**, *16*, 1977–1985.
131. Tattersall, F.; Rycroft, W.; Cumberbatch, M.; Mason, G.; Tye, S.; Williamson, D.; Hale, J.; Mills, S.; Finke, P.; MacCoss, M.; et al. The novel NK1 receptor antagonist MK-0869 (L-754,030) and its water soluble phosphoryl prodrug, L-758,298, inhibit acute and delayed cisplatin-induced emesis in ferrets. *Neuropharmacology* **2000**, *39*, 652–663.
132. Colon-Gonzalez, F.; Kraft, W.K. Pharmacokinetic evaluation of fosaprepitant dimeglumine. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 1277–1286.
133. Emonds-Alt, X.; Doutremepuich, J.-D.; Heaulme, M.; Neliat, G.; Santucci, V.; Steinberg, R.; Vilain, P.; Bichon, D.; Ducoux, J.-P.; Proietto, V.; et al. In vitro and in vivo biological activities of SR140333, a novel potent non-peptide tachykinin NK1 receptor antagonist. *Eur. J. Pharmacol.* **1993**, *250*, 403–413.
134. Ofner, S.; Hauser, K.; Schilling, W.; Vassout, A.; Veenstra, S.J. SAR of 2-benzyl-4-aminopiperidines: CGP 49823, an orally and centrally active non-peptide NK1 antagonist. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1623–1628.
135. Duffy, R.A.; Morgan, C.; Naylor, R.; Higgins, G.A.; Varty, G.B.; Lachowicz, J.E.; Parker, E.M. Rolapitant (SCH 619734): A potent, selective and orally active neurokinin NK1 receptor antagonist with centrally-mediated antiemetic effects in ferrets. *Pharmacol. Biochem. Behav.* **2012**, *102*, 95–100.
136. Syed, Y.Y. Rolapitant: First Global Approval. *Drugs* **2015**, *75*, 1941–1945.
137. Sakurai, H.; Yonezawa, K.; Tani, H.; Mimura, M.; Bauer, M.; Uchida, H. Novel Antidepressants in the Pipeline (Phase II and III): A Systematic Review of the US Clinical Trials Registry. *Pharmacopsychiatry* **2022**, *55*, 193–202.
138. Ratti, E.; Bettica, P.; Alexander, R.; Archer, G.; Carpenter, D.; Evoniuk, G.; Gomeni, R.; Lawson, E.; Lopez, M.; Millns, H.; et al. Full central neurokinin-1 receptor blockade is required for efficacy in depression: evidence from orvepitant clinical studies. *J. Psychopharmacol.* **2013**, *27*, 424–434.
139. MacLeod, A.M.; Merchant, K.J.; Cascieri, M.A.; Sadowski, S.; Ber, E.; Swain, C.J.; Baker, R. N-Acyl-L-tryptophan benzyl esters: potent substance P receptor antagonists. *J. Med. Chem.* **1993**, *36*, 2044–2045.
140. MacLeod, A.M.; Merchant, K.J.; Brookfield, F.; Kelleher, F.; Stevenson, G.; Owens, A.P.; Swain, C.J.; Cascieri, M.A.; Sadowski, S. Identification of L-Tryptophan Derivatives with Potent and Selective Antagonist Activity at the NK1 Receptor. *J. Med. Chem.* **1994**, *37*, 1269–1274.
141. MacLeod, A.M.; Cascieri, M.A.; Merchant, K.J.; Sadowski, S.; Hardwicke, S.; Lewis, R.T.; MacIntyre, D.E.; Metzger, J.M.; Fong, T.M. Synthesis and Biological Evaluation of NK1 Antagonists Derived from L-Tryptophan. *J. Med. Chem.* **1995**, *38*, 934–941.
142. Hipskind, P.A.; Howbert, J.J.; Bruns, R.F.; Cho, S.S.Y.; Crowell, T.A.; Foreman, M.M.; Gehlert, D.R.;

- Iyengar, S.; Johnson, K.W.; Krushinski, J.H.; et al. 3-Aryl-1,2-diacetamidopropane Derivatives as Novel and Potent NK-1 Receptor Antagonists. *J. Med. Chem.* **1996**, *39*, 736–748.
143. Goldstein, D. Study of the analgesic effect of lanepitant in patients with osteoarthritis pain. *Clin. Pharmacol. Ther.* **2000**, *67*, 419–426.
144. Xiao, J.; Zeng, S.; Wang, X.; Babazada, H.; Li, Z.; Liu, R.; Yu, W. Neurokinin 1 and opioid receptors: relationships and interactions in nervous system. *Transl. Perioper. pain Med.* **2016**, *1*, 11–21.
145. Lipkowski, A.W.; Misterek, K. Bifunctional pharmacophores. Biological activities of the peptide analog containing both casomorphin-like and substance P antagonist-like active elements. *Pol. J. Pharmacol. Pharm.* **1992**, *44*, 25–32.
146. LIPKOWSKI, A.W.; CARR, D.B.; BONNEY, I.; KOSSON, P. Opioid-Substance P Chimeric Peptides. In *Handbook of Biologically Active Peptides*; Elsevier, 2006; pp. 1373–1378.
147. Bonney, I.M.; Foran, S.E.; Marchand, J.E.; Lipkowski, A.W.; Carr, D.B. Spinal antinociceptive effects of AA501, a novel chimeric peptide with opioid receptor agonist and tachykinin receptor antagonist moieties. *Eur. J. Pharmacol.* **2004**, *488*, 91–99.
148. Matalińska, J.; Lipiński, P.F.J.; Kosson, P.; Kosińska, K.; Misicka, A. In Vivo, In Vitro and In Silico Studies of the Hybrid Compound AA3266, an Opioid Agonist/NK1R Antagonist with Selective Cytotoxicity. *Int. J. Mol. Sci.* **2020**, *21*, 7738.
149. Matalinska, J.; Skurzak, H.; Markowicz, S.; Lesniak, A.; Sacharczuk, M.; Molnar, G.; Varga, E.; Lipkowski, A.W. Opioid agonist – tachykinin antagonist as a new analgesic with adjuvant anticancer properties. *Folia Neuropathol.* **2013**, *2*, 132–139.
150. Dyniewicz, J.; Lipiński, P.F.J.; Kosson, P.; Leśniak, A.; Bochyńska-Czyż, M.; Muchowska, A.; Tourwé, D.; Ballet, S.; Misicka, A.; Lipkowski, A.W. Hydrazone Linker as a Useful Tool for Preparing Chimeric Peptide/Nonpeptide Bifunctional Compounds. *ACS Med. Chem. Lett.* **2017**, *8*, 73–77.
151. Dyniewicz, J.; Lipiński, P.F.J.; Kosson, P.; Bochyńska-Czyż, M.; Matalińska, J.; Misicka, A. Antinociceptive and Cytotoxic Activity of Opioid Peptides with Hydrazone and Hydrazide Moieties at the C-Terminus. *Molecules* **2020**, *25*, 3429.
152. Wtorek, K.; Adamska-Bartłomiejczyk, A.; Piekielna-Ciesielska, J.; Ferrari, F.; Ruzza, C.; Kluczyk, A.; Piasecka-Zelga, J.; Calo', G.; Janecka, A. Synthesis and Pharmacological Evaluation of Hybrids Targeting Opioid and Neurokinin Receptors. *Molecules* **2019**, *24*, 4460.
153. Kowalczyk, A.; Kleczkowska, P.; Rękawek, M.; Kulik, K.; Lesniak, A.; Erdei, A.; Borics, A.; Martin, C.; Pawlik, K.; Lipkowski, A.W.; et al. Biological evaluation and molecular docking studies of AA3052, a compound containing a μ -selective opioid peptide agonist DALDA and d-Phe-Phe-d-Phe-Leu-Leu-NH₂, a substance P analogue. *Eur. J. Pharm. Sci.* **2016**, *93*, 11–20.
154. Yamamoto, T.; Nair, P.; Davis, P.; Ma, S.; Navratilova, E.; Moye, S.; Tumati, S.; Lai, J.; Vanderah, T.W.; Yamamura, H.I.; et al. Design, Synthesis, and Biological Evaluation of Novel Bifunctional C-Terminal-Modified Peptides for δ/μ Opioid Receptor Agonists and Neurokinin-1 Receptor Antagonists. *J. Med. Chem.* **2007**, *50*, 2779–2786.
155. Largent-Milnes, T.M.T.; Yamamoto, T.; Nair, P.; Moulton, J.W.; Hruby, V.J.V.; Lai, J.; Porreca, F.; Vanderah, T.W. Spinal or systemic TY005, a peptidic opioid agonist/neurokinin 1 antagonist, attenuates pain with reduced tolerance. *Br. J. Pharmacol.* **2010**, *161*, 986–1001.

156. Vardanyan, R.; Kumirov, V.K.; Nichol, G.S.; Davis, P.; Liktor-Busa, E.; Rankin, D.; Varga, E.; Vanderah, T.; Porreca, F.; Lai, J.; et al. Synthesis and biological evaluation of new opioid agonist and neurokinin-1 antagonist bivalent ligands. *Bioorganic Med. Chem.* **2011**, *19*, 6135–6142.
157. Ballet, S.; Feytens, D.; Buysse, K.; Chung, N.N.; Lemieux, C.; Tumati, S.; Keresztes, A.; Van Duppen, J.; Lai, J.; Varga, E.; et al. Design of Novel Neurokinin 1 Receptor Antagonists Based on Conformationally Constrained Aromatic Amino Acids and Discovery of a Potent Chimeric Opioid Agonist-Neurokinin 1 Receptor Antagonist. *J. Med. Chem.* **2011**, *54*, 2467–2476.
158. Guillemyn, K.; Kleczkowska, P.; Novoa, A.; Vandormael, B.; Van den Eynde, I.; Kosson, P.; Asim, M.F.; Schiller, P.W.; Spetea, M.; Lipkowski, A.W.; et al. In vivo antinociception of potent mu opioid agonist tetrapeptide analogues and comparison with a compact opioid agonist-neurokinin 1 receptor antagonist chimera. *Mol. Brain* **2012**, *5*, 4.
159. Gadais, C.; Piekielna-Ciesielska, J.; De Neve, J.; Martin, C.; Janecka, A.; Ballet, S. Harnessing the Anti-Nociceptive Potential of NK2 and NK3 Ligands in the Design of New Multifunctional μ/δ -Opioid Agonist-Neurokinin Antagonist Peptidomimetics. *Molecules* **2021**, *26*, 5406.
160. Thomsen, A.R.B.; Jensen, D.D.; Hicks, G.A.; Bunnett, N.W. Therapeutic Targeting of Endosomal G-Protein-Coupled Receptors. *Trends Pharmacol. Sci.* **2018**, *39*, 879–891.
161. Mai, Q.N.; Shenoy, P.; Quach, T.; Retamal, J.S.; Gondin, A.B.; Yeatman, H.R.; Aurelio, L.; Conner, J.W.; Poole, D.P.; Canals, M.; et al. A lipid-anchored neurokinin 1 receptor antagonist prolongs pain relief by a three-pronged mechanism of action targeting the receptor at the plasma membrane and in endosomes. *J. Biol. Chem.* **2021**, *296*, 100345.
162. Ramírez-García, P.D.; Retamal, J.S.; Shenoy, P.; Imlach, W.; Sykes, M.; Truong, N.; Constandil, L.; Pelissier, T.; Nowell, C.J.; Khor, S.Y.; et al. A pH-responsive nanoparticle targets the neurokinin 1 receptor in endosomes to prevent chronic pain. *Nat. Nanotechnol.* **2019**, *14*, 1150–1159.
163. Latorre, R.; Ramírez-García, P.D.; Hegron, A.; Grace, J.L.; Retamal, J.S.; Shenoy, P.; Tran, M.; Aurelio, L.; Flynn, B.; Poole, D.P.; et al. Sustained endosomal release of a neurokinin-1 receptor antagonist from nanostars provides long-lasting relief of chronic pain. *Biomaterials* **2022**, *285*, 121536.
164. Hegron, A.; Peach, C.J.; Tonello, R.; Seemann, P.; Teng, S.; Latorre, R.; Huebner, H.; Weikert, D.; Rientjes, J.; Veldhuis, N.A.; et al. Therapeutic antagonism of the neurokinin 1 receptor in endosomes provides sustained pain relief. *Proc. Natl. Acad. Sci.* **2023**, *120*.
165. Muñoz, M.; Rosso, M.; Covenas, R. The NK-1 Receptor: A New Target in Cancer Therapy. *Curr. Drug Targets* **2011**, *12*, 909–921.
166. Muñoz, M.; Rosso, M.; González-Ortega, A.; Coveñas, R. The NK-1 Receptor Antagonist L-732,138 Induces Apoptosis and Counteracts Substance P-Related Mitogenesis in Human Melanoma Cell Lines. *Cancers (Basel)*. **2010**, *2*, 611–623.
167. Majkowska-Pilip, A.; Halik, P.K.; Gniazdowska, E. The Significance of NK1 Receptor Ligands and Their Application in Targeted Radionuclide Tumour Therapy. *Pharmaceutics* **2019**, *11*, 443.
168. Berger, M.; Neth, O.; Ilmer, M.; Garnier, A.; Salinas-Martín, M.V.; de Agustín Asencio, J.C.; von Schweinitz, D.; Kappler, R.; Muñoz, M. Hepatoblastoma cells express truncated neurokinin-1 receptor and can be growth inhibited by aprepitant in vitro and in vivo. *J. Hepatol.* **2014**, *60*, 985–994.
169. Muñoz, M.; Rosso, M.; Coveñas, R. The NK-1 receptor antagonist L-732,138 induces apoptosis in

- human gastrointestinal cancer cell lines. *Pharmacol. Reports* **2017**, *69*, 696–701.
170. Muñoz, M.; Berger, M.; Rosso, M.; Gonzales-Ortega, A.; Carranza, A.; Coveñas, R. Antitumor activity of neurokinin-1 receptor antagonists in MG-63 human osteosarcoma xenografts. *Int. J. Oncol.* **2014**, *44*, 137–146.
171. Muñoz, M.; Pérez, A.; Rosso, M.; Zamarriego, C.; Rosso, R. Antitumoral action of the neurokinin-1 receptor antagonist L-733 060 on human melanoma cell lines. *Melanoma Res.* **2004**, *14*, 183–188.
172. Muñoz, M.; Crespo, J.; Crespo, J.; Coveñas, R. Neurokinin-1 receptor antagonist aprepitant and radiotherapy, a successful combination therapy in a patient with lung cancer: A case report. *Mol. Clin. Oncol.* **2019**.
173. Lee, M. Prolonged Use of Aprepitant in Metastatic Breast Cancer and a Reduction in CA153 Tumour Marker Levels. *Int. J. Cancer Clin. Res.* **2016**, *3*.
174. Efficacy and Safety of PVT-1 Treatment in Patients With Advanced Non-Small Cell Lung Cancer <https://clinicaltrials.gov/study/NCT04840004>.
175. Recio, R.; Vengut-Climent, E.; Mouillac, B.; Orcel, H.; López-Lázaro, M.; Calderón-Montaño, J.M.; Álvarez, E.; Khair, N.; Fernández, I. Design, synthesis and biological studies of a library of NK1-Receptor Ligands Based on a 5-arylthiosubstituted 2-amino-4,6-diaryl-3-cyano-4 H -pyran core: Switch from antagonist to agonist effect by chemical modification. *Eur. J. Med. Chem.* **2017**, *138*, 644–660.
176. Cappelli, A.; Giuliani, G.; Pericot Mohr, G. I.; Gallelli, A.; Anzini, M.; Vomero, S.; Cupello, A.; Scarrone, S.; Matarrese, M.; Moresco, R.M.; et al. A Non-Peptide NK 1 Receptor Agonist Showing Subpicomolar Affinity. *J. Med. Chem.* **2004**, *47*, 1315–1318.
177. Recio, R.; Lerena, P.; Pozo, E.; Calderón-Montaño, J.M.; Burgos-Morón, E.; López-Lázaro, M.; Valdivia, V.; Pernia Leal, M.; Mouillac, B.; Organero, J.Á.; et al. Carbohydrate-Based NK1R Antagonists with Broad-Spectrum Anticancer Activity. *J. Med. Chem.* **2021**, *64*, 10350–10370.
178. Valdivia, V.; Recio, R.; Lerena, P.; Pozo, E.; Serrano, R.; Calero, R.; Pintado, C.; Leal, M.P.; Moreno-Rodríguez, N.; Organero, J.Á.; et al. Biological evaluation of carbohydrate-based aprepitant analogs for neuroblastoma treatment. *Eur. J. Med. Chem.* **2024**, *264*, 116021.
179. van Hagen, P.M.; Breeman, W.A.P.; Reubi, J.C.; Postema, P.T.E.; van den Anker-Lugtenburg, P.J.; Kwekkeboom, D.J.; Laissue, J.; Waser, B.; Lamberts, S.W.J.; Visser, T.J.; et al. Visualization of the thymus by substance P receptor scintigraphy in man. *Eur. J. Nucl. Med.* **1996**, *23*, 1508–1513.
180. Kneifel, S.; Cordier, D.; Good, S.; Ionescu, M.C.S.; Ghaffari, A.; Hofer, S.; Kretzschmar, M.; Tolnay, M.; Apostolidis, C.; Waser, B.; et al. Local Targeting of Malignant Gliomas by the Diffusible Peptidic Vector 1,4,7,10-Tetraazacyclododecane-1-Glutaric Acid-4,7,10-Triacetic Acid-Substance P. *Clin. Cancer Res.* **2006**, *12*, 3843–3850.
181. Królicki, L.; Bruchertseifer, F.; Kunikowska, J.; Koziara, H.; Królicki, B.; Jakuciński, M.; Pawlak, D.; Apostolidis, C.; Mirzadeh, S.; Rola, R.; et al. Safety and efficacy of targeted alpha therapy with 213Bi-DOTA-substance P in recurrent glioblastoma. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 614–622.
182. Królicki, L.; Kunikowska, J.; Bruchertseifer, F.; Koziara, H.; Królicki, B.; Jakuciński, M.; Pawlak, D.; Rola, R.; Morgenstern, A.; Rosiak, E.; et al. 225Ac- and 213Bi-Substance P Analogues for Glioma Therapy. *Semin. Nucl. Med.* **2020**, *50*, 141–151.
183. Cordier, D.; Merlo, A. Long-Term Results of Targeted Low-Grade Glioma Treatment with 213 Bi-

- DOTA-[Thi 8 ,Met(O 2) 11]-Substance P. *Cancer Biother. Radiopharm.* **2019**, *34*, 413–416.
184. Majkowska-Pilip, A.; Rius, M.; Bruchertseifer, F.; Apostolidis, C.; Weis, M.; Bonelli, M.; Laurenza, M.; Królicki, L.; Morgenstern, A. In vitro evaluation of ²²⁵Ac-DOTA-substance P for targeted alpha therapy of glioblastoma multiforme. *Chem. Biol. Drug Des.* **2018**, *92*, 1344–1356.
185. Zhang, H.; Kanduluru, A.K.; Desai, P.; Ahad, A.; Carlin, S.; Tandon, N.; Weber, W.A.; Low, P.S. Synthesis and Evaluation of a Novel ⁶⁴Cu- and ⁶⁷Ga-Labeled Neurokinin 1 Receptor Antagonist for in Vivo Targeting of NK1R-Positive Tumor Xenografts. *Bioconjug. Chem.* **2018**, *29*, 1319–1326.
186. Kanduluru, A.K.; Srinivasarao, M.; Wayua, C.; Low, P.S. Evaluation of a Neurokinin-1 Receptor-Targeted Technetium-99m Conjugate for Neuroendocrine Cancer Imaging. *Mol. Imaging Biol.* **2020**, *22*, 377–383.
187. Halik, P.K.; Koźmiński, P.; Matalińska, J.; Lipiński, P.F.J.; Misicka, A.; Gniazdowska, E. In Vitro Biological Evaluation of Aprepitant Based ¹⁷⁷Lu-Radioconjugates. *Pharmaceutics* **2022**, *14*, 607.
188. Halik, P.K.; Lipiński, P.F.J.; Matalińska, J.; Koźmiński, P.; Misicka, A.; Gniazdowska, E. Radiochemical Synthesis and Evaluation of Novel Radioconjugates of Neurokinin 1 Receptor Antagonist Aprepitant Dedicated for NK1R-Positive Tumors. *Molecules* **2020**, *25*, 3756.
189. Michelgård, Å.; Appel, L.; Pissioti, A.; Frans, Ö.; Långström, B.; Bergström, M.; Fredrikson, M. Symptom Provocation in Specific Phobia Affects the Substance P Neurokinin-1 Receptor System. *Biol. Psychiatry* **2007**, *61*, 1002–1006.
190. Frick, A.; Ahs, F.; Linnman, C.; Jonasson, M.; Appel, L.; Lubberink, M.; Långström, B.; Fredrikson, M.; Furmark, T. Increased neurokinin-1 receptor availability in the amygdala in social anxiety disorder: a positron emission tomography study with [¹¹C]GR205171. *Transl. Psychiatry* **2015**, *5*, e597–e597.
191. Danfors, T.; Åhs, F.; Appel, L.; Linnman, C.; Fredrikson, M.; Furmark, T.; Kumlien, E. Increased neurokinin-1 receptor availability in temporal lobe epilepsy: A positron emission tomography study using [¹¹C]GR205171. *Epilepsy Res.* **2011**, *97*, 183–189.
192. Fujimura, Y.; Yasuno, F.; Farris, A.; Liow, J.S.; Geraci, M.; Drevets, W.; Pine, D.S.; Ghose, S.; Lerner, A.; Hargreaves, R.; et al. Decreased Neurokinin-1 (Substance P) Receptor Binding in Patients with Panic Disorder: Positron Emission Tomographic Study with [¹⁸F]SPA-RQ. *Biol. Psychiatry* **2009**, *66*, 94–97.
193. Zamuner, S.; Rabiner, E.A.; Fernandes, S.A.; Bani, M.; Gunn, R.N.; Gomeni, R.; Ratti, E.; Cunningham, V.J. A pharmacokinetic PET study of NK1 receptor occupancy. *Eur. J. Nucl. Med. Mol. Imaging* **2012**, *39*, 226–235.
194. Poma, A.; Christensen, J.; Davis, J.; Kansra, V.; Martell, R.E.; Hedley, M.L. Phase 1 positron emission tomography (PET) study of the receptor occupancy of rolapitant, a novel NK-1 receptor antagonist. *J. Clin. Oncol.* **2014**, *32*, e20690–e20690.
195. Rizk, S.S.; Luchniak, A.; Uysal, S.; Brawley, C.M.; Rock, R.S.; Kossiakoff, A.A. An engineered substance P variant for receptor-mediated delivery of synthetic antibodies into tumor cells. *Proc. Natl. Acad. Sci.* **2009**, *106*, 11011–11015.
196. Rizk, S.S.; Misiura, A.; Paduch, M.; Kossiakoff, A.A. Substance P Derivatives as Versatile Tools for Specific Delivery of Various Types of Biomolecular Cargo. *Bioconjug. Chem.* **2012**, *23*, 42–46.
197. Zhang, W.; Song, J.; Mu, L.; Zhang, B.; Liu, L.; Xing, Y.; Wang, K.; Li, Z.; Wang, R. Improving

- anticancer activity and selectivity of camptothecin through conjugation with releasable substance P. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1452–1455.
198. Song, J.; Huang, S.; Zhang, Z.; Jia, B.; Xie, H.; Kai, M.; Zhang, W. SPA: a peptide antagonist that acts as a cell-penetrating peptide for drug delivery. *Drug Deliv.* **2020**, *27*, 91–99.
199. Song, J.; Huang, S.; Ma, P.; Zhang, B.; Jia, B.; Zhang, W. Improving NK1R-targeted gene delivery of stearyl-antimicrobial peptide CAMEL by conjugating it with substance P. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 127353.
200. Ding, G.; Wang, T.; Han, Z.; Tian, L.; Cheng, Q.; Luo, L.; Zhao, B.; Wang, C.; Feng, S.; Wang, L.; et al. Substance P containing peptide gene delivery vectors for specifically transfecting glioma cells mediated by a neurokinin-1 receptor. *J. Mater. Chem. B* **2021**, *9*, 6347–6356.
201. Kitchen, D.B.; Decornez, H.; Furr, J.R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, *3*, 935–949.
202. Shoichet, B.K. Virtual screening of chemical libraries. *Nature* **2004**, *432*, 862–865.
203. Schneider, G.; Fechner, U. Computer-based de novo design of drug-like molecules. *Nat. Rev. Drug Discov.* **2005**, *4*, 649–663.
204. Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided. Drug Des.* **2011**, *7*, 146–57.
205. Goodsell, D.S.; Olson, A.J. Automated docking of substrates to proteins by simulated annealing. *Proteins Struct. Funct. Bioinforma.* **1990**, *8*, 195–202.
206. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
207. Moustakas, D.T.; Lang, P.T.; Pegg, S.; Pettersen, E.; Kuntz, I.D.; Brooijmans, N.; Rizzo, R.C. Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided. Mol. Des.* **2006**, *20*, 601–619.
208. Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins Struct. Funct. Genet.* **1999**, *37*, 228–241.
209. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D. Improved protein–ligand docking using GOLD. *Proteins Struct. Funct. Bioinforma.* **2003**, *52*, 609–623.
210. Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
211. Berman, H.M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
212. Stanzione, F.; Giangreco, I.; Cole, J.C. Use of molecular docking computational tools in drug discovery. In *Progress in Medicinal Chemistry*; 2021; pp. 273–343.
213. Guedes, I.A.; de Magalhães, C.S.; Dardenne, L.E. Receptor–ligand molecular docking. *Biophys. Rev.* **2014**, *6*, 75–87.
214. Brooijmans, N.; Kuntz, I.D. Molecular Recognition and Docking Algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335–373.
215. Huang, S.-Y.; Grinter, S.Z.; Zou, X. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12899.

216. Liu, J.; Wang, R. Classification of Current Scoring Functions. *J. Chem. Inf. Model.* **2015**, *55*, 475–482.
217. Meng, E.C.; Shoichet, B.K.; Kuntz, I.D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **1992**, *13*, 505–524.
218. Quiroga, R.; Villarreal, M.A. Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. *PLoS One* **2016**, *11*, e0155183.
219. Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein–Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *47*, 2977–2980.
220. Neudert, G.; Klebe, G. DSX: a knowledge-based scoring function for the assessment of protein-ligand complexes. *J. Chem. Inf. Model.* **2011**, *51*, 2731–45.
221. Muegge, I. PMF scoring revisited. *J. Med. Chem.* **2006**, *49*, 5895–5902.
222. Fischer, A.; Smieško, M.; Sellner, M.; Lill, M.A. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. *J. Med. Chem.* **2021**, *64*, 2489–2500.
223. Fu, D.Y.; Meiler, J. Predictive Power of Different Types of Experimental Restraints in Small Molecule Docking: A Review. *J. Chem. Inf. Model.* **2018**, *58*, 225–233.
224. Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* **2018**, *9*.
225. Kooistra, A.J.; Vischer, H.F.; McNaught-Flores, D.; Leurs, R.; de Esch, I.J.P.; de Graaf, C. Function-specific virtual screening for GPCR ligands using a combined scoring method. *Sci. Rep.* **2016**, *6*, 28288.
226. Podlewska, S.; Bojarski, A.J. Post-processing of Docking Results: Tools and Strategies. In *Molecular Docking for Computer-Aided Drug Design*; Elsevier, 2021; pp. 57–74.
227. Li, H.; Leung, K.; Wong, M.; Ballester, P.J. Improving AutoDock Vina Using Random Forest: The Growing Accuracy of Binding Affinity Prediction by the Effective Exploitation of Larger Data Sets. *Mol. Inform.* **2015**, *34*, 115–126.
228. Li, H.; Sze, K.; Lu, G.; Ballester, P.J. Machine-learning scoring functions for structure-based drug lead optimization. *WIREs Comput. Mol. Sci.* **2020**, *10*.
229. Li, H.; Sze, K.; Lu, G.; Ballester, P.J. Machine-learning scoring functions for structure-based virtual screening. *WIREs Comput. Mol. Sci.* **2021**, *11*.
230. Gabel, J.; Desaphy, J.; Rognan, D. Beware of Machine Learning-Based Scoring Functions—On the Danger of Developing Black Boxes. *J. Chem. Inf. Model.* **2014**, *54*, 2807–2815.
231. Sieg, J.; Flachsenberg, F.; Rarey, M. In Need of Bias Control: Evaluating Chemical Data for Machine Learning in Structure-Based Virtual Screening. *J. Chem. Inf. Model.* **2019**, *59*, 947–961.
232. Tran-Nguyen, V.-K.; Bret, G.; Rognan, D. True Accuracy of Fast Scoring Functions to Predict High-Throughput Screening Data from Docking Poses: The Simpler the Better. *J. Chem. Inf. Model.* **2021**, *61*, 2788–2797.
233. Chen, L.; Cruz, A.; Ramsey, S.; Dickson, C.J.; Duca, J.S.; Hornak, V.; Koes, D.R.; Kurtzman, T. Hidden bias in the DUD-E dataset leads to misleading performance of deep learning in structure-based virtual screening. *PLoS One* **2019**, *14*, e0220113.
234. Volkov, M.; Turk, J.-A.; Drizard, N.; Martin, N.; Hoffmann, B.; Gaston-Mathé, Y.; Rognan, D. On the Frustration to Predict Binding Affinities from Protein–Ligand Structures with Deep Neural Networks. *J.*

- Med. Chem.* **2022**, *65*, 7946–7958.
235. Yang, J.; Shen, C.; Huang, N. Predicting or Pretending: Artificial Intelligence for Protein-Ligand Interactions Lack of Sufficiently Large and Unbiased Datasets. *Front. Pharmacol.* **2020**, *11*.
236. Shen, C.; Hu, Y.; Wang, Z.; Zhang, X.; Pang, J.; Wang, G.; Zhong, H.; Xu, L.; Cao, D.; Hou, T. Beware of the generic machine learning-based scoring functions in structure-based virtual screening. *Brief. Bioinform.* **2021**, *22*.
237. Ghislat, G.; Rahman, T.; Ballester, P.J. Recent progress on the prospective application of machine learning to structure-based virtual screening. *Curr. Opin. Chem. Biol.* **2021**, *65*, 28–34.
238. Warren, G.L.; Andrews, C.W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M.H.; Lindvall, M.; Nevins, N.; Semus, S.F.; Senger, S.; et al. A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
239. Kramer, C.; Gedeck, P. Leave-Cluster-Out Cross-Validation Is Appropriate for Scoring Functions Derived from Diverse Protein Data Sets. *J. Chem. Inf. Model.* **2010**, *50*, 1961–1969.
240. Plewczynski, D.; Łaźniewski, M.; Augustyniak, R.; Ginalski, K. Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database. *J. Comput. Chem.* **2011**, *32*, 742–755.
241. Smith, R.D.; Dunbar, J.B.; Ung, P.M.-U.; Esposito, E.X.; Yang, C.-Y.; Wang, S.; Carlson, H.A. CSAR Benchmark Exercise of 2010: Combined Evaluation Across All Submitted Scoring Functions. *J. Chem. Inf. Model.* **2011**, *51*, 2115–2131.
242. Ross, G.A.; Morris, G.M.; Biggin, P.C. One Size Does Not Fit All: The Limits of Structure-Based Models in Drug Discovery. *J. Chem. Theory Comput.* **2013**, *9*, 4266–4274.
243. Seifert, M.H.J. Targeted scoring functions for virtual screening. *Drug Discov. Today* **2009**, *14*, 562–569.
244. Martin, E.J.; Sullivan, D.C. AutoShim: Empirically Corrected Scoring Functions for Quantitative Docking with a Crystal Structure and IC 50 Training Data. *J. Chem. Inf. Model.* **2008**, *48*, 861–872.
245. Santos-Martins, D.; Forli, S.; Ramos, M.J.; Olson, A.J. AutoDock4 Zn : An Improved AutoDock Force Field for Small-Molecule Docking to Zinc Metalloproteins. *J. Chem. Inf. Model.* **2014**, *54*, 2371–2379.
246. Logean, A.; Sette, A.; Rognan, D. Customized versus universal scoring functions. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 675–679.
247. Pham, T.A.; Jain, A.N. Customizing scoring functions for docking. *J. Comput. Aided. Mol. Des.* **2008**, *22*, 269–286.
248. Ravitz, O.; Zsoldos, Z.; Simon, A. Improving molecular docking through eHiTS' tunable scoring function. *J. Comput. Aided. Mol. Des.* **2011**, *25*, 1033–1051.
249. Kinnings, S.L.; Liu, N.; Tonge, P.J.; Jackson, R.M.; Xie, L.; Bourne, P.E. A Machine Learning-Based Method To Improve Docking Scoring Functions and Its Application to Drug Repurposing. *J. Chem. Inf. Model.* **2011**, *51*, 408–419.
250. Antes, I.; Merkwirth, C.; Lengauer, T. POEM: Parameter Optimization Using Ensemble Methods: Application to Target Specific Scoring Functions. *J. Chem. Inf. Model.* **2005**, *45*, 1291–1302.
251. Seifert, M.H.J. Robust optimization of scoring functions for a target class. *J. Comput. Aided. Mol. Des.* **2009**, *23*, 633–644.
252. Seifert, M.H.J. Optimizing the Signal-to-Noise Ratio of Scoring Functions for Protein–Ligand Docking.

- J. Chem. Inf. Model.* **2008**, *48*, 602–612.
253. Wang, W.; Huang, Q.; Zou, J.; Li, L.; Yang, S. <sc>TS</sc> -Chemscore, a Target-Specific Scoring Function, Significantly Improves the Performance of Scoring in Virtual Screening. *Chem. Biol. Drug Des.* **2015**, *86*, 1–8.
254. Shamsara, E.; Shamsara, J. Developing target-specific scoring using black-box optimisation. *Int. J. Comput. Biol. Drug Des.* **2017**, *10*, 12.
255. Mysinger, M.M.; Carchia, M.; Irwin, J.J.; Shoichet, B.K. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* **2012**, *55*, 6582–94.
256. Rzęsikowska, K.; Kalinowska-Tłuścik, J.; Krawczuk, A. Hierarchical analysis of the target-based scoring function modification for the example of selected class A GPCRs. *Phys. Chem. Chem. Phys.* **2023**, *25*, 3513–3520.
257. Mendez, D.; Gaulton, A.; Bento, A.P.; Chambers, J.; De Veij, M.; Félix, E.; Magariños, M.P.; Mosquera, J.F.; Mutowo, P.; Nowotka, M.; et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* **2019**, *47*, D930–D940.
258. Davies, M.; Nowotka, M.; Papadatos, G.; Dedman, N.; Gaulton, A.; Atkinson, F.; Bellis, L.; Overington, J.P. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* **2015**, *43*, W612–W620.
259. Spiegel, J.; Senderowitz, H. Towards an Enrichment Optimization Algorithm (EOA)-based Target Specific Docking Functions for Virtual Screening. *Mol. Inform.* **2022**, *41*.
260. Buzko, O. V; Bishop, A.C.; Shokat, K.M. Modified AutoDock for accurate docking of protein kinase inhibitors. *J. Comput. Aided. Mol. Des.* **2002**, *16*, 113–27.
261. Xue, M.; Zheng, M.; Xiong, B.; Li, Y.; Jiang, H.; Shen, J. Knowledge-Based Scoring Functions in Drug Design. 1. Developing a Target-Specific Method for Kinase–Ligand Interactions. *J. Chem. Inf. Model.* **2010**, *50*, 1378–1386.
262. Wang, Y.; Li, L.; Zhang, B.; Xing, J.; Chen, S.; Wan, W.; Song, Y.; Jiang, H.; Jiang, H.; Luo, C.; et al. Discovery of Novel Disruptor of Silencing Telomeric 1-Like (DOT1L) Inhibitors using a Target-Specific Scoring Function for the (S)-Adenosyl- 1 -methionine (SAM)-Dependent Methyltransferase Family. *J. Med. Chem.* **2017**, *60*, 2026–2036.
263. Shen, Q.; Xiong, B.; Zheng, M.; Luo, X.; Luo, C.; Liu, X.; Du, Y.; Li, J.; Zhu, W.; Shen, J.; et al. Knowledge-Based Scoring Functions in Drug Design: 2. Can the Knowledge Base Be Enriched? *J. Chem. Inf. Model.* **2011**, *51*, 386–397.
264. Pfeffer, P.; Gohlke, H. DrugScore RNA Knowledge-Based Scoring Function To Predict RNA–Ligand Interactions. *J. Chem. Inf. Model.* **2007**, *47*, 1868–1876.
265. Stefaniak, F.; Bujnicki, J.M. AnnapuRNA: A scoring function for predicting RNA-small molecule binding poses. *PLOS Comput. Biol.* **2021**, *17*, e1008309.
266. Yan, Z.; Wang, J. SPA-LN: a scoring function of ligand–nucleic acid interactions via optimizing both specificity and affinity. *Nucleic Acids Res.* **2017**, *45*, e110–e110.
267. Chen, L.; Calin, G.A.; Zhang, S. Novel Insights of Structure-Based Modeling for RNA-Targeted Drug Discovery. *J. Chem. Inf. Model.* **2012**, *52*, 2741–2753.
268. Santos-Martins, D. Interaction with specific HSP90 residues as a scoring function: validation in the D3R

- Grand Challenge 2015. *J. Comput. Aided. Mol. Des.* **2016**, *30*, 731–742.
269. Yan, Y.; Wang, W.; Sun, Z.; Zhang, J.Z.H.; Ji, C. Protein–Ligand Empirical Interaction Components for Virtual Screening. *J. Chem. Inf. Model.* **2017**, *57*, 1793–1806.
270. Fourches, D.; Politi, R.; Tropsha, A. Target-Specific Native/Decoy Pose Classifier Improves the Accuracy of Ligand Ranking in the CSAR 2013 Benchmark. *J. Chem. Inf. Model.* **2015**, *55*, 63–71.
271. Berishvili, V.P.; Voronkov, A.E.; Radchenko, E. V.; Palyulin, V.A. Machine Learning Classification Models to Improve the Docking-based Screening: A Case of PI3K-Tankyrase Inhibitors. *Mol. Inform.* **2018**, *37*.
272. Koes, D.R.; Baumgartner, M.P.; Camacho, C.J. Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *J. Chem. Inf. Model.* **2013**, *53*, 1893–1904.
273. Liu, Z.; Li, Y.; Han, L.; Li, J.; Liu, J.; Zhao, Z.; Nie, W.; Liu, Y.; Wang, R. PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics* **2015**, *31*, 405–412.
274. Wójcikowski, M.; Ballester, P.J.; Siedlecki, P. Performance of machine-learning scoring functions in structure-based virtual screening. *Sci. Rep.* **2017**, *7*, 46710.
275. Pellicani, F.; Dal Ben, D.; Perali, A.; Pilati, S. Machine Learning Scoring Functions for Drug Discovery from Experimental and Computer-Generated Protein–Ligand Structures: Towards Per-Target Scoring Functions. *Molecules* **2023**, *28*, 1661.
276. Wang, D.; Cui, C.; Ding, X.; Xiong, Z.; Zheng, M.; Luo, X.; Jiang, H.; Chen, K. Improving the Virtual Screening Ability of Target-Specific Scoring Functions Using Deep Learning Methods. *Front. Pharmacol.* **2019**, *10*, 1–11.
277. Shen, C.; Weng, G.; Zhang, X.; Leung, E.L.-H.; Yao, X.; Pang, J.; Chai, X.; Li, D.; Wang, E.; Cao, D.; et al. Accuracy or novelty: what can we gain from target-specific machine-learning-based scoring functions in virtual screening? *Brief. Bioinform.* **2021**, *22*.
278. Kerzmann, A.; Neumann, D.; Kohlbacher, O. SLICK – Scoring and Energy Functions for Protein–Carbohydrate Interactions. *J. Chem. Inf. Model.* **2006**, *46*, 1635–1642.
279. Kerzmann, A.; Fuhrmann, J.; Kohlbacher, O.; Neumann, D. BALLDock/SLICK: A New Method for Protein–Carbohydrate Docking. *J. Chem. Inf. Model.* **2008**, *48*, 1616–1625.
280. Nivedha, A.K.; Thieker, D.F.; Makeneni, S.; Hu, H.; Woods, R.J. Vina-Carb: Improving Glycosidic Angles during Carbohydrate Docking. *J. Chem. Theory Comput.* **2016**, *12*, 892–901.
281. Hetényi, C.; Paragi, G.; Maran, U.; Timár, Z.; Karelson, M.; Penke, B. Combination of a Modified Scoring Function with Two-Dimensional Descriptors for Calculation of Binding Affinities of Bulky, Flexible Ligands to Proteins. *J. Am. Chem. Soc.* **2006**, *128*, 1233–1239.
282. Zalevsky, A.; Zlobin, A.; Gedzun, V.; Reshetnikov, R.; Lovat, M.; Malyshev, A.; Doronin, I.; Babkin, G.; Golovin, A. PeptoGrid—Rescoring Function for AutoDock Vina to Identify New Bioactive Molecules from Short Peptide Libraries. *Molecules* **2019**, *24*, 277.
283. Kordnezhadian, R.; Li, B.; Zogu, A.; Demaerel, J.; De Borggraeve, W.M.; Ismalaj, E. Chemistry of Pentafluorosulfonyl Derivatives and Related Analogs: From Synthesis to Applications. *Chem. – A Eur. J.* **2022**, *28*.
284. Savoie, P.R.; Welch, J.T. Preparation and Utility of Organic Pentafluorosulfonyl-Containing Compounds. *Chem. Rev.* **2015**, *115*, 1130–1190.

285. Sowaileh, M.F.; Hazlitt, R.A.; Colby, D.A. Application of the Pentafluorosulfanyl Group as a Bioisosteric Replacement. *ChemMedChem* **2017**, *12*, 1481–1490.
286. Shinada, N.K.; de Brevern, A.G.; Schmidtke, P. Halogens in Protein–Ligand Binding Mechanism: A Structural Perspective. *J. Med. Chem.* **2019**, *62*, 9341–9356.
287. Chiodi, D.; Ishihara, Y. “Magic Chloro”: Profound Effects of the Chlorine Atom in Drug Discovery. *J. Med. Chem.* **2023**, *66*, 5305–5331.
288. Cramer, J.; Krimmer, S.G.; Heine, A.; Klebe, G. Paying the Price of Desolvation in Solvent-Exposed Protein Pockets: Impact of Distal Solubilizing Groups on Affinity and Binding Thermodynamics in a Series of Thermolysin Inhibitors. *J. Med. Chem.* **2017**, *60*, 5791–5799.
289. Gfeller, D.; Michielin, O.; Zoete, V. Shaping the interaction landscape of bioactive molecules. *Bioinformatics* **2013**, *29*, 3073–3079.
290. Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res.* **2019**, *47*, W357–W364.
291. Novak, D.; Anderluh, M.; Kolenc Peitl, P. CCK2R antagonists: from SAR to clinical trials. *Drug Discov. Today* **2020**, *25*, 1322–1336.
292. Varga, A.J.; Nemeth, I.B.; Kemeny, L.; Varga, J.; Tiszlavicz, L.; Kumar, D.; Dodd, S.; Simpson, A.W.M.; Buknicz, T.; Beynon, R.; et al. Elevated Serum Gastrin Is Associated with Melanoma Progression: Putative Role in Increased Migration and Invasion of Melanoma Cells. *Int. J. Mol. Sci.* **2023**, *24*, 16851.
293. Mathieu, V.; Mijatovic, T.; van Damme, M.; Kiss, R. Gastrin Exerts Pleiotropic Effects on Human Melanoma Cell Biology. *Neoplasia* **2005**, *7*, 930–943.
294. Svoboda, D.L.; Saddler, T.; Auerbach, S.S. An Overview of National Toxicology Program’s Toxicogenomic Applications: DrugMatrix and ToxFX. In; 2019; pp. 141–157.
295. Engelberg, A. Iconix Pharmaceuticals, Inc. – removing barriers to efficient drug discovery through chemogenomics. *Pharmacogenomics* **2004**, *5*, 741–744.
296. Gramatica, P. External Evaluation of QSAR Models, in Addition to Cross-Validation: Verification of Predictive Capability on Totally New Chemicals. *Mol. Inform.* **2014**, *33*, 311–314.
297. Mysinger, M.M.; Shoichet, B.K. Rapid Context-Dependent Ligand Desolvation in Molecular Docking. *J. Chem. Inf. Model.* **2010**, *50*, 1561–1573.
298. Jaiteh, M.; Rodríguez-Espigares, I.; Selent, J.; Carlsson, J. Performance of virtual screening against GPCR homology models: Impact of template selection and treatment of binding site plasticity. *PLOS Comput. Biol.* **2020**, *16*, e1007680.
299. Landrum, G.A.; Riniker, S. Combining IC₅₀ or K_i Values from Different Sources Is a Source of Significant Noise. *J. Chem. Inf. Model.* **2024**.
300. Holley, R.W.; Sondheimer, E. The Synthesis of L-Histidyl Peptides 1. *J. Am. Chem. Soc.* **1954**, *76*, 1326–1328.
301. Verma, A. Diisopropylcarbodiimide. *Synlett* **2012**, *23*, 1099–1100.
302. El-Faham, A.; Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem. Rev.* **2011**, *111*, 6557–6602.
303. Dunetz, J.R.; Magano, J.; Weisenburger, G.A. Large-Scale Applications of Amide Coupling Reagents

- for the Synthesis of Pharmaceuticals. *Org. Process Res. Dev.* **2016**, *20*, 140–177.
304. Wünsch, E.; Jaeger, E.; Kisfaludy, L.; Löw, M. Side Reactions in Peptide Synthesis: tert -Butylation of Tryptophan. *Angew. Chemie Int. Ed. English* **1977**, *16*, 317–318.
 305. LÖW, M.; KISFALUDY, L.; SOHÁR, P. tert-Butylierung des Tryptophan-Indolringes während der Abspaltung der tert-Butyloxycarbonyl-Gruppe bei Peptidsynthesen. *Hoppe-Seyler's Zeitschrift für Physiol. Chemie* **1978**, *359*, 1643–1652.
 306. Lundt, B.F.; Johansen, N.L.; Vølund, A.; Markussen, J. REMOVAL OF t-BUTYL AND t-BUTOXYCARBONYL PROTECTING GROUPS WITH TRIFLUOROACETIC ACID. *Int. J. Pept. Protein Res.* **1978**, *12*, 258–268.
 307. Zhang, Z.; Chen, D. Hydrazone–aromatic interaction: the vital role in molecular self-assembly of benzaldehyde hydrazones. *CrystEngComm* **2020**, *22*, 1691–1694.
 308. Zhang, Z.; Wang, L.; Xuan, X. The C□N···C–X σ -hole interaction acts as a conformational lock. *New J. Chem.* **2017**, *41*, 42–46.
 309. Meanwell, N.A. Synopsis of Some Recent Tactical Application of Bioisosteres in Drug Design. *J. Med. Chem.* **2011**, *54*, 2529–2591.
 310. Prinz, C.; Starke, L.; Ramspoth, T.-F.; Kerkering, J.; Martos Riaño, V.; Paul, J.; Neuenschwander, M.; Oder, A.; Radetzki, S.; Adelhoefer, S.; et al. Pentafluorosulfanyl (SF 5) as a Superior 19 F Magnetic Resonance Reporter Group: Signal Detection and Biological Activity of Teriflunomide Derivatives. *ACS Sensors* **2021**, *6*, 3948–3956.
 311. Wilcken, R.; Zimmermann, M.O.; Lange, A.; Joerger, A.C.; Boeckler, F.M. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2013**, *56*, 1363–1388.
 312. Galati, S.; Di Stefano, M.; Martinelli, E.; Poli, G.; Tuccinardi, T. Recent Advances in In Silico Target Fishing. *Molecules* **2021**, *26*, 5124.
 313. Yang, S.; Ye, Q.; Ding, J.; Ming-Zhu Yin; Lu, A.; Chen, X.; Hou, T.; Cao, D. Current advances in ligand-based target prediction. *WIREs Comput. Mol. Sci.* **2021**, *11*.
 314. Huang, N.; Shoichet, B.K.; Irwin, J.J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49*, 6789–6801.
 315. Stein, R.M.; Yang, Y.; Balius, T.E.; O'Meara, M.J.; Lyu, J.; Young, J.; Tang, K.; Shoichet, B.K.; Irwin, J.J. Property-Unmatched Decoys in Docking Benchmarks. *J. Chem. Inf. Model.* **2021**, *61*, 699–714.
 316. Xu, M.; Shen, C.; Yang, J.; Wang, Q.; Huang, N. Systematic Investigation of Docking Failures in Large-Scale Structure-Based Virtual Screening. *ACS Omega* **2022**, *7*, 39417–39428.
 317. Branik, M.; Kessler, H. Zur Konformation geschützter Aminosäuren, III. NMR- und IR-Untersuchungen von Boc-L- α -Aminosäuren. *Chem. Ber.* **1975**, *108*, 2176–2188.
 318. Lipiński, P.F.J.; Matalińska, J. NK1 receptor binding of a few low molecular weight 3,5-bistrifluoromethylbenzene derivatives. *Acta Pol. Pharm. - Drug Res.* **2023**, *80*, 363–372.
 319. Biovia Discovery Studio Visualizer Biovia Discovery Studio Visualizer 2018.
 320. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791.

321. RDKit: Open-source cheminformatics. <https://www.rdkit.org>.
322. Ropp, P.J.; Kaminsky, J.C.; Yablonski, S.; Durrant, J.D. Dimorphite-DL: an open-source program for enumerating the ionization states of drug-like small molecules. *J. Cheminform.* **2019**, *11*, 14.
323. Halgren, T.A. MMFF VI. MMFF94s option for energy minimization studies. *J. Comput. Chem.* **1999**, *20*, 720–729.
324. Pándy-Szekeres, G.; Caroli, J.; Mamyrbekov, A.; Kermani, A.A.; Keserű, G.M.; Kooistra, A.J.; Gloriam, D.E. GPCRdb in 2023: state-specific structure models using AlphaFold2 and new ligand resources. *Nucleic Acids Res.* **2023**, *51*, D395–D402.
325. Schrödinger LLC The PyMOL Molecular Graphics System 2018, sourceforge.net/p/pymol/code/HEAD/tree/trunk/pymol.