



**POLITECHNIKA  
GDAŃSKA**  
Wydział Chemiczny



**WYDZIAŁ  
CHEMICZNY**

Imię i nazwisko autora rozprawy: mgr inż. Joachim Eichenlaub

Dyscyplina naukowa: Nauki chemiczne

## **ROZPRAWA DOKTORSKA**

Tytuł rozprawy w języku polskim: Wykorzystanie modelowania QSPR w celu przewidywania właściwości fizykochemicznych cieczy jonowych

Tytuł rozprawy w języku angielskim: Application of Quantitative Structure – Property Relationship modelling for prediction of ionic liquids properties

Promotor

*podpis*

prof. dr hab. inż. Adam Kloskowski

Gdańsk, rok 2025

## OŚWIADCZENIE

Autor rozprawy doktorskiej: mgr inż. Joachim Eichenlaub

Ja, niżej podpisany, wyrażam zgodę na bezpłatne korzystanie z mojej rozprawy doktorskiej zatytułowanej:

Wykorzystanie modelowania QSPR w celu przewidywania właściwości fizykochemicznych cieczy jonowych

do celów naukowych lub dydaktycznych.<sup>1</sup>

Gdańsk, dnia .....

.....  
*podpis doktoranta*

Świadomy odpowiedzialności karnej z tytułu naruszenia przepisów ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (Dz. U. z 2006 r., nr 90, poz. 631) i konsekwencji dyscyplinarnych określonych w ustawie Prawo o szkolnictwie wyższym (Dz. U. z 2012 r., poz. 572 z późn. zm.),<sup>2</sup> a także odpowiedzialności cywilno-prawnej oświadczam, że przedkładana rozprawa doktorska została napisana przeze mnie samodzielnie.

Oświadczam, że treść rozprawy opracowana została na podstawie wyników badań prowadzonych pod kierunkiem i w ścisłej współpracy z promotorem prof. dr hab. inż. Adamem Kloskowskim

Niniejsza rozprawa doktorska nie była wcześniej podstawą żadnej innej urzędowej procedury związanej z nadaniem stopnia doktora.

Wszystkie informacje umieszczone w ww. rozprawie uzyskane ze źródeł pisanych i elektronicznych, zostały udokumentowane w wykazie literatury odpowiednimi odnośnikami zgodnie z art. 34 ustawy o prawie autorskim i prawach pokrewnych.

Potwierdzam zgodność niniejszej wersji pracy doktorskiej z załączoną wersją elektroniczną.

Gdańsk, dnia .....

.....  
*podpis doktoranta*

Ja, niżej podpisany, wyrażam zgodę na umieszczenie ww. rozprawy doktorskiej w wersji elektronicznej w otwartym, cyfrowym repozytorium instytucjonalnym Politechniki Gdańskiej, Pomorskiej Bibliotece Cyfrowej oraz poddawania jej procesom weryfikacji i ochrony przed przywłaszczaniem jej autorstwa.

Gdańsk, dnia .....

.....  
*podpis doktoranta*

---

<sup>1</sup> Zarządzenie Rektora Politechniki Gdańskiej nr 34/2009 z 9 listopada 2009 r., załącznik nr 8 do instrukcji archiwalnej PG.

<sup>2</sup> Ustawa z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym: Rozdział 7 Odpowiedzialność dyscyplinarna doktorantów, Art. 226.

## **OPIS ROZPRAWY DOKTORSKIEJ**

**Autor rozprawy doktorskiej:** mgr inż. Joachim Eichenlaub

**Tytuł rozprawy doktorskiej w języku polskim:** Wykorzystanie modelowania QSPR w celu przewidywania właściwości fizykochemicznych cieczy jonowych

**Tytuł rozprawy w języku angielskim:** Application of Quantitative Structure – Property Relationship modelling for prediction of ionic liquids properties

**Język rozprawy doktorskiej:** polski

**Promotor rozprawy doktorskiej:** prof. dr hab. inż. Adam Kloskowski

**Słowa kluczowe rozprawy doktorskiej w języku polski:** ciecze jonowe, modelowanie matematyczne, QSPR, ditlenek węgla, zielona chemia, wolna objętość

**Słowa kluczowe rozprawy doktorskiej w języku angielskim:** ionic liquids, mathematical modelling, QSPR, carbon dioxide, green chemistry, free volume

### **Streszczenie rozprawy w języku polskim:**

Przedmiotem rozprawy doktorskiej są badania prowadzone w trakcie kształcenia w Szkole Doktorskiej na Politechnice Gdańskiej. Tematyka badań obejmuje wykorzystanie modelowania QSPR (ang. Quantitative Structure – Property Relationship) w celu analizy rozpuszczalności ditlenku węgla w cieczach jonowych. Do głównych celów badań należało: a) zbudowanie modelu QSPR w celu przewidywania wartości stałej Henry’ego ditlenku węgla w cieczach jonowych, b) przeprowadzenie analizy wpływu wolnej objętości w cieczach jonowych na proces absorpcji ditlenku węgla, c) zbadanie wpływu zmian środowiska oraz sposobu optymalizacji struktur cieczy jonowych na otrzymywane modele QSPR. W wyniku przeprowadzonych badań otrzymano predykcyjny model QSPR i na podstawie analizy i interpretacji deskryptorów molekularnych zidentyfikowano cechy strukturalne cieczy jonowych, które wpływają na rozpuszczalność CO<sub>2</sub>. Udowodniono także, że wolna objętość w cieczach jonowych jest zależna głównie od struktury kationu, a uwzględnienie jej w modelowaniu QSPR pozwala na dokładniejszą analizę wpływu anionu na rozpuszczalność CO<sub>2</sub>. W ostatnim etapie wykazano, że deskryptory molekularne 3D wykazują znaczną podatność na zmiany w sposobie optymalizacji struktur cieczy jonowych, lecz w ogólnym rozrachunku wartości deskryptorów zmieniają się z zachowaniem tego samego trendu dla każdej z metod, co w rezultacie jedynie nieznacznie przekłada się na wpływ na parametry otrzymanych modeli QSPR.

**Streszczenie rozprawy w języku angielskim:**

The subject of this doctoral dissertation is the research that was carried out during the course of study at the Doctoral School at Gdańsk University of Technology. The work focuses on the utilization of QSPR (Quantitative Structure – Property Relationship) modelling to analyze the solubility of carbon dioxide in ionic liquids. The main research objectives included: a) constructing a QSPR model to predict the Henry's Law Constant of carbon dioxide in ionic liquids, b) investigating the effect of the free volume in ionic liquids on the carbon dioxide absorption process, c) examining the impact of various ionic liquids' structure optimization methods on the resulting QSPR models. As a result of the research, a predictive QSPR model was obtained and through the analysis and interpretation of molecular descriptors, the structural features of ionic liquids that influence the CO<sub>2</sub> solubility were identified. It was also demonstrated that free volume in ionic liquids depends primarily on the structure of the cation, and its inclusion in QSPR modeling allows for a more precise evaluation of the anion's contribution to CO<sub>2</sub> solubility. Finally, it was shown that 3D molecular descriptors show significant susceptibility to changes in the method of structure optimization, however, in general, the descriptor values change with the same trend for each method, which as a result only slightly affects the parameters of the obtained QSPR models.

Pragnę w tym miejscu podziękować:

Promotorowi, prof. dr hab. inż. Adamowi Kloskowskiemu,  
za współpracę przez ponad dekadę,  
okazane wsparcie i ogromną cierpliwość,  
a także za wiele ciekawych dyskusji.

Mojej narzeczonej Ewelinie,  
moim rodzicom Annie i Arturowi  
oraz dziadkom Wiesławie i Augustynowi,  
bez których napisanie rozprawy nie byłoby możliwe.  
Tylko dzięki Wam udało się mi osiągnąć wszystko, co sobie zaplanowałem.

Zespołowi badawczemu z Katedry Chemii Fizycznej, a w szczególności  
mgr. inż. Karolowi Baranowi oraz prof. dr hab. inż. Maciejowi Śmiechowskiemu,  
oraz byłym pracownikom i doktorantom, dr inż. Karolinie Delińskiej,  
dr inż. Katerynie Yavir oraz dr inż. Paulinie Rakowskiej  
za przyjazną atmosferę pracy i otrzymaną pomoc.

## SPIS TREŚCI

|                                                                                             |    |
|---------------------------------------------------------------------------------------------|----|
| WYKAZ NAJWAŻNIEJSZYCH OZNACZEŃ I SKRÓTÓW.....                                               | 8  |
| 1. WPROWADZENIE.....                                                                        | 10 |
| 1.1. Ciecze jonowe.....                                                                     | 10 |
| 1.1.1. Wstęp.....                                                                           | 10 |
| 1.1.2. Rozwiewanie mitów wokół cieczy jonowych.....                                         | 14 |
| 1.1.3. Ciecze jonowe w przemyśle.....                                                       | 15 |
| 1.1.4. Zależność wybranych właściwości fizykochemicznych cieczy jonowych od ich budowy..... | 16 |
| 1.2. Modelowanie QSPR.....                                                                  | 20 |
| 1.2.1. Wytyczne OECD.....                                                                   | 21 |
| 1.2.2. Etapy modelowania.....                                                               | 22 |
| 1.2.3. Metody optymalizacji struktury.....                                                  | 23 |
| 1.2.4. Deskryptory molekularne.....                                                         | 26 |
| 1.2.5. Badania QSPR dotyczące cieczy jonowych.....                                          | 27 |
| 1.3. Rozpuszczalność gazów.....                                                             | 29 |
| 1.4. Zielone rozpuszczalniki.....                                                           | 34 |
| 1.5. Ditlenek węgla jako gaz cieplarniany.....                                              | 37 |
| 1.6. Wolna objętość w cieczach jonowych.....                                                | 37 |
| 2. CELE I TEZY PRACY.....                                                                   | 39 |
| 3. METODY BADAWCZE.....                                                                     | 40 |
| 3.1. Przygotowanie danych.....                                                              | 40 |
| 3.2. Optymalizacja struktur i obliczenie deskryptorów molekularnych.....                    | 43 |
| 3.2.1. Modelowanie stałej Henry'ego.....                                                    | 43 |
| 3.2.2. Analiza wolnej objętości.....                                                        | 44 |
| 3.2.3. Analiza metod optymalizacji.....                                                     | 45 |
| 3.3. Podział zbioru danych.....                                                             | 46 |
| 3.3.1. Modelowanie stałej Henry'ego.....                                                    | 47 |
| 3.3.2. Analiza wolnej objętości.....                                                        | 47 |
| 3.3.3. Analiza metod optymalizacji.....                                                     | 48 |
| 3.4. Modelowanie QSPR.....                                                                  | 48 |
| 3.4.1. Modelowanie stałej Henry'ego.....                                                    | 48 |
| 3.4.2. Analiza wolnej objętości.....                                                        | 50 |
| 3.4.3. Analiza metod optymalizacji.....                                                     | 51 |
| 3.5. Walidacja modeli.....                                                                  | 52 |
| 4. WYNIKI.....                                                                              | 56 |
| 4.1. Modelowanie stałej Henry'ego.....                                                      | 56 |
| 4.1.1. Walidacja.....                                                                       | 57 |
| 4.1.2. Interpretacja deskryptorów.....                                                      | 61 |
| 4.1.2.1. Deskryptory kationowe.....                                                         | 62 |
| 4.1.2.2. Deskryptory anionowe.....                                                          | 64 |
| 4.2. Analiza wolnej objętości.....                                                          | 66 |
| 4.2.1. Frakcyjna wolna objętość.....                                                        | 66 |

|                                                                      |     |
|----------------------------------------------------------------------|-----|
| 4.2.2. Modele QSPR.....                                              | 69  |
| 4.2.2.1. FFV jako zmienna zależna.....                               | 71  |
| 4.2.2.2. Model standardowy.....                                      | 72  |
| 4.2.2.3. Model FFV.....                                              | 73  |
| 4.2.2.4. Omówienie i porównanie modeli.....                          | 74  |
| 4.2.3. Interpretacja deskryptorów.....                               | 75  |
| 4.2.3.1. Model standardowy.....                                      | 76  |
| 4.2.3.2. Model FFV.....                                              | 77  |
| 4.3. Analiza metod optymalizacji.....                                | 78  |
| 4.3.1. Zestawienie parametrów statystycznych otrzymanych modeli..... | 79  |
| 4.3.2. Porównanie metod modelowania.....                             | 87  |
| 4.3.3. Porównanie optymalizacji jonów osobno i razem.....            | 90  |
| 4.3.4. Porównanie optymalizacji jonów w próżni i w etanolu.....      | 92  |
| 4.3.5. RMSE.....                                                     | 94  |
| 4.3.6. Wpływ sposobu optymalizacji na wartości deskryptorów.....     | 95  |
| 4.3.6.1. Optymalizacja osobno i razem.....                           | 95  |
| 4.3.6.2. Optymalizacja w próżni i w etanolu.....                     | 97  |
| 4.3.6.3. Deskryptory o największej zmienności.....                   | 99  |
| 4.3.7. Zmienność deskryptorów a otrzymane modele QSPR.....           | 101 |
| 5. WNIOSKI.....                                                      | 104 |
| 5.1. Modelowanie stałej Henry'ego.....                               | 104 |
| 5.2. Analiza wolnej objętości.....                                   | 105 |
| 5.3. Analiza metod optymalizacji.....                                | 105 |
| 6. PODSUMOWANIE.....                                                 | 108 |
| 7. BIBLIOGRAFIA.....                                                 | 109 |
| 8. DOROBEK NAUKOWY.....                                              | 127 |
| SPIS ILUSTRACJI I TABEL.....                                         | 128 |
| MATERIAŁY UZUPEŁNIAJĄCE.....                                         | 130 |

## WYKAZ NAJWAŻNIEJSZYCH OZNACZEŃ I SKRÓTÓW

### Ciecze jonowe:

#### Kationy:

|                                                     |                                                            |
|-----------------------------------------------------|------------------------------------------------------------|
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | N,N,N',N'-tetrametyloformamidynowy                         |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [[[(dimetyloamino)metyleno]amino]metyleno]dimetyloamoniowy |
| [1,2BzMPY]                                          | 1-benzylo-2-metylopirydyniowy                              |
| [1,3BzMPY]                                          | 1-benzylo-3-metylopirydyniowy                              |
| [1,4BzMPY]                                          | 1-benzylo-4-metylopirydyniowy                              |
| [B3MPy]                                             | 1-butylo-3-metylopirydyniowy                               |
| [B4MPy]                                             | 1-butylo-4-metylopirydyniowy                               |
| [BDMIM]                                             | 1-butylo-2,3-dimetyloimidazoliowy                          |
| [BHEA]                                              | bis(2-hydroksyetylo)amoniowy                               |
| [BMIM]                                              | 1-butylo-3-metyloimidazoliowy                              |
| [BMPyrr]                                            | 1-butylo-1-metylopirolidynowy                              |
| [BzMIM]                                             | 1-benzylo-3-metyloimidazoliowy                             |
| [BzMPyrr]                                           | 1-benzylo-1-metylopirolidynowy                             |
| [BzPY]                                              | 1-benzylopirydyniowy                                       |
| [C3CNDMIM]                                          | 3-(3-cyjanopropilo)-1,2-dimetyloimidazoliowy               |
| [C3CNMIM]                                           | 1-(3-cyjanopropilo)-3-metyloimidazoliowy                   |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | 1-metylo-3-(nonafluoroheksylo)imidazoliowy                 |
| [C <sub>8</sub> H <sub>4</sub> F <sub>13</sub> MIM] | 1-metylo-3-(tridekafluorooktylo)imidazoliowy               |
| [DMPIM]                                             | 1,2-dimetylo-3-propyloimidazoliowy                         |
| [EDMIM]                                             | 1-etylo-2,3-dimetyloimidazoliowy                           |
| [EMIM]                                              | 1-etylo-3-metyloimidazoliowy                               |
| [ETT]                                               | S-etylo-N,N,N',N'-tetrametylotiomocznikowy                 |
| [HEA]                                               | 2-hydroksyetyloamoniowy                                    |
| [HEEA]                                              | 2-(2-hydroksyetoksy)etyloamoniowy                          |
| [HHEMEA]                                            | 2-hydroksy-N-(2-hydroksyetylo)-N-metyloetyloamoniowy       |
| [HMIM]                                              | 1-heksylo-3-metyloimidazoliowy                             |
| [HMPy]                                              | 1-heksylo-3-metylopirydyniowy                              |
| [HOEtMIM]                                           | 1-(2-Hydroksyetylo)-3-metyloimidazoliowy                   |
| [MHEA]                                              | N-metylo-2-hydroksyetyloamoniowy                           |
| [N <sub>4,1,1,1</sub> ]                             | trimetylo(butylo)amoniowy                                  |
| [OMIM]                                              | 1-metylo-3-oktyloimidazoliowy                              |
| [P66614]                                            | triheksylo(tetradecylo)fosfoniowy                          |
| [PentMIM]                                           | 1-pentylo-3-metyloimidazoliowy                             |
| [PhBIM]                                             | 1-butylo-3-feniloimidazoliowy                              |
| [PhDBIM]                                            | 3,5-dibutylo-1-feniloimidazoliowy                          |
| [PMIM]                                              | 1-metylo-3-propyloimidazoliowy                             |
| [THEA]                                              | tri(2-hydroksyetylo)amoniowy                               |
| [DMIM]                                              | 1,3-dimetyloimidazoliowy                                   |

#### Aniony:

|                       |                                            |
|-----------------------|--------------------------------------------|
| [OAc]                 | octan                                      |
| [NTf <sub>2</sub> ]   | bis(trifluorometanosulfonylo)imidek        |
| [Br]                  | bromek                                     |
| [Cl]                  | chlorek                                    |
| [N(CN) <sub>2</sub> ] | dicyjanoamid                               |
| [HCOO]                | mrówczan                                   |
| [PF <sub>6</sub> ]    | heksafluorofosforan                        |
| [Lac]                 | mleczan                                    |
| [BF <sub>4</sub> ]    | tetrafluoroboran                           |
| [C(CN) <sub>3</sub> ] | tricyjanometanian                          |
| [TA]                  | trifluorooctan                             |
| [TfO]                 | trifluorometanosulfonian                   |
| [pFAP]                | tris(heptafluoropropylo)trifluorofosfonian |
| [bFAP]                | tris(nonafluorobutylo)trifluorofosfonian   |
| [eFAP]                | tris(pentafluoroetylo)trifluorofosfonian   |

#### Pozostałe skróty:

|                  |                                                        |
|------------------|--------------------------------------------------------|
| QSPR             | Quantitative Structure - Property Relationship         |
| IL(s)            | ciecz jonowa (ciecze jonowe)                           |
| MLR              | wielokrotna regresja liniowa                           |
| LR               | regresja logistyczna                                   |
| PLS              | metoda cząstkowych najmniejszych kwadratów             |
| CO <sub>2</sub>  | ditlenek węgla                                         |
| H <sub>CO2</sub> | stała Henry'ego dla ditlenku węgla w cieczach jonowych |
| FV               | wolna objętość                                         |
| FFV              | cząstkowa wolna objętość                               |

# 1. WPROWADZENIE

## 1.1. Ciecze jonowe

### 1.1.1. Wstęp

Ciecze jonowe to klasa związków chemicznych, która w ciągu kilkudziesięciu lat od ich odkrycia, ze stosunkowo nieznaną grupy stała się jednym z głównych przedmiotów badań we współczesnej chemii. Zainteresowanie cieczami jonowymi wynikało z ich właściwości – bardzo niskiej temperaturze topnienia, niskiej prężności pary, wysokiej stabilności termicznej oraz elektrochemicznej, a także powszechnemu przeświadczeniu o ich nieszkodliwości dla środowiska.[1,2] Choć obecnie część ze wspomnianych właściwości została zakwestionowana, co szczególnie widać na przykładzie badań dotyczących ich toksyczności [3–5], które udowadniają, że zachwyty środowiska naukowego nad wynalezieniem „zielonych rozpuszczalników” były przedwczesne i przesadzone, badania nad cieczami jonowymi wciąż pozostają jedną z bardziej interesujących gałęzi rozwoju chemii.[6]

Pierwsze doniesienia o stopionych solach, które pozostawały w ciekłym stanie skupienia w temperaturze pokojowej, pochodzą z 1914 roku, kiedy Paul Walden odkrył, że badany przez niego azotan etyloamoniowy charakteryzuje się temperaturą topnienia wynoszącą 12 °C.[7] Ówczesne środowisko naukowe nie przywiązało jednak odpowiedniej wagi do tego odkrycia i aby ciecze jonowe znalazły swoje pierwsze zastosowanie minąć musiało ponad czterdzieści lat. Pierwsze odnotowane w historii użycie terminu „ciecz jonowa” miało miejsce w 1943 roku.[8] W roku 1948, opatentowano pierwszą ciecz opartą na anionie tetrachloroglinianowym.[9]

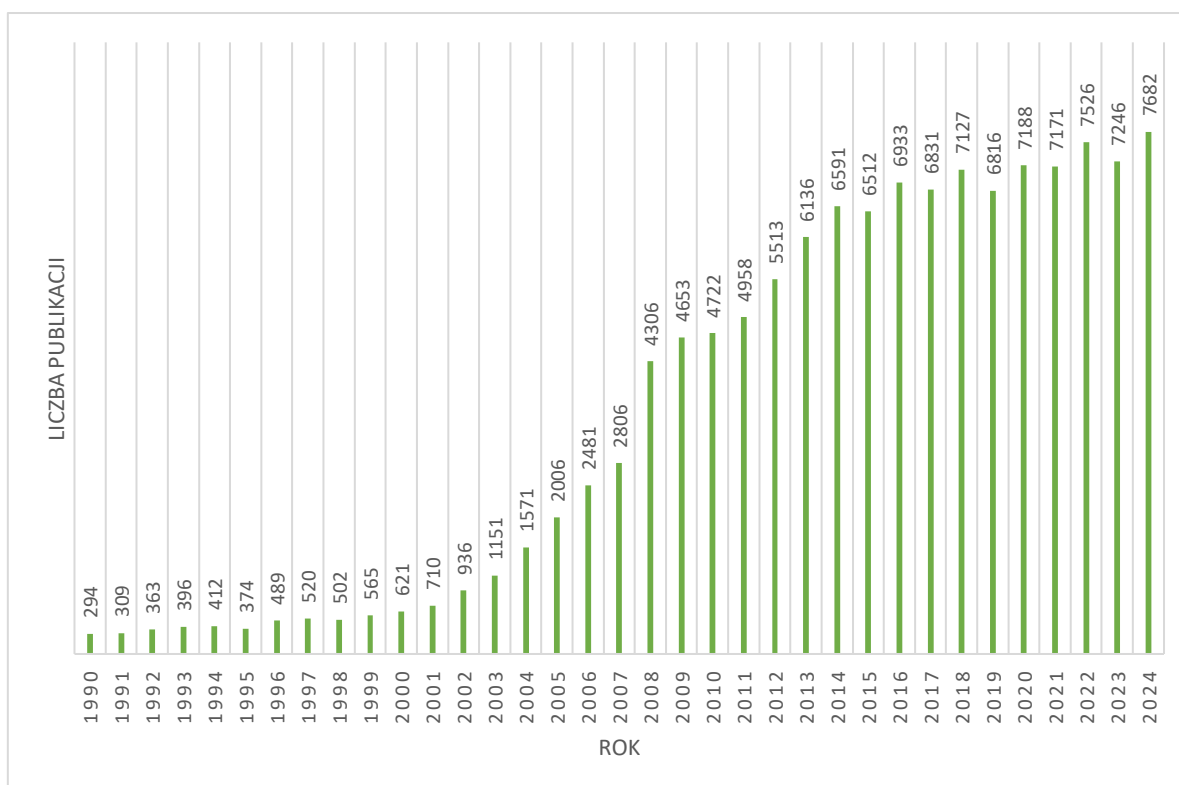
Doniesienia na temat cieczy jonowych, które pojawiały się w tamtym czasie, nie były jednak szeroko rozpowszechnione. Na ten moment należało zaczekać do lat 70. XX wieku, kiedy kilka zespołów badawczych niezależnie od siebie skupiło się na chlorometalicznych cieczach jonowych uzyskiwanych z chlorku glinu, a ich odkrycia trafiły w środowiskach akademickich do szerokiej rzeszy odbiorców i spopularyzowały koncepcję badań nad „stopionymi solami ciekłymi w temperaturze pokojowej” (*ang. room-temperature molten salts*).[10–12] Ich zastosowanie wciąż jednak było ograniczone – wszystkie badane ciecze jonowe były podatne na wilgoć, a najmniejszy nawet dodatek wody powodował ich hydrolizę, co zmuszało naukowców do poszukiwania nowych typów cieczy jonowych. W latach 80. John Wilkes rozpoczął pracę nad cieczami opartymi o kation 1,3-dialkiloimidazolowy, a badania te doprowadziły w 1992 roku do wynalezienia klasy cieczy jonowych, które pozostawały stabilne zarówno w powietrzu, jak i w kontakcie z wodą.[13] Od tego momentu zainteresowanie cieczami systematycznie rosło. Do syntezy nowych cieczy jonowych wykorzystywano coraz to szerszą gamę związków, a różnorodność kationów i anionów rosła w szybkim tempie. Ciecze jonowe zaczęły być badane pod

kątem wykorzystania nie tylko jako reagenty, ale również jako rozpuszczalniki i katalizatory reakcji.[14]  
Najważniejsze momenty w historii badań nad cieczami jonowymi przedstawiono na Rysunku 1.1.[15]



**Rysunek 1. 1.** Wybrane wydarzenia w historii badań nad cieczami jonowymi

Na początku XXI wieku ciecze jonowe stały się szerzej dostępne dla środowiska naukowego ze względu na pojawienie się ich w sprzedaży u największych dystrybutorów odczynników chemicznych, co wyeliminowało problem związany z koniecznością ich samodzielnej syntezy. Kombinacja wszystkich wspomnianych czynników spowodowała ogromny wzrost zainteresowania cieczami jonowymi, co widać na Rysunku 1.2, który przedstawia wzrost liczby artykułów naukowych dotyczących cieczy jonowych w bazie SCOPUS na przestrzeni lat 1990-2024 (dostęp 15 maja 2025). Największy trend wzrostowy można zaobserwować na początku lat dwutysięcznych. W roku 1994 opublikowano 412 artykułów, w roku 2004 było to już 1571, a w 2014 aż 6591. Tempo publikacji zmalało i ustabilizowało się w ostatnich latach, co może być spowodowane pandemią Covid-19 w latach 2020-2022. W 2024 opublikowano 7628 artykułów na temat cieczy jonowych, co w odniesieniu do wcześniejszych wzrostów na przestrzeni dekad jest niewielką zmianą. Mimo to, zauważalny jest wyraźny trend wzrostowy, co pokazuje, że ciecze jonowe wciąż cieszą wielkim zainteresowaniem.



**Rysunek 1. 2.** Liczba publikacji dotyczących cieczy jonowych indeksowanych w bazie SCOPUS na przestrzeni lat

Rosnąca popularność cieczy zbiegła się w czasie z opublikowaniem w 1998 roku przez Paula Anastasa i Johna Warnera Dwunastu Zasad Zielonej Chemii, z których piąta zasada dotyczyła stosowania zielonych rozpuszczalników.[16] W tamtym czasie ciecze jonowe uznawane były za bardziej ekologiczną, „zieloną” alternatywę dla stosowanych klasycznych rozpuszczalników organicznych, które cechują się przeważnie wysoką lotnością oraz toksycznością dla środowiska. W momencie publikacji Dwunastu Zasad w środowisku naukowym przeważał pogląd, że ciecze są nielotne, natomiast dostępność badań dotyczących ich toksyczności była niewielka. Mimo, że obecnie wiadomo na ten temat więcej – ciecze jonowe charakteryzują się niewielką, aczkolwiek nie zaniedbywalną prężnością par [17], a ich toksyczność jest w większości przypadków porównywalna do rozpuszczalników organicznych, które mają zastąpić – związki te wciąż wpisują się piątą zasadę Zielonej Chemii dotyczącą używania substancji pomocniczych.[18] Mówi ona o tym, że [16]:

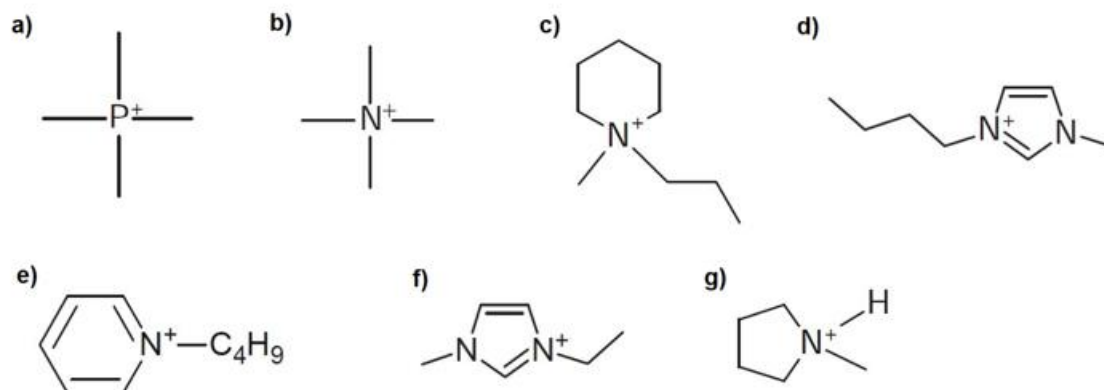
„Używanie substancji pomocniczych (roztworów, środków separacyjnych i tym podobnych) powinno zostać wyeliminowane, gdzie tylko jest to możliwe, a we wszystkich innych sytuacjach używać należy substancji nieszkodliwych”

Ze względu na swoją stabilność termiczną oraz elektrochemiczną oraz bardzo niską lotność, ryzyko niekontrolowanego opuszczenia środowiska reakcji przez ciecz jonową jest niewielkie.

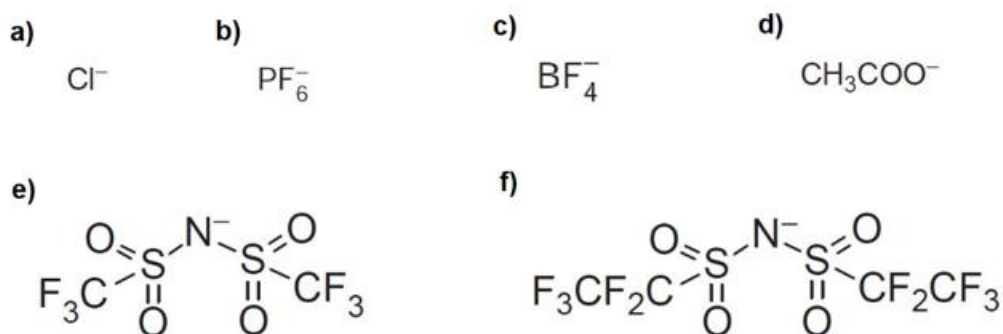
Po publikacji Zasad, badania nad właściwościami fizykochemicznymi cieczy jonowych zintensyfikowano, ponieważ spodziewano się, że mogą przyczynić się do zredukowania emitowanych do środowiska antropogenicznych zanieczyszczeń. W 2000 roku ciecze jonowe zostały po raz pierwszy nazwane „zielonymi rozpuszczalnikami”. [2] Właściwa implementacja Zasad Zielonej Chemii stanowiła jedno z najważniejszych wyzwań stojących przed naukowcami i przemysłem ponad dwadzieścia lat temu i do dziś nie utraciła swojej wagi. Mimo, że część wniosków i założeń dotyczących wykorzystania cieczy jonowych, jako zielonych rozpuszczalników okazała się zbyt optymistyczna i konsekwentnie zrewidowana, ówczesne zainteresowanie środowiska naukowego cieczami jonowymi doprowadziło do rozkwitu zainteresowania cieczami, a każdego roku publikowanych było ponad 3500 nowych artykułów na ich temat. Obecnie rozwijane i badane zastosowania cieczy obejmują szeroki i zróżnicowany zakres dziedzin chemii, medycyny i przemysłu. Wymienić można chociażby wykorzystanie cieczy jonowych do wychwytywania ditlenku węgla ze spalin [19], usuwania jonów metali ciężkich z gleby i ścieków za pomocą separacji membranowej [20], wykorzystania jako chłodziwa podczas obróbki metali [21] czy jako środki antybakteryjne w materiałach stomatologicznych. [22]

Klasyczne ciecze jonowe składają się z dużego organicznego kationu oraz mniejszego, często, nieorganicznego anionu. Najczęściej spotykane kationy zbudowane są na bazie związków cyklicznych – imidazolu, pirydyny, piperidyny i pirolidyny – oraz z czwartorzędowych amin i jonów fosfoniowych. [23] Każdy z tych związków może zostać w mniej lub bardziej skomplikowany sposób zmodyfikowany podczas syntezy, przez wprowadzenie do struktury szeregu zróżnicowanych chemicznie podstawników. [24] Mogą być to proste łańcuchy alkilowe o określonej długości, nieorganiczne grupy funkcyjne, a nawet rozbudowane organiczne podstawniki o złożonej budowie, jak na przykład 2,2,2-trifluoroetyl, 3-(3-trietoksylanopropyl) lub 3-fenylpropyl. [25,26] Różnorodność anionów jest mniejsza, ale wciąż bardzo duża. Ze wspomnianych powodów ciecze jonowe opisywane są często jako „rozpuszczalniki projektowalne”, gdyż mogą poprzez dobór odpowiedniego kationu i anionu zostać zsyntezowane w taki sposób, aby wykazywały pożądane właściwości fizykochemiczne o ściśle określonych wartościach. [27] Najpopularniejsze kationy i aniony przedstawiono poniżej na Rysunku 1.3.

### KATIONY



### ANIONY



**Rysunek 1. 3.** Wybrane wzory kationów i anionów tworzących cieczy jonowe

Istnieją miliony możliwych kombinacji kationów i anionów, które mogą utworzyć ciecz jonową.[28] Zsyntezowanie tak ogromnego zbioru związków chemicznych znacznie przekracza obecne możliwości nawet najbardziej wydajnych i nowoczesnych laboratoriów. Niemożliwe jest więc otrzymanie i przebadanie wszystkich możliwych cieczy jonowych. Jest to zarówno czasochłonne, jak i niezwykle kosztowne. Obecne ceny cieczy wśród największych dostawców odczynników laboratoryjnych wahają się w granicach około 250-1000\$ za 100 g w zależności od stopnia skomplikowania syntezy oraz dostępności półproduktów.[29] Potrzebna jest metoda przesiewowa, która pozwala odrzucić na wstępie cieczy o niepożądanych właściwościach i wskazać kierunek syntezy i prowadzenia dalszych badań.

#### 1.1.2. Rozwiewanie mitów wokół cieczy jonowych

Jak wspomniano już wcześniej, przez wiele lat cieczy jonowe uważane były za nietlotne i nietoksyczne. W połączeniu z ich wysoką stabilnością termiczną i elektrochemiczną uważane były za

zielone rozpuszczalniki niemalże pozbawione wad. Jednakże, z biegiem lat ich „cudowne właściwości” zostały zakwestionowane, a część z nich okazała się nieprawdą.[30]

Stabilność termiczna cieczy została zweryfikowana przez Kosmulskiego i innych, a w artykule udowodniono, że rzeczywista temperatura rozkładu wielu cieczy jonowych opartych na pierścieniu imidazolowym jest często niższa o niemal 200 °C w porównaniu do wcześniejszych doniesień.[31] Również brak lotności cieczy jonowych okazał się mitem. Badania Earla i innych zakończyły się przeprowadzeniem z powodzeniem destylacji analizowanych cieczy jonowych, a wyniki opublikowano na łamach Nature.[17] Co więcej, znaczna część cieczy okazała się być szkodliwa dla organizmów żywych.[32] Badania toksyczności cieczy jonowych są jednymi z pierwszych zastosowań modelowania QSPR w analizie tych związków. Zbudowano ponad czterdzieści modeli, a na wspomnienie zasługują w tym miejscu prace prof. Tomasza Puzyna z Uniwersytetu Gdańskiego, którego zespół był bardzo aktywny w tej dziedzinie badań, co poskutkowało zbudowaniem wielu świetnej jakości modeli do testowania toksyczności cieczy jonowych *in silico*. [33–35]

Badania nad szkodliwością cieczy jonowych wskazały, że kierunkiem rozwoju mogą być ciecze zbudowane na bazie naturalnie występujących związków organicznych, takich jak aminokwasy lub kwasy karboksylowe.[24] Kapitanov i inni zsyntezowali ciecze jonowe na bazie aminokwasów i udowodnili, że cechowały się one znacznie mniejszą toksycznością wobec bakterii.[36] W swojej pracy Fan i inni zbadali toksyczność cieczy jonowych na bazie kwasów tłuszczowych wobec populacji kreketek i doszli do podobnych wniosków.[37] „Naturalne” ciecze jonowe są obecnie bardzo popularną dziedziną badań, ale ze względu na ich niewielką dostępność komercyjną i konieczność samodzielnej syntezy w laboratorium, również w ich przypadku pomiary eksperymentalne wymagają wsparcia modelowaniem matematycznym.

#### 1.1.3. Ciecze jonowe w przemyśle

Ciecze jonowe wykorzystywane są w przemyśle, zarówno w instalacjach pilotażowych, jak i na większą skalę. W 2020 roku było 57 znanych przypadków przemysłowego wykorzystania cieczy.[38]

Największe przedsiębiorstwo chemiczne na świecie, niemiecki BASF SE w 2002 roku zaczął stosować 1-metylimidazol w procesie produkcji alkoksyfenylfosfiny, zastępując używane wcześniej trialkiloaminy. Aminy były wykorzystywane do wychwytywania kwasu chlorowodorowego, który powstawał jako produkt uboczny reakcji. Wyzwanie stanowiła jednak późniejsza separacja mieszaniny i otrzymanie czystego produktu końcowego. W wyniku zamiany amin na 1-metylimidazol powstającym produktem ubocznym była ciecz jonowa, chlorek 1-metylimidazolu, która okazała się być zdecydowanie łatwiejsza w separacji i w późniejszej regeneracji, efektywnie redukując koszty procesu.[39]

Swoje zastosowanie ciecz jonowe znajdują również w elektrochemii. Przedsiębiorstwo Scionix wykorzystuje chlorek choliny jako elektrolit w procesie elektroosadzania powłok chromowych ze względu na jego wysoką stabilność elektrochemiczną oraz wysokie przewodnictwo. Zastosowanie chlorku choliny zwiększyło wydajność procesu oraz wpłynęło pozytywnie na jakość otrzymywanych powłok.[40]

Ciecze jonowe charakteryzują się nie tylko szerokim zakresem okna elektrochemicznego, ale są również praktycznie niepalne, a obie te cechy zwracają uwagę producentów baterii i akumulatorów. Po zastosowaniu cieczy jonowych jako elektrolitów w produkowanych przez siebie bateriach przedsiębiorstwo NOHMs Technologies (obecnie Sionic Energy) uzyskało produkty o 400% dłuższym czasie pracy.[41]

Ciecze wykazują potencjał do zastąpienia nie tylko rozpuszczalników organicznych. Amerykańska korporacja Chevron w jednym z procesów alkilacji zastąpiła wykorzystywany jako katalizator kwas siarkowy (VI) mieszaniną chloru i cieczy na bazie chlorku glinu. W rezultacie otrzymano katalizator o znacznie mniej korozyjnych właściwościach. Co prawda zamiana nie wpłynęła na zmniejszenie kosztów procesu, ale zwiększyła jego wydajność oraz bezpieczeństwo.[6,42,43]

Inna amerykański koncern, 3M, wykorzystuje ciecz jonowe jako antystatki przy produkcji materiałów polimerowych.[44] Ciecze znajdują zastosowanie również w procesie recyklingu polimerów. Współfinansowane przez Unilever, The Coca-Cola Company oraz Indorama Ventures, czyli największego na świecie producenta politereftalanu etylenu (PET), przedsiębiorstwo Ionica wykorzystuje ciecz jonowe w procesie recyklingu obejmującym 10 000 ton PET rocznie.[45]

Wybrane przypadki ilustrują zaangażowanie największych korporacji w wykorzystanie cieczy jonowych w przemyśle, a także mnogość ich możliwych zastosowań. Istnieje obecnie wiele mniejszych, pilotażowych instalacji przemysłowych, w których ciecz jonowe odgrywają kluczową rolę w procesach separacji gazów lub katalizy.[38,46] Nie ulega wątpliwości, że w przyszłości zainteresowanie nimi będzie rosło.

#### 1.1.4. Zależność wybranych właściwości fizykochemicznych cieczy jonowych od ich budowy

##### a) temperatura topnienia

Temperatura topnienia zależy zarówno od typu anionu, jak i kationu. Znaczny wpływ ma tutaj asymetria cząsteczek, rozumiana zarówno jako różnica w rozmiarze kationu i anionu oraz jako obecność kilku asymetrycznych podstawników w jednym z jonów, np. krótkiego i długiego łańcucha alkilowego. W obu przypadkach wraz ze wzrostem różnic temperatura topnienia maleje.[47] Wpływ ma również długość łańcuchów alkilowych, którymi podstawione są najczęściej kationy. W przypadku cieczy

opartych na kationie imidazolowym temperatura topnienia osiąga swoje minimum dla łańcuchów o długości 4-6 atomów węgla [48], lecz przełożenie tej zależności na inne kationy nie przynosi efektów. Obecnie uważa się, że nie istnieje jednoznaczny trend we wpływie długości łańcucha alkilowego na temperaturę topnienia.[49] Aby uzyskać niską wartość temperatury topnienia w strukturze cieczy jonowej powinny pojawić się grupy alkoksylowe [50] oraz nienasycone łańcuchy węglowodorowe.[51] Unikać należy, również, obecności atomów fluoru oraz jodu [49,52].

#### b) rozpuszczalność w układach ciecz-ciecz

Wpływ budowy cieczy jonowych na ich rozpuszczalność, co oczywiste, musi być rozpatrywany w odniesieniu do konkretnego rozpuszczalnika (lub typu rozpuszczalników). W przypadku kationu, znaczny wpływ na rozpuszczalność ma długość łańcucha alkilowego. Im jest dłuższy, tym gorsza rozpuszczalność w wodzie.[53,54] Rodzaj pierścienia, z którego zbudowana jest ciecz jonowa, ma duży wpływ na rozpuszczalność w rozpuszczalnikach polarnych. Rozpuszczalność cieczy jonowych maleje w nich w kolejności: imidazol > pirydyna > pirolidyna > piperydyna.[53] Występowanie w strukturze cieczy jonowej hydroksylowych lub karboksylowych grup funkcyjnych poprawia ich rozpuszczalność w wodzie.[55] Odwrotny efekt wykazują podstawniki lipofilowe.[55,56] Rozpuszczalność w związkach organicznych nie rządzi się uniwersalnymi prawami i różni się znacznie w zależności od analizowanej substancji. [57] W zależności od rozpuszczalnika, może on wchodzić w interakcje z różnymi fragmentami struktury cieczy jonowych. Np. rozpuszczalność w metanolu zależy od możliwości wytworzenia wiązań wodorowych z anionem oraz w znacznie mniejszej mierze z kationem. Rozpuszczalność w alkanach jest związana z obecnością niepolarnych struktur, jak na przykład długich łańcuchów alkilowych, a aceton lub acetonitryl oddziałują zarówno z polarnymi, jak i z niepolarnymi fragmentami strukturalnymi cieczy jonowej.[58]

#### c) gęstość

Gęstość maleje wraz ze wzrostem długości łańcucha, niezależnie, czy jest on powiązany z anionem, czy z kationem.[59] Jest to związane z przestrzennym ułożeniem cieczy jonowej w układzie warstwowym i z podziałem na część polarną i na część niepolarną. Wzrost długości łańcucha sprawia, że warstwa niepolarna, która jest dość luźno upakowana, zwiększa swoją objętość, co przekłada się na niższą gęstość.[60] Większa gęstość wiąże się z występowaniem grup -OH w łańcuchach alkilowych, które tworzą wiązania wodorowe z pierścieniami kationów, stabilizując je i dając w rezultacie większe upakowanie.[61] Wpływ anionu jest dużo mniejszy i głównie wiąże się liniowo ze wzrostem masy molowej.[62]

#### d) lepkość

W większości przypadków lepkość cieczy jonowej wzrasta wraz ze wzrostem długości łańcucha alkilowego przy kationie.[63] Aniony wpływają na lepkość głównie poprzez swój rozmiar.[64,65] Duży anion  $[\text{NTf}_2]$  znacznie redukuje lepkość, natomiast mniejszy  $[\text{PF}_6]$  zauważalnie ją zwiększa.[66] W przypadku kationów ważny jest typ rdzenia cieczy jonowej – np. ciecze oparte na czwartorzędowych zasadach amonowych mają znacznie wyższe lepkości, niż ciecze oparte na pierścieniach imidazolowych.[67]

#### e) napięcie powierzchniowe

Napięcie powierzchniowe jest zależne od długości łańcucha alkilowego przy kationie i maleje wraz ze wzrostem jego długości aż do osiągnięcia *plateau* przy 6-8 atomach węgla.[62,68] Napięcie wzrasta w przypadku obecności większości grup funkcyjnych (niezależnie od ich budowy) na terminalnym atomie węgla w łańcuchu alkilowym [69], a także przy zastąpieniu pierścienia imidazolowego piroolidynowym.[62] Wpływ anionu uważany jest za mniejszy od wpływu kationu. Nieznaczne wahania w wartości napięcia powierzchniowego zaobserwować można dla szeregu homologicznych cieczy jonowych opartych na anionach halogenkowych, gdzie wzrost napięcia powierzchniowego ściśle związany jest ze wzrostem objętości anionu ( $\text{I}^- > \text{Br}^- > \text{Cl}^-$ ).[69] Spadek napięcia uzyskać można poprzez wysycenie łańcucha alkilowego atomami fluoru.[62]

#### f) Ciecze jonowe jako rozpuszczalniki - rozpuszczalność gazów

Podobnie jak w przypadku omawianej wcześniej rozpuszczalności cieczy jonowych w układach ciecz-ciecz, niemożliwym jest opisanie jednego wspólnego mechanizmu rozpuszczania gazów oraz wskazanie uniwersalnych zależności pomiędzy strukturą cieczy a rozpuszczalnością. Należy zauważyć również, że o ile w większości układów absorpcja zachodzi na zasadzie fizysorpcji, możliwe są także przypadki, w których za większą część absorpcji odpowiada chemisorpcja.[70,71] W identycznych warunkach ciśnienia i temperatury rozpuszczalność gazów w większości cieczy jonowych można przedstawić w następującej kolejności: ditlenek siarki ( $\text{SO}_2$ ) > ditlenek węgla ( $\text{CO}_2$ ) = tlenek azotu (I) ( $\text{N}_2\text{O}$ ) > eten ( $\text{C}_2\text{H}_4$ ) > etan ( $\text{C}_2\text{H}_6$ ) > metan ( $\text{CH}_4$ ) > argon (Ar) > tlen ( $\text{O}_2$ ) > azot ( $\text{N}_2$ ) > tlenek węgla (CO).[72]

Rozpuszczalność ditlenku siarki w cieczach jonowych jest zauważalnie wyższa, niż rozpuszczalność ditlenku węgla. Wpływ rodzaju anionu jest niewielki, podobnie jak długość łańcucha węglowego. Dużo istotniejsze okazują się oddziaływania sił van der Waalsa pomiędzy cząsteczkami gazu a zdelokalizowanymi elektronami  $\pi$  w kationie i to właśnie jego rodzaj ma decydujący wpływ na

rozpuszczalność. W przypadku cieczy na bazie amin zachodzi głównie chemisorpcja, a  $\text{SO}_2$  przereagowuje z grupami  $-\text{NH}$  w kationie.[73]

Ditlenek węgla jest jednym z najczęściej analizowanych pod kątem rozpuszczalności w cieczach jonowych gazów. Obecność atomów fluoru w obu jonach, atomów bromu w anionie, a także długie łańcuchy alkilowe i grupy funkcyjne karbonylowe i estrowe w kationie zwiększają rozpuszczalność ditlenku węgla. Obniżenie rozpuszczalności związane jest z obecnością w kationie grup metylowych, eterowych, hydroksylowych, alkinowych i nitrylowych. W dużo rzadszych przypadkach chemisorpcji, za zwiększenie rozpuszczalności odpowiada obecność anionów karboksylowych lub aminowych.[72]

W przypadku węglowodorów, największy wpływ na rozpuszczalność ma rodzaj anionu, ale także rodzaj węglowodoru. Co ciekawe, zależność jest bardzo podobna do opisywanej już zależności długości łańcucha alkilowego – im więcej atomów węgla posiada węglowódor, tym lepiej rozpuszcza się w cieczach jonowych. Wyższą rozpuszczalnością od swoich analogów charakteryzują się również węglowodory nienasycone oraz aromatyczne.[74] W przypadku kationów, pierścień pirolidynowy zapewnia większą rozpuszczalność węglowodorów o pierścienia imidazolowego.[75]

Tlen, azot i tlenek węgla rozpuszczają się w cieczach jonowych bardzo słabo. Aby zwiększyć ich rozpuszczalność, można zastosować dwa niezbyt zaskakujące rozwiązania – w cieczy jonowej powinien pojawić się perfluorowany anion lub kation z długim łańcuchem alkilowym.[72]

#### g) toksyczność

Istnieje kilka czynników strukturalnych, które wpływają na toksyczność cieczy jonowych. Dla kationu wzrost toksyczności związany jest z występowaniem pierścieni aromatycznych, heteroatomów w związkach cyklicznych, długich łańcuchów alkilowych, dużej ilości podstawników i wzrostem hydrofobowości.[76–78] Toksyczność maleje w przypadku cieczy opartych na kationach morfolinowych, cholinowych, sulfoniowych, a także w przypadku występowania podstawników estrowych i hydroksylowych. Toksyczność można obniżyć poprzez obecność podstawników hydrofilowych. Jest ona również niższa w przypadku cieczy opartych na aminokwasach.[79] Kolejną grupą cieczy jonowych, które uważane są za mało toksyczne, są ciecze zbudowane na bazie choliny lub betainy, często w połączeniu z aminokwasami. Są to tzw. ciecze trzeciej generacji, znane również jako biokompatybilne.[80] W przypadku anionów toksyczność zauważalnie wzrasta, jeśli ciecz zawiera atomy fluoru.[78] Obniżenie toksyczności uzyskać można stosując organiczne aniony, takie jak np. karboksylany.[80,81]

## 1.2. Modelowanie QSPR

Odpowiedzią na problem ograniczonej liczby danych eksperymentalnych wynikający z niewielkiej różnorodności komercyjnie dostępnych komercyjnie cieczy jonowych oraz wysokiego kosztu ich syntezy jest zastosowanie modelowania matematycznego QSPR (*ang. Quantitative Structure – Property Relationship*), metody pozwalającej na przewidywanie właściwości fizykochemicznych nieprzebadanych związków chemicznych na podstawie istniejących danych eksperymentalnych. Wykorzystanie modelowania jest uważane za sprawdzoną, przyjazną środowisku alternatywę dla przeprowadzania badań eksperymentalnych, dzięki której udaje się zaoszczędzić zarówno czas, jak i pieniądze.[82] Modelowanie QSPR opiera się na założeniu, że różnice we właściwościach fizykochemicznych substancji mają swoje odzworowanie w różnicach w ich strukturach molekularnych. Proces zakłada, że możliwe jest zidentyfikowanie kluczowych elementów struktury i zbudowanie modelu matematycznego, który będzie w stanie przewidzieć właściwości fizykochemiczne związków, które nie zostały przebadane.[83] Aby „przetłumaczyć” cząsteczkę z reprezentacji chemicznej na reprezentację matematyczną stosowane są deskryptory molekularne, czyli obszerny zestaw predefiniowanych zmiennych służący do opisu cech cząsteczki w postaci zestawu wartości liczbowych. Deskryptory mogą dotyczyć wielu właściwości, nie tylko cech czysto strukturalnych, takich jak układ i wzajemne położenie atomów czy rodzaje i rozkład wiązań chemicznych w cząsteczce, ale także opisywać mogą bardziej skomplikowane relacje i wpływ np. elektroujemności poszczególnych atomów w cząsteczce w odniesieniu do ich relacji przestrzennej. Co więcej, deskryptorami mogą być również inne od analizowanej właściwości fizykochemiczne związków, co pozwala wykryć ich wzajemne powiązania.[84] Na rynku od dawna istnieje szereg gotowych rozwiązań i pakietów oprogramowania, które pozwalają na przeprowadzenie modelowania QSPR (alvaModel [85], DRAGON [86], ModelBuilder [87]), a wraz z rozwojem popularności języka programowania Python oraz generatywnych modeli AI wspomagających programowanie, niesłabnącym zainteresowaniem cieszą się autorskie programy do modelowania, często udostępnione w ramach licencji *open source* (Orange Data Mining).[88]

Podstawowe założenia modelowania QSPR zostały opracowane przez Organizację Współpracy Gospodarczej i Rozwoju (*ang. Organisation for Economic Co-operation and Development, OECD*) i pozostają niezmiennie.[89] Wyniki eksperymentalne należy odpowiednio przygotować i przekształcić je w zbiór danych wejściowych uzupełniony o obliczone deskryptory molekularne, następnie należy zbudować model matematyczny, poddać go walidacji statystycznej, a na koniec dokonać interpretacji wyników. Każdy z tych etapów modelowania można przeprowadzić na wiele sposobów, a w środowisku naukowym nie ma i najprawdopodobniej nigdy nie będzie konsensusu co do tego, która z metod zapewnia najlepsze wyniki i najbardziej predyktywne modele. Powinno się jednak podejmować próby

odnalezienia najbardziej optymalnych metod modelowania, zapewniających odpowiedni balans pomiędzy skomplikowaniem obliczeń a czasochłonnością przygotowania danych.[90]

#### 1.2.1. Wytyczne OECD

Każde badanie QSPR powinno zostać przeprowadzone zgodnie z wytycznymi przygotowanymi przez Organizację Współpracy Gospodarczej i Rozwoju (*ang. Organisation for Economic Co-operation and Development, OECD*). Według nich, każdy zbudowany model powinien posiadać:[91]

- zdefiniowany punkt końcowy,
- jednoznaczny algorytm,
- określoną dziedzinę stosowalności,
- odpowiednią walidację,
- interpretację wyników, jeśli to możliwe.

Model, który jest w rzeczywistości matematycznym wzorem, musi posiadać zdefiniowany punkt końcowy oraz jednoznaczny algorytm według którego obliczane są wyniki. Co więcej, nie powinien stanowić „czarnej skrzynki”, która nie zapewnia wiedzy i kontroli nad tym, na jakiej zasadzie obliczane są wyniki predykcji. Najprostszym sposobem spełnienia obu tych warunków jest skorzystanie z regresji liniowej, jednak metody nieliniowe również zapewniają możliwości chemicznej analizy algorytmu.

Dziedzina stosowalności również musi zostać określona, gdyż każdy model jest w stanie z odpowiednią dokładnością przewidzieć właściwości jedynie tych nieprzebadanych substancji, które są w pewnym stopniu podobne do związków chemicznych wykorzystanych do jego budowy. Wartości predykcji dla związków zbyt odbiegających od zbioru danych obliczane są w przeważającej liczbie przypadków niedokładnie i nie mogą być traktowane z ufnością.

Aby ocenić zgodność wyników doświadczalnych z przewidywanymi przez model, a także jego jakość i zdolności predykcyjne, niezbędna jest walidacja. Dopasowanie danych eksperymentalnych do obliczonych określane jest przez współczynnik determinacji  $R^2$ . Testy statystyczne służą weryfikacji rzetelności modelu, zarówno w odniesieniu do danych użytych do jego zbudowania (walidacja wewnętrzna), jak i względem nowych danych (walidacja zewnętrzna).[92]

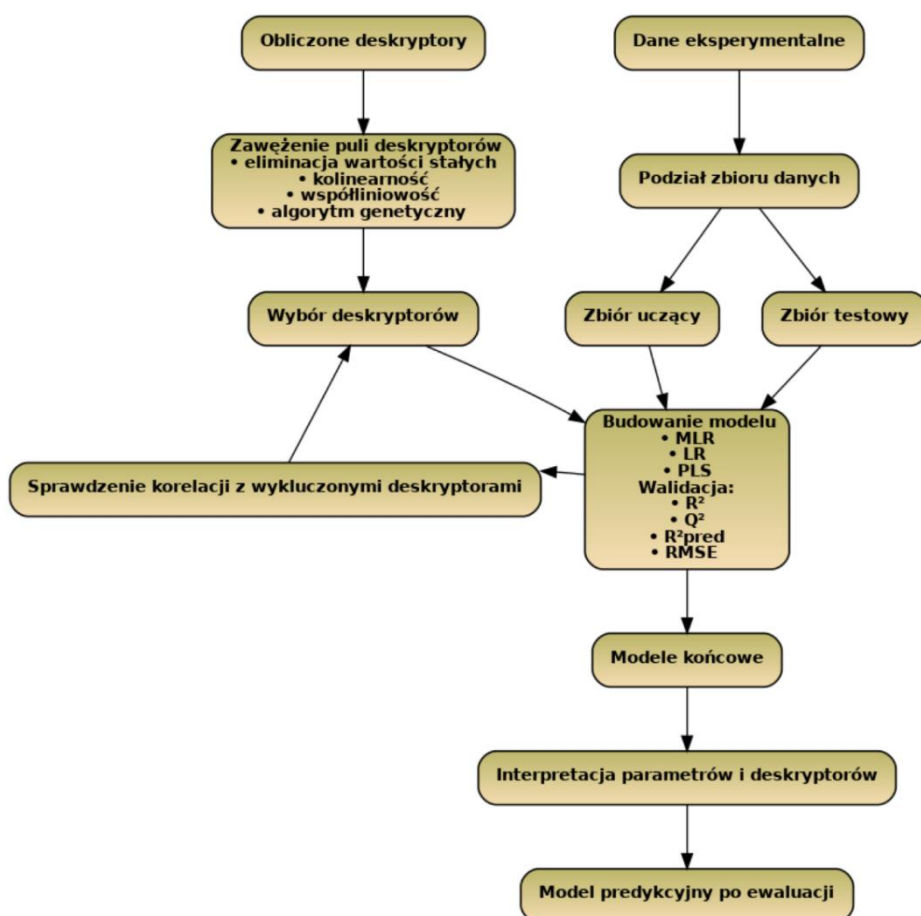
Ostatnim krokiem, jaki zgodnie z wytycznymi OECD powinien zostać przeprowadzony, jest interpretacja modelu. Analiza obliczonych deskryptorów, tego jak wpływają na algorytm modelu oraz od czego zależą ich wartości, jest kluczowa dla lepszego zrozumienia na poziomie molekularnym mechanizmów, od których zależą makroskopowe właściwości związków. Bardzo często jednak autorzy nie przywiązują do interpretacji deskryptorów większej wagi, ograniczając się jedynie do podania krótkich definicji deskryptorów dostępnych w literaturze, z których niewiele w istocie wynika.[93]

### 1.2.2. Etapy modelowania

Każde przeprowadzone modelowanie QSPR składa się z pięciu podstawowych etapów:[82]

1. Przygotowania zestawu danych do modelowania poprzez zebranie wyników pomiarów eksperymentalnych, ich weryfikację i ustandaryzowanie
2. Optymalizacja struktur cieczy jonowych wybraną metodą optymalizacji z uwzględnieniem środowiska (otoczenia) chemicznego lub bez
3. Obliczenie deskryptorów molekularnych oraz ich standaryzacja, usunięcie duplikatów oraz wartości skorelowanych
4. Modelowanie matematyczne oraz walidacja, prowadzące do otrzymania modeli o zadowalających parametrach statystycznych
5. Analiza otrzymanych modeli, obejmująca interpretację deskryptorów molekularnych w przełożeniu na strukturę cząsteczkową oraz z uwzględnieniem ich statystycznego wpływu na model

Schemat modelowania QSPR przedstawiono na Rysunku 1.4.



Rysunek 1. 4. Schemat procedury modelowania QSPR

Każdy z etapów został omówiony w rozdziale trzecim – Metody badawcze.

### 1.2.3. Metody optymalizacji struktury

#### a) metody optymalizacji *ab initio*

Metody te nie są oparte na uproszczeniach czy eksperymentalnie wyznaczonych współczynnikach w równaniach regresyjnych, stąd ich nazwa, tłumaczona jako „od początku”. Dzięki temu znajdują zastosowanie dla szerokiego zakresu związków i w odróżnieniu od metod pół-empirycznych nie tracą na dokładności w przypadku zastosowania ich dla zestawu danych, dla którego nie wyznaczono wcześniej żadnych współczynników. Jest to najwyższy, kwantowo-chemiczny poziom optymalizacji. Do najczęściej wykorzystywanych metod optymalizacji należą metoda Hartree-Focka (HF) oraz Møllera-Plesseta (MP), która jest jej rozwinięciem i uwzględnia korelację ruchu elektronów.[94] Obliczenia MP2 (czyli wyższy poziom obliczeń w rachunku zaburzeń uwzględniający oddziaływania międzyelektronowe [94]) są znacznie dłuższe i kosztowniejsze od HF, wymagają też większej bazy funkcyjnej (*ang. basis set*), takiej jak np. 6-31G\*.[95] Obecnie rzadko stosuje się czyste obliczenia MP, dużo większą popularnością cieszą się bazujące na nich nowsze metody, jak np. MP2-F12 zawierająca poprawki funkcji falowej uwzględniające odległości pomiędzy elektronami w sposób jawny.[96]

W metodach *ab initio* bardzo ważny jest wybór odpowiedniej bazy funkcyjnej, która zawiera informacje opisujące orbitale elektronowe.[97] Baza może opierać się na orbitalach Slatera (*ang. Slater Type Orbitals*, STO), które są uważane za dokładniejsze, ponieważ lepiej opisują orbital elektronu wodoru, czyli jedyny, dla którego znaleziono analityczne rozwiązanie równania Schrödingera.[98] Obliczenia oparte na STO są długotrwałe i z racji tego przeważnie wykorzystywane są do optymalizacji małych cząsteczek. Zamiast tego stosuje się orbitale Gaussowskie (*ang. Gaussain Type Orbital*, GTO), które, jak udowodniono, oferują zbliżoną dokładność oraz dużo krótszy czas obliczeń.[97]

Istnieją różne bazy funkcyjne, które różnią się precyzją i czasem obliczeń. Do najbardziej znanych należą STO-3G, która wyszła już właściwie z użycia, a także 3-21G i 6-31 G.[97] 3-21G to baza, w której orbitale elektronowe podzielone są na trzy części: orbital niewalencyjny, opisywany przez trzy funkcje gaussowskie, orbital walencyjny wewnętrzny, opisywany przez dwie funkcje oraz walencyjny zewnętrzny opisywany przez jedną.[94] Baza ta dobrze opisuje proste atomy o liczbie atomowej mniejszej od liczby atomowej neonu, jednak dla większych struktura cząsteczki może być błędna. Rozwiązaniem jest uwzględnienie funkcji polaryzacyjnych typu orbitalu d, co w nazewnictwie zawsze symbolizowane jest znakiem „\*” umieszczonym na końcu, np. 3-21G\*. Kolejna baza to 6-31G, co oznacza, że orbital niewalencyjny opisany jest przez 6 funkcji, orbital walencyjny wewnętrzny przez 3, a zewnętrzny przez jedną.[94] Czas obliczeń wymagany dla tej bazy jest pięciokrotnie dłuższy, niż w

przypadku 3-21G. W tym przypadku również uwzględnić można funkcje polaryzacyjne.[99] W przypadku, gdy elektrony walencyjne są znacznie oddalone od jądra atomowego (np. dla anionów), zaleca się uwzględnienie funkcji dyfuzyjnych oznaczanych w zapisie jako „+”, aby obliczenia dla odległych od jądra regionów były dokładniejsze.[100] Jeszcze większych baz funkcyjnych używa się w celu jeszcze dokładniejszego wyznaczenia energii układu, już po optymalizacji geometrii prostszą metodą.[94]

W przypadku badań QSPR optymalizacja struktury jest ostatnim wymaganym etapem obliczeń, gdyż jedynie otrzymane w jej wyniku współrzędne atomów są wykorzystywane do obliczenia deskryptorów molekularnych. Z tego powodu zwykle nie ma potrzeby używania bazy 6-31G\*, gdyż daje ona porównywalne wyniki, co szybsza i tańsza zestaw 3-21G\*. Baza 6-31G\* zapewnia dokładniejsze obliczenia związane z energią cząsteczki [94,95], lecz informacje te nie są w żadnym etapie wykorzystywane.

Bazy funkcyjne uwzględniające korelację elektronów, oznaczane jako cc-pVXZ opierają się na podobnych zasadach, dając w niektórych przypadkach dokładniejsze wyniki przy praktycznie takim samym czasie obliczeniowym.[97] Zawsze uwzględniają one polaryzację, natomiast funkcje dyfuzyjne oznacza się w ich opisie jako AUG.[94]

Metody *ab initio* często stosowane są do optymalizacji małych cząstek, gdyż czas i poziom skomplikowania obliczeń wzrasta niewspółmiernie wraz z rozmiarami cząsteczek.[94] Wspomnieć należy również, że większość metody *ab initio* ma trudności z optymalizacją struktury cząsteczek zawierających atomy fluoru, co może stanowić problem przy wykorzystaniu ich w przypadku badań nad cieczami jonowymi pod kątem rozpuszczalności CO<sub>2</sub>, gdyż ciecze te często zawierają perfluorowane podstawniki.[101]

#### b) metody optymalizacji pół-empiryczne

Pół-empiryczne metody optymalizacji stanowią uproszczenie metod *ab initio*. Obliczenia są szybsze, ponieważ przyjmuje się w nich wiele ogólnych założeń i uproszczeń, uwzględniając pomiary eksperymentalne oraz wyniki uzyskane na wyższym poziomie obliczeń, takim jak HF, MP2 albo DFT.[102] Bazy funkcyjne w metodach pół-empirycznych są z góry narzucone i oparte na orbitalach o wcześniej obliczonych parametrach.[103] W metodach tych obliczenia przeprowadza się tylko dla elektronów walencyjnych, a elektrony niewalencyjne traktowane są jak część atomu. Do najbardziej znanych należą metody AM1 oraz PM3( od 3 do 7). Metody PM stanowią udoskonalenie metody AM1, a każda kolejna ich wersja jest dostosowana do coraz szerszej gamy związków. Dla porównania, metoda PM3 jest oparta na danych dla 500 związków, a dla PM6 jest to już 9 tysięcy.[94] Obliczenia te są, w zależności od wybranej metody, od 100 do 1000 razy szybsze od metod *ab initio* oraz DFT.[104]

Dokładność obliczeń metod PM6 lub PM7 podczas optymalizacji struktury jest porównywalna z metodami *ab initio* w przypadku większości związków chemicznych. Starsze metody pół-empiryczne, takiej jak AM1 i PM3, charakteryzują się zdecydowanie większą niedokładnością.[94] W związku z tym, nie ma podstaw, aby wciąż je stosować, mając do wyboru nowsze, równie szybkie ulepszone metody.

Niestety, znacznym ograniczeniem zastosowania wszystkich metod pół-empirycznych, jest ograniczona możliwość optymalizacji geometrii związków o niecodziennej strukturze, których analogi nie były użyte jako zbiór uczący podczas opracowywania empirycznych współczynników funkcji bazy. Błędnie optymalizowane są chociażby jony niektórych związków organicznych, a także atomy hiperwalencyjne.[94,105] Metody pół-empiryczne gorzej radziły sobie również z optymalizacją struktur zawierających wiązania wodorowe, co zgodnie z doniesieniami literaturowymi zostało skorygowane w nowszych wersjach.[103] Doświadczenie autora z pracy z metodą PM7 wskazuje jednak na to, że obecność wiązań wodorowych wciąż może stanowić wyzwanie w przypadku obliczeń prowadzonych dla cieczy jonowych. Optymalizacja metodami pół-empirycznymi jest wskazana szczególnie w przypadku związków o dużej masie cząsteczkowej i zbudowanych z wielu atomów, takich jak na przykład białka.

#### c) metody optymalizacji DFT

W odróżnieniu od metod *ab initio* i metod pół-empirycznych, metody teorii funkcjonału gęstości, czyli DFT (*ang. Density Functional Theory*) nie są oparte na funkcji falowej, lecz na lokalnej gęstości elektronowej.[94] Dzięki temu rozwiązaniu równania rozwiązywane są w trójwymiarowym układzie współrzędnych, podczas gdy poprzednie metody charakteryzują się trójwymiarowym układem współrzędnych dla każdego z elektronów z osobna. Jest to niewątpliwa zaleta, gdyż następuje znaczna redukcja wymiarowości problemu.[94]

Metody DFT nie są uniwersalne i każda z nich sprawdza się na innym polu, zależnie od zastosowanego funkcjonału. Są one pewnego rodzaju pomostem łączącym metody *ab initio* z metodami pół-empirycznymi, ponieważ liczone są „od początku”, ale parametry funkcjonału gęstości ustalane są empirycznie. Korzystanie z nich wymaga doświadczenia, gdyż istnieje wiele różnych funkcjonałów oraz baz funkcyjnych i należy posiadać odpowiednią wiedzę, jaką kombinację dobrać do konkretnego zadania. Stosując metody DFT trzeba znaleźć również odpowiedni balans pomiędzy czasem, skomplikowaniem i kosztami obliczeń a dokładnością wyników.[106]

Do najbardziej znanych metod należy B3LYP/6-31G\*, która często jest stosowana przy optymalizacji struktury cieczy jonowych.[102] Funkcjonał ten nie bierze jednak pod uwagę działania sił dyspersyjnych Londona, przez co optymalizacja struktury nie zawsze daje rzetelne wyniki, zwłaszcza w przypadku związków niepolarnych. Dużo lepszą dokładność przy bardzo zbliżonym czasie obliczeń zapewniają nowsze metody, takie jak r<sup>2</sup>SCAN-3c lub B97M-V zawierające wbudowane poprawki na

oddziaływania dyspersyjne.[95] Znacznie większą dokładnością, zwłaszcza w obliczeniach energii układu, cechują się metody hybrydowe, w których część wymienna funkcjonału gęstości zawiera energię wymienną z metody HF, lub podwójnie hybrydowe, w których energia ta liczona jest metodą rachunku zaburzeń MP2. Niestety dzieje się to kosztem zwiększonego zapotrzebowania na moc obliczeniową. W połączeniu z bazami funkcyjnymi jakości *Triple Zeta* (TZ) lub *Quadruple Zeta* (QZ) metody te cechują się dokładnością bliską kompletnej bazy funkcyjnej (*CBS, complete basis set*). [95]

W kontekście optymalizacji struktury, niemal każdy nowszy funkcjonał w połączeniu z bazą funkcyjną o jakości TZ daje wyniki porównywalne dokładnością z metodami *ab initio* przy jednoczesnym krótszym czasie obliczeniowym. Mogą to być między innymi wspomniane  $r^2$ -SCAN-3c oraz B97-3c. Wyjątkiem jest optymalizacja związków, które posiadają zdelokalizowane elektrony w anionach lub zawierają w strukturze ciężkie pierwiastki. W takich przypadkach aby uzyskać prawidłową geometrię, należy skorzystać z bardziej wymagających metod hybrydowych.[95] Dla związków metaloorganicznych oraz dla metali przejściowych dobrze sprawdzają się funkcjonały rodziny M06, tzw. Minnesota Functionals. Z obliczaniem parametrów termodynamicznych najlepiej radzą sobie wspomniane już metody hybrydowe i podwójnie hybrydowe oparte na bazach typu Quadruple Zeta. Niestety, są one bardzo czasochłonne i kosztowne.[95]

#### 1.2.4. Deskryptory molekularne

Deskryptory molekularne stanowią ogniwo łączące chemię z matematyką. Są one matematycznym opisem chemicznej struktury cząsteczki o ściśle określonym sposobie obliczania. Są one niezbędne w celu przeprowadzenia badania QSPR, które również stanowi matematyczną próbę analizy właściwości chemicznych. Istnieje ponad pięć tysięcy różnorodnych deskryptorów, a co roku publikowane są artykuły poszerzające zbiór deskryptorów o kolejne, coraz to bardziej szczegółowe.[85]

Podstawowy podział deskryptorów to podział na deskryptory dwuwymiarowe, które obliczyć można wykorzystując np. wzory SMILES oraz deskryptory trójwymiarowe, które do obliczenia wymagają wcześniejszej optymalizacji geometrii cząsteczki. Kategorie, na które podzielić można deskryptory, wymienione są w Tabeli A1 w Materiałach uzupełniających.

Deskryptory są interdyscyplinarne - źródłem danych do obliczenia deskryptora może być m.in. chemia organiczna, chemia kwantowa, teoria grafów, chemia fizyczna, topologia czy matematyka dyskretna. Deskryptory mogą przedstawić cząsteczkę na wiele różnorodnych sposobów, zaczynając od nieskomplikowanych, zliczających liczbę i typy wiązań, obecność konkretnych atomów lub ugrupowań, poprzez wartości wektorów własnych obliczonych na podstawie macierzy topologiczno-geometrycznych, na skomplikowanych indeksach, które do obliczenia wymagają obliczenia kilku innych mniejszych deskryptorów. W deskryptorach zakodować można nie tylko informacje o przestrzennym ułożeniu

atomów, ale także szereg ich właściwości fizykochemicznych, jak np. elektroujemność czy potencjał jonizacji.[107]

#### 1.2.5. Badania QSPR dotyczące cieczy jonowych

Modelowanie QSPR w celu przewidywania właściwości fizykochemicznych cieczy jonowych jest obecnie bardzo popularne w środowisku naukowym. W Tabeli 1.1 przedstawiono zestawienie wybranych publikacji, które ukazały się w ciągu ostatnich lat. Zamieszczono w niej informacje o modelowanej właściwości, rozmiarze zbioru danych oraz zastosowanych metodach modelowania.

| WŁAŚCIWOŚĆ                      | AUTORZY                 | ZBIÓR DANYCH                 | METODY                   | ŹRÓDŁO |
|---------------------------------|-------------------------|------------------------------|--------------------------|--------|
| <b>Gęstość</b>                  |                         |                              |                          |        |
|                                 | Keshavarz i inni        | 249 IL                       | MLR                      | [108]  |
|                                 | Qiao i inni             | 123 IL                       | MLR, GC                  | [109]  |
|                                 | Paduszyński             | 2267 IL                      | GC                       | [110]  |
|                                 | Zhang i inni            | 314 IL                       | f-T-P                    | [111]  |
| <b>Temperatura topnienia</b>    |                         |                              |                          |        |
|                                 | Cerecedo-Cordoba i inni | 415 IL                       | Clustering               | [112]  |
|                                 | Katritzky i inni        | 104 IL                       | MLR                      | [113]  |
|                                 | Varnek i inni           | 717 IL                       | ANN, SVM, PLS, MLR, k-NN | [114]  |
|                                 | Farahani i inni         | 705 IL                       | MLR                      | [115]  |
|                                 | Lazzus                  | 400 IL                       | MLR, GC                  | [116]  |
|                                 | Yan i inni              | 394 IL                       | MLR                      | [117]  |
|                                 | Venkatraman i inni      | 2212 IL                      | PLS, k-NN                | [118]  |
| <b>Lepkość</b>                  |                         |                              |                          |        |
|                                 | Paduszyński             | 2068 IL                      | GC                       | [119]  |
|                                 | Das i inni              | 291 IL                       | PLS                      | [120]  |
|                                 | Koi i inni              | 40 IL                        | MLR                      | [121]  |
|                                 | Yan i inni              | 349 IL                       | f-T-P                    | [122]  |
|                                 | Billard i inni          | 99 IL                        | ANN                      | [123]  |
|                                 | Paduszyński, Domańska   | 1484 IL                      | ANN                      | [124]  |
| <b>Współczynniki aktywności</b> |                         |                              |                          |        |
|                                 | Jeon i inni             | 195 IL, 147 rozpuszczalników | ANN                      | [125]  |
|                                 | Eike i inni             | 3 IL, 38 rozpuszczalników    | MLR                      | [126]  |
|                                 | Xi i inni               | 1 IL, 39 rozpuszczalników    | OLS                      | [127]  |
|                                 | Paduszyński             | 188 IL, 128 rozpuszczalników | MLR, ANN, SVM            | [128]  |
|                                 | Gonfa i inni            | 53 IL, 1 rozpuszczalnik      | MLR                      | [129]  |
| <b>Rozpuszczalność gazów</b>    |                         |                              |                          |        |
|                                 | Mehraein, Riahi         | 21 IL, CO <sub>2</sub>       | MLR, SVM                 | [130]  |

|                               |                        |                         |                  |       |
|-------------------------------|------------------------|-------------------------|------------------|-------|
|                               | Safamirzaei            | 1 IL, 7 gazów           | ANN              | [131] |
|                               | Kang i inni            | 28 IL, H <sub>2</sub> s | ELM              | [132] |
|                               | Aghaie i inni          | 46 IL, CO <sub>2</sub>  | SVM, RF, DT, MLR | [133] |
| <b>Współczynniki podziału</b> |                        |                         |                  |       |
|                               | Jia i inni             | 18 związków             | Norm-index       | [134] |
|                               | Zhu i inni             | 190 związków            | MLR, ANN, SVM    | [135] |
|                               | Katritzky i inni       | 60 związków             | MLR              | [136] |
|                               | Mutelet i inni         | 42 związków             | MLR              | [137] |
|                               | Khooshechin i inni     | 90 związków             | SVM              | [138] |
| <b>Stała Henry'ego</b>        |                        |                         |                  |       |
|                               | Ghaslani i inni        | 32 IL, CO <sub>2</sub>  | MLR, SVM         | [139] |
|                               | Kang i inni            | 34 IL, CO <sub>2</sub>  | MLR, SVM, ELM    | [140] |
|                               | Safamirzaei, Modarress | 1 IL, 7 gazów           | ANN              | [141] |

**Tabela 1. 1.** Wybrane publikacje dotyczące modelowania QSPR w celu przewidywania właściwości fizykochemicznych cieczy jonowych . Wyjaśnienie skrótów (ze względu na to, że nazwy metod w j. angielskim przyjęły się szeroko w Polsce, pozostawiono je bez tłumaczenia): OLS - Ordinary Least-square Regression, ANN - Artificial Neural Network, MLR - Multiple Linear Regression, SVM -Support Vector Machines, K-NN - k-Nearest Neighbors, GC – Group contribution, ELM - Extreme learning machine, RF - Random Forest, DT - Decision Tree, f-T-P - variation of property f with temperature T and pressure P]

Najczęściej stosowanymi metodami modelowania jest wielokrotna regresja liniowa (*ang. Multiple Linear Regression*, MLR) oraz nieliniowa sztuczna sieć neuronowa (*ang. Artificial Neural Network*, ANN).[142] Wybór dostępnego na rynku oprogramowania do QSPR jest szeroki i obejmuje programy takie jak CODESSA [113], COSMO-RS [143], DRAGON [86] czy alvaDesc [85] służące do obliczania deskryptorów oraz MATLAB [124,131], QSARINS [144], ModelBuilder [145], Orange Data Mining [88] i alvaModel [85] używane do modelowania. Oprócz tego wielu badaczy stosuje programy własnego autorstwa, najczęściej napisane w środowisku Python.[133] Do popularności tworzenia własnych skryptów, w ostatnich latach, przyczynił się także rozwój modeli AI, dzięki wsparciu których programowanie stało się bardziej przystępne dla chemików. Niezależnie od wyboru oprogramowania i metod, każdy z wymienionych w Tabeli 1.1 modeli został poddany odpowiedniej walidacji i wykazał odpowiednio wysokie parametry statystyczne. Jak można zauważyć, zastosowane metody tworzenia modeli są bardzo różnorodne i obecnie nie można wskazać metody dominującej. Zadowalający rezultat uzyskać można na wiele sposobów i zasadniczym celem badań w tym obszarze jest znalezienie najbardziej odpowiedniej dla posiadanego zestawu danych.

Uwagę należy zwrócić na znaczące rozbieżności w rozmiarach zbiorów danych. Dane eksperymentalne dotyczące podstawowych właściwości fizykochemicznych, takich jak gęstość, lepkość lub temperatura topnienia pozwalają na zbudowanie modeli QSPR opartych na kilkuset, a niekiedy kilku tysiącach cieczy jonowych (np. Paduszyński modelując gęstość i lepkość wykorzystał odpowiednio 2267

i 2068 cieczy [110,119]). Badacze, którzy chcą skupić się na rzadziej badanych właściwościach, ograniczeni są do kilkunastu lub w najlepszym przypadku kilkudziesięciu punktów pomiarowych. Im więcej związków znajduje się w zbiorze uczącym modelu, tym szerszy zakres zmienności w strukturze molekularnej jest on w stanie uwzględnić i z powodzeniem przewidzieć. Biorąc za przykład pomiary wartości stałej Henry'ego dla ditlenku węgla w cieczach jonowych, na których skupia się ta rozprawa, największy opublikowany przed rozpoczęciem badań model obejmował 34 ciecze.[139] Niniejsza praca obejmująca 73 związki jest największym do tej pory istniejącym zbiorem danych dotyczącym stałej Henry'ego. Nowe dane eksperymentalne pojawiają się co roku, często uwzględniając nieprzebadane wcześniej rodzaje cieczy jonowych. Przekłada się to na ciągle wzrastający potencjał do budowania modeli matematycznych o rozszerzonym zakresie stosowalności i coraz większej dokładności.

Niekiedy badania QSPR są poparte walidacją eksperymentalną, lecz nie jest to popularna praktyka. Venkatraman i inni zbudowali modele QSPR przewidujące indeksy refrakcji cieczy jonowych, na podstawie interpretacji wyników zsyntezowali 14 nowych cieczy jonowych i eksperymentalnie zweryfikowali przewidziane przez modele wartości uzyskując satysfakcjonujące wyniki.[146] W innej pracy, Gorji i inni wyznaczyli współczynniki podziału tiofenu dla układów ciecz jonowa/węglowodory z pomocą modelowania QSPR i zwalidowali wyniki wykonując pomiary eksperymentalne dla trzech wybranych układów trójskładnikowych.[144]

### **1.3. Rozpuszczalność gazów [147–150]**

Ze względu na poziom skomplikowania procesu rozpuszczania gazów w cieczach oraz na znaczne zróżnicowanie możliwych do opisu układów, bardzo trudnym okazuje się znalezienie uniwersalnych zasad rządzących rozpuszczalnością, co niekiedy zmusza naukowców do stosowania wielu założeń i uproszczeń. Niemniej jednak próby matematycznego opisu procesu rozpuszczania gazów w cieczach z powodzeniem podejmowane są od wielu lat. [151–153]

Zacząć należy od opisu dwóch czynników mających wpływ na rozpuszczalność gazów w cieczach – ciśnienia ( $P$ ) i temperatury ( $T$ ). Zależności te przedstawić można na przykładzie układu dwuskładnikowego i dwufazowego. Gdzie fazę ciekłą tworzy roztwór gazu w cieczy (ułamek molowy gazu oznaczony jako  $x_2$ ), natomiast faza gazowa zawiera wyłącznie rozpuszczany gaz (ułamek molowy oznaczony jako  $y_2$ ). Takie założenie jest możliwe w przypadku cieczy jonowych, dla których zakłada się pomijalnie małą prężność par. Warunkiem równowagi termodynamicznej jest równość potencjałów chemicznych ( $\mu$ ) gazu znajdującego się w fazie ciekłej i gazowej. Przy wcześniejszym założeniu, że prężność pary rozpuszczalnika jest znikoma w porównaniu do prężności pary substancji rozpuszczanej można przyjąć, że faza gazowa zawiera niemal wyłącznie substancję rozpuszczaną, a więc  $y_2$  równe jest 1. W konsekwencji uzyskuje się Równania 1.1 i 1.2:

$$\mu_2^g(P, T, y_2) = \mu_2^c(P, T, x_2) \quad (1.1)$$

$$\mu_2^*(P, T) = \mu_2^c(P, T, x_2) \quad (1.2)$$

gdzie  $\mu_2^g$  oznacza potencjał chemiczny gazu w fazie gazowej,  $\mu_2^c$  oznacza potencjał chemiczny gazu w fazie ciekłej, a  $\mu_2^*$  oznacza potencjał chemiczny czystego gazu.

Jeśli potencjały chemiczne składników układu w obu fazach są równe, równe są także ich różniczki zupełne, jak przedstawiono w Równaniu 1.3:

$$\left(\frac{\partial \mu_2^*}{\partial P}\right)_T dP + \left(\frac{\partial \mu_2^*}{\partial T}\right)_P dT = \left(\frac{\partial \mu_2^c}{\partial P}\right)_T dP + \left(\frac{\partial \mu_2^c}{\partial T}\right)_P dT + \left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T} dx_2 \quad (1.3)$$

W Równaniu 1.3 pojawiają się znane zależności termodynamiczne:

$$\left(\frac{\partial \mu_i}{\partial P}\right)_T = V_i, \quad (1.4)$$

$$\left(\frac{\partial \mu_i}{\partial T}\right)_P = -S_i \quad (1.5)$$

Równanie 1.4 opisuje objętość molową  $V_i$ , natomiast Równanie 1.5 entropię molową  $S_i$ . Równanie 1.3 można uprościć, więc do postaci Równania 1.6:

$$V_2^* dP - S_2^* dT = V_2^c dP - S_2^c dT + \left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T} dx_2 \quad (1.6)$$

gdzie  $V_2^*$  oznacza objętość molową czystego gazu,  $V_2^c$  oznacza objętość molową gazu w fazie ciekłej,  $S_2^*$  oznacza entropię molową czystego gazu, a  $S_2^c$  oznacza entropię molową gazu w fazie ciekłej.

Po uporządkowaniu i przekształceniu otrzymuje się Równanie 1.7:

$$(V_2^c - V_2^*) dP - (S_2^c - S_2^*) dT + \left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T} dx_2 = 0 \quad (1.7)$$

Równanie 1.7 można wykorzystać do wyprowadzenia zależności opisujących wpływ ciśnienia i temperatury na rozpuszczalność. Przy stałej temperaturze  $dT=0$ , równanie przyjmuje postać:

$$\left(\frac{\partial x_2}{\partial P}\right)_T = -\frac{V_2^c - V_2^*}{\left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T}} \quad (1.8)$$

Z analizy Równania 1.8 można wyciągnąć wnioski dotyczące zależności rozpuszczalności gazu od jego ciśnienia  $\left(\frac{\partial x_2}{\partial P}\right)_T$ . Potencjał chemiczny składnika jest rosnącą funkcją jego ułamka molowego, a więc znak mianownika wyrażenia po prawej stronie równania zawsze będzie dodatni. W warunkach umiarkowanego ciśnieniach objętość molowa gazu w fazie gazowej jest zawsze większa od jego cząstkowej objętości molowej w ciekłych roztworze, co oznacza ujemne wartości licznika w wyrażeniu

po prawej stronie równania. W konsekwencji można stwierdzić, że rozpuszczalność gazów rośnie wraz ze wzrostem ciśnienia.

Równanie to możemy dalej przekształcać, jeśli przyjmie się przybliżenia następujące założenia:

- 1)  $V_2^*$  jest znacznie większe od  $V_2^c$ , a  $V_2^* = RT/P$ , co jest słuszne w zakresie umiarkowanych ciśnień.
- 2)  $d\mu = RT d \ln x_2$ , co jest słuszne dla większości układów w zakresie obejmowanym przez prawa Raoult'a oraz Henry'ego, czyli przy  $x_2 < 0,1$  oraz  $x_2 > 0,9$ .

W powyższych wzorach R oznacza stałą gazową wynoszącą 8,3145 J/molK.

Po podstawieniu tych wartości do wzoru otrzymuje się Równanie 1.9:

$$\left(\frac{\partial x_2}{\partial P}\right)_T = \frac{RT/P}{RT/x_2} = \frac{x_2}{P} \quad (1.9)$$

Stąd:  $d \ln P = d \ln x_2$ , czyli  $\ln P = \ln x_2 + \ln C$

C jest stałą całkowania zależną od temperatury, natomiast P równe jest cząstkowemu ciśnieniu gazu nad roztworem. Równanie to można zapisać więc jako  $P_2 = k_H(T)x_2$ , gdzie  $k_H$  oznacza stałą Henry'ego.

W ten sposób dochodzimy do prawa Henry'ego, które mówi o tym, że prężność cząstkowa  $P_2$  nad rozcieńczonym roztworem gazu (składnika 2) jest proporcjonalna do jego ułamka molowego  $x_2$  w roztworze. Jest to prawo graniczne, gdyż wyprowadza się je przyjmując założenia o doskonałości fazy ciekłej i gazowej oraz o niezależności potencjału chemicznego składników fazy ciekłej od ciśnienia. Prawo Henry'ego spełnione jest dla bardzo rozcieńczonych roztworów pod umiarkowanymi ciśnieniami. Występująca w tym równaniu stała  $k_H$  jest jednym ze sposobów wyrażania rozpuszczalności.

To samo Równanie 1.9 można wykorzystać do opisu wpływu temperatury na rozpuszczalność. Zakładając stałe ciśnienie  $dP=0$ , równanie to przyjmuje postać:

$$\left(\frac{\partial x_2}{\partial T}\right)_P = \frac{S_2^c - S_2^*}{\left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T}} \quad (1.10)$$

Licznik w Równaniu 1.10,  $S_2^c - S_2^*$  jest to entropia przeniesienia jednego mola gazu z fazy gazowej do fazy ciekłej roztworu nasyconego. Ponieważ, jak już wspomniano, w warunkach równowagi termodynamicznej potencjały chemiczne składników układu muszą być równe, można sformułować Równanie 1.11:

$$H_2^c - TS_2^c = H_2^* - TS_2^* \quad (1.11)$$

gdzie  $H_2^c$  oznacza cząstkową entalpię molową gazu w fazie ciekłej, a  $H_2^*$  oznacza entalpię molową czystego gazu. Po przekształceniu i podstawieniu Równania 1.11 do Równania 1.10 otrzyma się Równanie 1.12:

$$\left(\frac{\partial x_2}{\partial T}\right)_P = \frac{\Delta H_{2, \text{nasyc}}}{T \left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T}} \quad (1.12)$$

gdzie  $\Delta H_{2, \text{nasyc}}$  jest to entalpia molowa cząstkowa rozpuszczania gazu w roztworze nasyconym, definiowana jako ciepło rozpuszczenia jednego mola gazu w nieskończonej objętości roztworu nasyconego. Ponieważ, jak już wspomniano, potencjał chemiczny  $\left(\frac{\partial \mu_2^c}{\partial x_2}\right)_{P,T}$  jest dodatni, to znak pochodnej  $\left(\frac{\partial x_2}{\partial T}\right)_P$  zależy od wartości entalpii  $\Delta H_{2, \text{nasyc}}$ . W przypadku większości gazów, rozpuszczanie jest procesem egzotermicznym, a więc entalpia jest ujemna, a rozpuszczalność gazów maleje wraz ze wzrostem temperatury.

W tym miejscu należałoby poświęcić kilka zdań, aby wyjaśnić od czego zależy, czy proces rozpuszczania gazu jest endo- czy egzotermiczny. Proces rozpuszczania zajdzie tylko, jeśli całkowita entalpia swobodna procesu ( $\Delta G$ ) będzie ujemna, co przedstawiono poniżej w Równaniu 1.13.

$$\Delta G = \Delta H - T\Delta S < 0 \quad (1.13)$$

gdzie  $\Delta H$  oznacza entalpię całkowitą procesu. W przypadku rozpuszczania gazu w cieczy entropia procesu jest przeważnie ujemna, ponieważ entropia fazy gazowej jest znacznie większa, niż entropia gazu rozpuszczonego w cieczy. Innymi słowy – podczas rozpuszczania gazów w cieczach ich uporządkowanie wzrasta. Zgodnie z Równaniem 1.14 na entalpię całkowitą  $\Delta H$  składają się trzy wartości:

$\Delta H_1$  - entalpia związana z zerwaniem oddziaływań wewnątrz substancji rozpuszczanej (w gazach równa jest 0)

$\Delta H_2$  - dodatnia entalpia związana z zerwaniem oddziaływań wewnątrz rozpuszczalnika (takich jak wiązania wodorowe lub oddziaływania dipol-dipol)

$\Delta H_3$  - ujemna entalpia związana z tworzeniem oddziaływań rozpuszczalnik – substancja rozpuszczona, czyli entalpia rozpuszczania

$$\Delta H = \Delta H_1 + \Delta H_2 + \Delta H_3 \quad (1.14)$$

W rozpuszczalnikach polarnych entalpia oddziaływań rozpuszczalnika z substancją rozpuszczoną  $\Delta H_3$  jest większa od entalpii oddziaływań wewnątrzcząsteczkowych rozpuszczalnika  $\Delta H_2$ , a więc proces jest egzotermiczny. W rozpuszczalnikach niepolarnych sytuacja nie jest tak prosta, ponieważ siła

oddziaływań  $\Delta H_2$  i  $\Delta H_3$  może być porównywalna lub nawet entalpia tworzenia oddziaływań substancja rozpuszczona - rozpuszczalnik może być mniejsza od oddziaływań wewnętrznych rozpuszczalnika.[154,155] Przykładem takiej sytuacji jest chociażby układ wodor-toluen, gdzie rozpuszczalność wodoru wzrasta w miarę ogrzewania roztworu.[156]

Aby odpowiednio opisać rozpuszczalność, stosuje się pojęcie rozpuszczalności doskonałej, dla której za stan odniesienia przyjęto roztwór doskonały.[147] Zgodnie z prawem Raoult'a rozpuszczalność doskonałą gazu definiuje się jako  $x_2^{id} = \frac{P_2}{P_2^*}$ , gdzie  $P_2$  to ciśnienie cząstkowe gazu nad roztworem, a  $P_2^*$  to prężność pary nasyconej nad czystym skroplonym gazem. W sensie fizycznym wartość ta ma sens jedynie w temperaturze poniżej temperatury krytycznej, kiedy w układzie współistnieje faza ciekła i gazowa. W przypadkach, kiedy chcemy wyznaczyć wartości rozpuszczalności doskonałej w temperaturze wyższej od temperatury krytycznej, wartość  $P_2^*$  jest ekstrapolowana. Stosowane jest równanie korelacyjne lub równanie Antoine'a, które jest uważane za dokładniejsze.[149] Przedstawiono je, odpowiednio, jako Równania 1.15 i 1.16:

$$\log\left(\frac{P_2^*}{P^0}\right) = a - \frac{b}{T} \quad (1.15)$$

$$\log\left(\frac{P_2^*}{P^0}\right) = a - \frac{b}{c + T} \quad (1.16)$$

gdzie a, b i c są współczynnikami dobieranymi empirycznie na podstawie istniejących danych dla temperatur poniżej temperatury krytycznej, a  $P^0$  to ciśnienie standardowe. Choć są to jedynie przybliżone wartości, rozpuszczalność doskonała obliczona na podstawie tych równań może być stosowana do szacowania rozpuszczalności gazów lub porównywania rozpuszczalności różnych gazów w tym samym rozpuszczalniku.

Do innych stosowanych metod opisu rozpuszczalności należą:[149]

- ułamek molowy gazu  $x_2$  w roztworze w temperaturze T oraz pod ciśnieniem cząstkowym gazu  $P_2$ ,
- rozpuszczalność Ostwalda, czyli stosunek objętości gazu rozpuszczonego do objętości rozpuszczalnika mierzonych w aktualnych warunkach ciśnienia i temperatury
- współczynnik Bunsena, czyli stosunek objętości rozpuszczonego gazu, zredukowanej do warunków normalnych, (273,15 K, 1,013 bar) do objętości rozpuszczalnika w danej temperaturze i w ciśnieniu 1,013 bar.
- współczynnik absorpcji, czyli stosunek objętości rozpuszczanego gazu, zredukowanej do warunków normalnych, do objętości rozpuszczalnika w danej temperaturze i ciśnieniu

całkowitym 1,013 bar, przy czym ciśnienie całkowite to suma prężności cząstkowych rozpuszczanego gazu i rozpuszczalnika.

- współczynnik Kuenena, czyli stosunek objętości rozpuszczanego gazu, zredukowanej do warunków normalnych, do masy rozpuszczalnika w danej temperaturze i pod ciśnieniem cząstkowym gazu równym 1,013 bar.

#### 1.4. Zielone rozpuszczalniki

Idealny zielony rozpuszczalnik powinien spełniać następujące kryteria:[157–160]

- nie wywierać negatywnego wpływu na środowisko na żadnym z etapów jego syntezy, stosowania lub utylizacji,
- być łatwy do zsyntezowania z substratów z odnawialnego źródła
- łagodzić warunki reakcji, które katalizuje, zmniejszając jednocześnie ich wymagania energetyczne,
- posiadać wiele zastosowań i nie być selektywny jedynie względem wąskiego spektrum substancji,
- zwiększać wydajność procesu, w którym jest stosowany,
- nie opuszczać środowiska reakcji, redukując konieczność jego uzupełniania,
- łatwo poddawać się regeneracji i ponownemu wykorzystaniu,
- być nietoksyczny i nieszkodliwy dla ludzi,
- być całkowicie biodegradowalny bez pozostawiania substancji reszkowych.

Niestety, w świetle obecnej wiedzy nie jest możliwe stworzenie rozpuszczalnika, który spełniałby wszystkie opisane wymagania. Mityczny, opisywany przez XVI-wiecznego alchemika Paracelsusa „alkalest” nie istnieje. Większość rozpuszczalników uznawanych za „zielone” jest zielona jedynie w pewnym, ściśle określonym kontekście i zastosowaniu.[161] Używane w przemyśle rozpuszczalniki takie jak toluen czy tetrahydrofuran są zastępowane mniej toksycznymi substancjami, jak na przykład niektóre estry, etera czy alkohole, ale nie oznacza to, że związki te należy uznać za „zielone”. Są one jedynie mniej szkodliwe od wcześniej używanych substancji.

Wiele rozpuszczalników uważanych za „zielone” i stosowanych obecnie to substancje, które były wcześniej znane, lecz odkryto dla nich nowe zastosowania zgodne z Zasadami Zielonej Chemii. Należą do nich m. in. ciecze jonowe, mieszaniny głęboko eutektyczne, ditlenek węgla w stanie nadkrytycznym, fluorowane rozpuszczalniki organiczne oraz woda.

#### a) ciecze jonowe

Ciecze jonowe znalazły zastosowanie nie tylko w instalacjach pilotażowych, ale także w tych pracujących na przemysłową skalę.[46] Należy jednak nadmienić, że takie przypadki są wciąż nieliczne. Ciecze jonowe charakteryzują się wysoko stabilnością elektrochemiczną, są niemal całkowicie niepalne i niepalne. Dodatkowo mogą być modyfikowane poprzez zmiany struktury kationu i anionu. Zastosowanie cieczy jonowych potrafi obniżyć zużycie energii i zwiększyć wydajność wielu procesów przemysłowych, jak np. destylacja ekstrakcyjna, gdzie zastępują glikol etylenowy.[162] Są również bardzo stabilne, dzięki czemu można stosować je w procesach przeprowadzanych w wysokiej temperaturze. W wielu przypadkach regeneracja cieczy jonowych ma bardzo wysoką wydajność, dzięki czemu można stosować je przez wiele cykli produkcyjnych.[163] Posiadają również zdolność pochłaniania ditlenku węgla, przez co można zastosować je do oczyszczania spalin. Z drugiej strony, regeneracja cieczy jonowych jest często skomplikowana i wymaga znacznych nakładów energii, a same ciecze są w większości przypadków toksyczne, a także drogie i słabo dostępne.[164,165]

#### b) mieszaniny głęboko eutektyczne

Mieszaniny głęboko eutektyczne składają się z dwóch lub więcej substancji, które w połączeniu tworzą ciecz o dużo niższej temperaturze topnienia, niż tworzące ją związki z osobna.[166] Ich zalety są porównywalne do cieczy jonowych – możliwość optymalizacji struktury chemicznej, obniżenie kosztów energetycznych procesu, w którym pełnią funkcję katalizatora i zwiększenie jego wydajności, możliwość łatwej regeneracji, a także stabilność termiczna. Niektóre z cieczy głęboko eutektycznych, jak na przykład ciecze oparte na czwartorzędowych solach amoniowych, są biodegradowalne.[167] Dodatkowymi zaletami jest ich stosunkowo łatwy proces produkcji oraz dostępność składników. Zgodnie z większością doniesień nie cechują się wysoką toksycznością, jednak możliwe są sytuacje, w których mieszanina głęboko eutektyczna okazywała się bardziej cytotoksyczna, niż czyste związki, z których została zbudowana.[168]

#### c) ditlenek węgla w stanie nadkrytycznym

CO<sub>2</sub> w stanie nadkrytycznym, czyli ditlenek węgla utrzymywany w warunkach powyżej jego parametrów krytycznych (304,128 K, 7,3773 MPa) jest uważany za zielony rozpuszczalnik zwłaszcza w związku z łatwością jego oczyszczania – po zmianie parametrów procesu CO<sub>2</sub> odparowuje pozostawiając produkty w oddzielnej fazie.[169] Użycie go jako rozpuszczalnika zwiększa wydajność procesów ekstrakcji przy jednoczesnym zmniejszeniu ich kosztów. CO<sub>2</sub> jest też stosunkowo tani. Niestety, odparowujący ditlenek węgla należy wychwycić, aby nie wydostał się do atmosfery, gdyż jest gazem cieplarnianym. Dodatkowo, cała aparatura musi operować w ściśle określonych warunkach

podwyższonego ciśnienia lub temperatury, co wymusza przeznaczenie znacznych nakładów energii na ich utrzymanie.[170]

#### d) fluorowane rozpuszczalniki organiczne

Podstawione fluorowanym lub perfluorowanym łańcuchem alkilowych rozpuszczalniki organiczne znane są od lat 50-tych XX wieku.[171] Do ich zalet należy rozpuszczanie związków organicznych w sposób selektywny i zależny do temperatury. Ich wykorzystanie w reakcjach katalitycznych upraszcza etap rozdzielania produktów końcowych od katalizatora. Efekt ten otrzymuje się poprzez ogrzanie układu początkowo niemieszających się faz, przeprowadzenie reakcji w fazie homogennej a następnie ochłodzenie układu w celu separacji wydzielenia rozpuszczalnika. Proces ten wymaga jednak znacznych nakładów energetycznych na ogrzewanie i chłodzenie.[172] Związki fluorowane są bardzo stabilne chemicznie i łatwe do regeneracji, lecz mają niskie temperatury wrzenia. Dodatkowo ze względu swoją stabilność, nie ulegają degradacji jeśli trafią do atmosfery, co powoduje ich bioakumulację.[173]

#### e) woda

Woda uważana jest za zielony rozpuszczalnik, gdyż jest nietoksyczna dla środowiska oraz łatwo dostępna i tania. W celu wykorzystania jej w procesie przemysłowym jako rozpuszczalnik, wymagane jest jednak jej oczyszczenie, co jest bardzo kosztowne. Dodatkowo, odparowanie wody wymaga bardzo dużych nakładów energii w porównaniu do rozpuszczalników organicznych. Co więcej, woda nie rozpuszcza znacznej części substancji organicznych, co zdecydowanie ogranicza jej zastosowanie.[174] Dodanie środków aktywnych powierzchniowo pozwala na poszerzenie spektrum jej wykorzystania, ale rozwiązanie to zwiększa koszty oczyszczania, ogranicza możliwość regeneracji, a także prowadzi do wprowadzenia do środowiska reakcji dodatkowych substancji, co stoi w sprzeczności z piątą zasadą Zielonej Chemii.[161,175]

„Zieloność” jest więc niestety kryterium niezwykle względnym. Każdy z omówionych rozpuszczalników ma swoje ograniczenia, a część badań skupia się na jednym z aspektów „zieloności”, ignorując pozostałe.[161] Często usprawnienie jednego z etapów procesu wiąże się z pogorszeniem warunków innego czy nawet koniecznością dodania dodatkowego etapu, którym zwykle jest ekstrakcja. Analizując potencjalne wykorzystanie nowego, „zielonego” rozpuszczalnika bardzo ważne jest wykonanie pełnej analizy cyklu życia. Pozwala ona na zidentyfikowanie wielu czynników związanych z użyciem nowego związku, które w ogólnym rozrachunku mogą sprawić, że rozpuszczalnik ten nie będzie bardziej „zielony” do swojego poprzednika.[176]

### 1.5. Dytlenek węgla jako gaz cieplarniany

Stopniowy wzrost średniej temperatury na Ziemi jest jednym z największych wyzwań, z jakimi w ostatnim czasie musi mierzyć się ludzkość.[177] Fenomen ten, znany jako globalne ocieplenie, spowodowany jest zwiększającym się stężeniem gazów cieplarnianych w atmosferze ziemskiej. Należą do nich między innymi tlenki azotu, tlenek węgla oraz metan, które odpowiedzialne są za wytworzenie w atmosferze bariery, która zatrzymuje ciepło emitowane z powierzchni planety, nie pozwalając mu na rozproszenie się w przestrzeni.[178] W związku z dynamicznym rozwojem przemysłu na przestrzeni ostatnich kilkudziesięciu lat, do atmosfery wyemitowano wystarczająco dużo gazów cieplarnianych, by zaburzyć istniejącą naturalną równowagę, przez co proces globalnego ocieplenia znacznie przybiera na intensywności.[179] Spośród wszystkich antropogenicznych zanieczyszczeń atmosfery, ditlenek węgla jest gazem cieplarnianym, który odpowiada za największy wzrost średniej temperatury - szacuje się, że jest on odpowiedzialny za 66% całkowitego efektu wymuszania radiacyjnego.[180,181]

79% całego ditlenku węgla w atmosferze bierze swój początek ze spalania paliw kopalnych, które mimo ciągłych starań rządów oraz organizacji politycznych i ekologicznych wciąż pozostają głównym źródłem energii na świecie. Stały wzrost światowej populacji oraz rozwój przemysłu sprawia, że zapotrzebowanie na energię stale wzrasta, przez co wzrasta również emisja CO<sub>2</sub> do środowiska, a jego naturalne absorbenty, czyli oceany, są w stanie pochłonąć jedynie ok. 30% antropogenicznych emisji. Ograniczenie i wychwycenie pozostałych 70% jest już zadaniem ludzkości.[180]

Jednym ze stosowanych rozwiązań jest wychwytywanie CO<sub>2</sub> bezpośrednio ze spalin powstałych w procesach przemysłowych.[182] Najpowszechniejszym procesem jest absorpcyjne oczyszczanie spalin za pomocą wodnych roztworów amin. Jest to sprawdzone rozwiązanie, stosowane w przemyśle od lat 30-tych XX wieku, ale nie jest niestety pozbawione wad.[183] Powstały w wyniku absorpcji roztwór zawierający karbaminiany i wodorowęglany należy podgrzać, aby oczyścić go i zregenerować, w wyniku czego czysty CO<sub>2</sub> zostaje uwolniony i może zostać selektywnie wychwycony, a roztwór amin można wykorzystać ponownie.[184] Niestety proces ten wymaga dostarczenia znacznych ilości energii, aminy są silnie korrozyjne, a dodatkowo ze względu na swoją lotność, w przypadku najmniejszej nawet nieszczelności, istnieje ryzyko uwolnienia się toksycznych związków do środowiska. Jednym z możliwych rozwiązań ograniczających wysokie koszty tego procesu jest wykorzystanie cieczy jonowych, które charakteryzują się niską lotnością, stabilnością termiczną, a także wykazują, w większości przypadków, wysoką rozpuszczalnością CO<sub>2</sub>. [165,185]

### 1.6. Wolna objętość w cieczach jonowych

Jak zostało już wcześniej wspomniane, ciecze jonowe mogą zawierać w swoich strukturach zróżnicowane łańcuchy alkilowe, podstawniki, grupy funkcyjne i heteroatomy, a często także pierścienie

aromatyczne. W wyniku tego strukturalnego zróżnicowania, w cieczy jonowej zachodzi wiele interakcji pomiędzy elementami strukturalnymi oddziałujących ze sobą jonów. Prowadzi to często do tworzenia się wolnych przestrzeni pomiędzy poszczególnymi cząsteczkami, na przykład z powodu wzajemnego odpychania się fragmentów obdarzonych ładunkiem lub zawady przestrzennej.[186] Rozmiary oraz częstotliwość występowania wspomnianych przestrzeni opisuje parametr nazywany wolną objętością (*ang. Free Volume, FV*). Definiowany jest jako różnica pomiędzy objętością molową zajmowaną przez cząsteczki cieczy, a ich objętością van der Waalsa pomnożoną przez 1.3.[187] Parametr ten może zostać znormalizowany poprzez podzielenie FV przez całkowitą objętość i jest opisywany w takim przypadku jako frakcyjna wolna objętość (*ang. Fractional Free Volume, FFV*).[188]

W cieczach jonowych wolna objętość może wpływać na ich zdolność do rozpuszczania gazów, gdyż wolne przestrzenie pomiędzy jonami mogą zostać wypełnione cząsteczkami gazu.[189] Z dotychczas przeprowadzonych badań wynika, że FFV jest powiązana z objętością molową cieczy, ale zależność ta jest zbyt złożona, by można było ją wykorzystać podczas obliczeń.[190] Mimo, że w większości przypadków to kationy mają większy wpływ na objętość cząsteczek cieczy, wiadomo również, że analiza zależności wolnej objętości od rozpuszczalności gazów musi uwzględniać wpływ anionów. Co więcej, stwierdzono, że kluczową interakcją wpływającą na rozpuszczalność jest zdolność łączenia mniejszych luk w strukturze cieczy co prowadzi do tworzenia większych niezajętych obszarów, które mogą zostać zajęte przez cząsteczki gazu.[189]

## 2. CELE I TEZY PRACY

Głównym celem badań opisanych w niniejszej pracy doktorskiej była dokładna analiza etapów modelowania QSPR oraz wykorzystanie go do przewidywania właściwości cieczy jonowych. W pracy zdecydowano się skupić na rozpuszczalności ditlenku węgla w cieczach jonowych, wyrażonym w postaci stałej Henry'ego.

Badania przeprowadzono w trzech etapach:

1. Etap pierwszy – modelowanie stałej Henry'ego
2. Etap drugi – analiza wpływu frakcyjnej wolnej objętości na modele
3. Etap trzeci – porównanie sposobów optymalizacji cząsteczek

W pierwszym etapie zbudowano trzy modele QSPR służące do przewidywania stałej Henry'ego ditlenku węgla w cieczach jonowych i przeanalizowano je pod kątem identyfikacji i interpretacji deskryptorów molekularnych posiadających największy wpływ na proces absorpcji CO<sub>2</sub>.

W drugim etapie skupiono się na frakcyjnej wolnej objętości w cieczach jonowych, czyli jednej z właściwości, którą zidentyfikowano w pierwszym etapie prac. Obliczono FFV dla 73 cieczy jonowych oraz skonstruowano modele QSPR aby zweryfikować, czy:

- a) uwzględnienie FFV jako zmiennej w modelowaniu poprawi parametry modeli
- b) FFV można przewidywać za pomocą modeli QSPR zamiast dynamiki molekularnej

W trzecim etapie skupiono się wpływie sposobów optymalizacji cząsteczek cieczy jonowych na wartości obliczonych deskryptorów molekularnych oraz na wyniki modelowania QSPR. W tym celu skonstruowano 32 modele QSPR uwzględniające:

- a) optymalizację jonów razem oraz osobno kation i anion
- b) optymalizację cząsteczek w próżni oraz w środowisku etanolu
- c) cztery metody modelowania

Dodatkowo na przykładzie cieczy jonowych z anionem [NTf<sub>2</sub>] przeprowadzono analizę różnic i odchyleń w obliczonych wartościach deskryptorów molekularnych w zależności od sposobu optymalizacji.

### 3. METODY BADAWCZE

Rozdział ten zawiera szczegółowe informacje dotyczące metod modelowania oraz oprogramowania wykorzystanego podczas prac nad rozprawą. Badania składały się z trzech głównych części, z których każda miała na celu zweryfikowanie przyporządkowanych im hipotez. Wszystkie trzy są ze sobą związane wspólnym zbiorem danych oraz metodami walidacji, co oznacza, że etapy modelowania QSPR polegające na zebraniu, obróbce i przygotowaniu danych opisany w rozdziałach 3.1 oraz 3.5 są identyczne dla całej pracy badawczej. W przypadku pozostałych etapów występowały rozbieżności wynikające z wykorzystania różnych metod optymalizacji struktur, podziału zbioru danych oraz modelowania, w związku z czym kolejne rozdziały, po opisie części wspólnej, zawierają osobne opisy prac wykonanych dla każdego z trzech głównych zadań badawczych.

#### 3.1. Przygotowanie danych

Dane eksperymentalne dotyczące pomiarów wartości stałej Henry'ego ditlenku węgla w cieczach jonowych  $[H_{CO_2}]$  zostały zebrane z artykułów dostępnych w bazie SCOPUS oraz uzupełnione o dane dostępne w w bazie ILThermo.[191,192] W przypadku pojawienia się dwóch lub więcej różnych wartości  $H_{CO_2}$  w różnych źródłach, wybierano wartość pojawiającą się najczęściej. W sytuacji, gdy raportowane wyniki badań znacznie od siebie odbiegały, artykuły źródłowe były analizowane i wybierano te, które pochodziły od zespołów badawczych z większym stażem naukowym, lub te, które przeprowadzone były w sposób bardziej zbliżonych do pozostałych badań pod kątem używanej aparatury oraz sposobu wyliczania wartości  $H_{CO_2}$ . Otrzymany zestaw danych zawierał jednak wartości pomiarów przeprowadzonych w szerokim zakresie temperatur. W celu wykluczenia wpływu tego parametru na planowane modele matematyczne, dane posortowano i przefiltrowano w taki sposób, aby otrzymać jak największy zbiór wartości uzyskanych w identycznych warunkach laboratoryjnych. Dzięki temu zabiegowi otrzymano zestawy 55 pomiarów w temperaturze 298 K oraz 39 pomiarów w temperaturze 303 K. Analiza porównawcza pokazała, że różnice w wartości  $H_{CO_2}$  w obu temperaturach są zanedbywalnie małe. W związku z tym oba zestawy zdecydowano się połączyć, jednak w przypadku posiadania danych pomiarowych w obu temperaturach, preferowane były dane uzyskane dla 298 K. W rezultacie otrzymano zbiór danych zawierający 73 unikatowe ciecze jonowe zbudowane z 39 kationów oraz 15 anionów. Cały zestaw danych przedstawiono poniżej w Tabeli 3.1.

| Kation    | Anion  | $H_{CO_2}$ [kPa] | FFV   | Źródło |
|-----------|--------|------------------|-------|--------|
| [BMIM]    | [TfO]  | 1428             | 0.193 | [193]  |
| [PentMIM] | [bFAP] | 2020             | 0.187 | [194]  |
| [HMIM]    | [pFAP] | 2160             | 0.182 | [194]  |

|                                                     |                       |      |       |       |
|-----------------------------------------------------|-----------------------|------|-------|-------|
| [HMIM]                                              | [eFAP]                | 2520 | 0.228 | [194] |
| [P66614]                                            | [N(CN) <sub>2</sub> ] | 2700 | 0.221 | [195] |
| [C <sub>8</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]                | 2730 | 0.204 | [194] |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]                | 2840 | 0.166 | [194] |
| [BMPyrr]                                            | [eFAP]                | 2850 | 0.19  | [143] |
| [PentMIM]                                           | [NTf2]                | 2880 | 0.18  | [196] |
| [DMIM]                                              | [BF4]                 | 2900 | 0.184 | [197] |
| [ETT]                                               | [eFAP]                | 2940 | 0.202 | [143] |
| [OMIM]                                              | [NTf2]                | 3000 | 0.186 | [198] |
| [P66614]                                            | [Cl]                  | 3040 | 0.23  | [199] |
| [HMPy]                                              | [NTf2]                | 3280 | 0.175 | [200] |
| [B3MPy]                                             | [NTf2]                | 3300 | 0.17  | [201] |
| [BMIM]                                              | [NTf2]                | 3300 | 0.184 | [194] |
| [EMIM]                                              | [eFAP]                | 3323 | 0.134 | [202] |
| [P66614]                                            | [NTf2]                | 3340 | 0.228 | [203] |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]                | 3370 | 0.17  | [204] |
| [HMIM]                                              | [NTf2]                | 3500 | 0.181 | [198] |
| [PMIM]                                              | [NTf2]                | 3700 | 0.208 | [201] |
| [EMIM]                                              | [NTf2]                | 3770 | 0.181 | [204] |
| [B4MPy]                                             | [NTf2]                | 3830 | 0.169 | [205] |
| [BMIM]                                              | [TA]                  | 3850 | 0.182 | [206] |
| [DMPIM]                                             | [NTf2]                | 3850 | 0.163 | [201] |
| [BMPyrr]                                            | [NTf2]                | 3860 | 0.159 | [74]  |
| [1,4BzMPY]                                          | [NTf2]                | 3930 | 0.151 | [207] |
| [EDMIM]                                             | [NTf2]                | 3960 | 0.205 | [201] |
| [1,3BzMPY]                                          | [NTf2]                | 3980 | 0.149 | [207] |
| [BzMIM]                                             | [NTf2]                | 3990 | 0.185 | [207] |
| [BzMPyrr]                                           | [NTf2]                | 4000 | 0.149 | [207] |
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | [NTf2]                | 4070 | 0.173 | [204] |
| [1,2BzMPY]                                          | [NTf2]                | 4170 | 0.145 | [207] |
| [C3CNMIM]                                           | [NTf2]                | 4210 | 0.2   | [208] |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]                | 4350 | 0.169 | [74]  |
| [HOEtMIM]                                           | [NTf2]                | 4420 | 0.19  | [209] |

|                                      |                       |       |       |       |
|--------------------------------------|-----------------------|-------|-------|-------|
| [BzPY]                               | [NTf2]                | 4610  | 0.146 | [207] |
| [C3CNDMIM]                           | [NTf2]                | 4620  | 0.156 | [208] |
| [OMIM]                               | [PF6]                 | 5140  | 0.204 | [205] |
| [PMIM]                               | [PF6]                 | 5200  | 0.192 | [198] |
| [(N <sub>111</sub> ) <sub>2</sub> N] | [C(CN) <sub>3</sub> ] | 5250  | 0.189 | [204] |
| [BMPyrr]                             | [TA]                  | 5320  | 0.173 | [210] |
| [BMIM]                               | [PF6]                 | 5340  | 0.191 | [194] |
| [OMIM]                               | [BF4]                 | 5390  | 0.204 | [208] |
| [BMIM]                               | [N(CN) <sub>2</sub> ] | 5590  | 0.178 | [211] |
| [BMIM]                               | [BF4]                 | 5600  | 0.191 | [201] |
| [P66614]                             | [Br]                  | 5880  | 0.232 | [195] |
| [BDMIM]                              | [BF4]                 | 6100  | 0.188 | [212] |
| [HOEtMIM]                            | [TfO]                 | 6140  | 0.169 | [209] |
| [BDMIM]                              | [PF6]                 | 6180  | 0.192 | [212] |
| [PhDBIM]                             | [NTf2]                | 6300  | 0.17  | [198] |
| [EMIM]                               | [TfO]                 | 7400  | 0.188 | [199] |
| [EMIM]                               | [BF4]                 | 8110  | 0.183 | [213] |
| [HEEA]                               | [OAc]                 | 8148  | 0.159 | [214] |
| [EMIM]                               | [N(CN) <sub>2</sub> ] | 8210  | 0.172 | [205] |
| [HOEtMIM]                            | [PF6]                 | 8940  | 0.171 | [209] |
| [THEA]                               | [Lac]                 | 8951  | 0.133 | [214] |
| [HEA]                                | [OAc]                 | 9500  | 0.152 | [215] |
| [HOEtMIM]                            | [BF4]                 | 10800 | 0.167 | [209] |
| [C3CNMIM]                            | [N(CN) <sub>2</sub> ] | 11060 | 0.163 | [208] |
| [HMIM]                               | [Cl]                  | 11800 | 0.184 | [205] |
| [HEA]                                | [HCOO]                | 13217 | 0.135 | [214] |
| [C3CNDMIM]                           | [N(CN) <sub>2</sub> ] | 13630 | 0.165 | [208] |
| [MHEA]                               | [HCOO]                | 13950 | 0.152 | [216] |
| [BHEA]                               | [OAc]                 | 14100 | 0.15  | [215] |
| [HEA]                                | [Lac]                 | 14200 | 0.133 | [215] |
| [HEEA]                               | [Lac]                 | 16214 | 0.153 | [214] |
| [BHEA]                               | [Lac]                 | 17700 | 0.135 | [215] |
| [PhBIM]                              | [NTf2]                | 18000 | 0.161 | [198] |

|          |        |       |       |       |
|----------|--------|-------|-------|-------|
| [THEA]   | [OAc]  | 18238 | 0.136 | [214] |
| [HHEMEA] | [Lac]  | 19500 | 0.139 | [215] |
| [HHEMEA] | [OAc]  | 20500 | 0.143 | [215] |
| [HEEA]   | [HCOO] | 21025 | 0.153 | [214] |

**Tabela 3. 1.** Zestawienie cieczy jonowych wykorzystanych w badaniach, wraz z ich zmierzonymi eksperymentalnie wartościami stałej Henry'ego dla ditlenku węgla, źródłami literaturowymi oraz z obliczonymi wartościami frakcyjnej wolnej objętości. Skrótowe nazwy cieczy rozwinięte są na początku rozprawy.

Frakcyjna wolna objętość (*ang. Fractional Free Volume, FFV*) została obliczona we współpracy z mgr. inż. Karolem Baranem oraz dr. hab. inż. Maciejem Śmiechowskim z Katedry Chemii Fizycznej z wykorzystaniem oprogramowania GROMACS 2018.5. Obliczone wartości FFV znajdują się powyżej w Tabeli 3.1.

### 3.2. Optymalizacja struktur i obliczenie deskryptorów molekularnych

Pierwszym etapem w procesie obliczania deskryptorów molekularnych jest przygotowanie ich trójwymiarowych struktur. Często stosowaną metodą jest wykorzystanie zapisów tekstowych SMILES jako danych wejściowych do programów służących do rysowania i wstępnej optymalizacji struktur 3D.[217] Rozwiązanie to jest szczególnie popularne w przypadku dużych zbiorów danych, gdzie ręczne przygotowanie struktur zajęłoby niewspółmiernie dużo czasu w porównaniu do pozostałych etapów modelowania. W przypadku opisywanych w rozprawie badań, które oparto na 54 jonach, podjęto decyzję o samodzielnym przygotowaniu wszystkich struktur w celu wyeliminowania ryzyka uzyskania cząsteczek o nieprawidłowej strukturze, a co za tym idzie, obliczenia nieprawidłowych deskryptorów.

#### 3.2.1. Modelowanie stałej Henry'ego

Każdy z 39 kationów i 15 anionów został narysowany w programie Avogadro [218] oraz poddany wstępnej optymalizacji struktury wykorzystując wbudowane pole siłowe MMFF94.[219] Tak przygotowane struktury zoptymalizowano używając oprogramowania GAUSSIAN 16.[220] W oparciu o wcześniejsze doświadczenia, podjęto decyzję o wyborze pół-empirycznej metody PM7, ze względu na jej liczne zalety, takie jak zweryfikowana statystycznie odpowiednia jakość wyników, a także szybkość działania algorytmu oraz brak błędów podczas obliczeń i konieczności ich powtarzania.[102] Mimo to, w trakcie weryfikacji danych zauważono, że niektóre z cieczy jonowych zawierające w swojej cząsteczce pierścienie imidazolowe przestają być rozpoznawalne jako związki aromatyczne przez oprogramowanie DRAGON 7 używane do obliczania deskryptorów molekularnych. Z powodu braku kompatybilności pomiędzy wspomnianymi programami, podczas konwersji plików wyjściowych .log z GAUSSIANA 16 do formatu .hin odczytywanego przez oprogramowanie DRAGON 7, konieczne było skorzystanie z dodatkowego oprogramowania HyperChem.[221] U podłoża problemu leżał inny sposób zapisywania informacji o aromatyczności i brak przekazywania ich podczas konwersji do pośredniego formatu plików

.mol, wymuszonego przez brak możliwości bezpośredniej konwersji. Problem ten udało się rozwiązać w prosty sposób poprzez edycję plików .hin, bez ingerencji w zoptymalizowane koordynaty poszczególnych atomów. Przypadek ten wyraźnie pokazuje konieczność weryfikacji zbioru danych na każdym etapie procesu modelowania QSPR, co jednak stanowić może ogromne wyzwanie w przypadku większych zbiorów.

Deskryptory molekularne zostały obliczone z wykorzystaniem wspomnianego już oprogramowania DRAGON 7. Wykorzystywany program jest w stanie obliczyć 5270 unikatowych deskryptorów z 30 kategorii dla każdej cząsteczki, co w przypadku cieczy jonowej składającej się z kationu i anionu przekłada się na 10540 deskryptorów dla pojedynczego związku i 769420 deskryptorów dla całego zbioru danych.[86,222] Aby ograniczyć liczbę niezależnych zmiennych podjęto szereg czynności:

- usunięto deskryptory o stałych wartościach dla każdej z cieczy jonowych,
- usunięto deskryptory o niemal stałych wartościach (90% lub więcej),
- usunięto deskryptory skorelowane ze sobą na poziomie powyżej 95%,
- usunięto deskryptory stanowiące liniową kombinację innych deskryptorów (współliniowość).

W rezultacie uzyskano redukcję liczby zmiennych o 94% - w zestawie danych wykorzystanym do modelowania QSPR pozostało 612 deskryptorów dla każdej z cieczy jonowych.

Aby wybrać najlepiej dopasowane deskryptory molekularne wykorzystano algorytm genetyczny. Algorytm genetyczny dokonuje wyboru optymalnego zestawu zmiennych na zasadzie darwinowskiego prawa selekcji naturalnej oraz mechanizmów ewolucyjnych.[223] Mimo, że jest to metoda oparta na losowości, algorytm jest skuteczną metodą uzyskiwania rzetelnych i wiarygodnych danych do modelowania, a jego efektywność rośnie z kolejnymi generacjami.[224]

W opisywanym badaniu ze względu na ograniczenia sprzętowe algorytm genetyczny postanowiono podzielić na cztery cykle zawierające ograniczoną liczbę deskryptorów. Każdy z czterech cykli algorytmu tworzył 1000 generacji deskryptorów z 10% stopniem mutacji, a cykle przeprowadzono osobno dla anionów i kationów. W wyniku obliczeń odrzucono 485 zmiennych, pozostawiając 157 najlepiej dopasowanych deskryptorów z wejściowej puli. Wyniki redukcji dla każdej z 30 kategorii przedstawiono poniżej w Tabeli A1 w Materiałach uzupełniających.

### 3.2.2. Analiza frakcyjnej wolnej objętości

Ciecze jonowe składają się z dwóch lub więcej niepołączonych ze sobą trwałymi wiązaniami jonów. Z tego powodu ich struktury można optymalizować osobno lub wspólnie, kation z anionem. W poprzednim badaniu zdecydowano się na optymalizację jonów osobno, przez co uzyskano te same

wartości deskryptorów 2D i 3D dla cieczy zawierających te same jony. Gdyby optymalizację przeprowadzono wspólnie, deskryptory 2D pozostałyby bez zmian, natomiast te z kategorii 3D różniłyby się dla każdej cieczy mimo posiadania tych samych jonów w swojej strukturze. W przypadku badań nad wolną objętością spodziewano się zdecydowanie większego wpływu deskryptorów trójwymiarowych na końcowe modele QSPR.[225] Z tego powodu zdecydowano się wspólną optymalizację obu jonów, aby uwzględnić interakcje pomiędzy nimi już na etapie optymalizacji i uniknąć dzięki temu przeoczenia i utraty teoretycznie istotnych informacji.

Wzory strukturalne cieczy jonowych przygotowano w oprogramowaniu Avogadro i wstępnie zoptymalizowano wykorzystując pole siłowe MMFF94. Następnie wykorzystując programowanie GAUSSIAN 16 jony zoptymalizowano wspólnie. W obliczeniach skorzystano z trzech dostępnych w programie metod optymalizacji – B3LYP/6-31G\*, B97-3c oraz PM7. Metody DFT okazały się jednak zdecydowanie mniej skuteczne od metody pół-empirycznej. W obu metodach atomy wodoru ulegały podczas obliczeń delokalizacji, tworząc niekiedy układ dwóch stabilnych związków chemicznych w miejscu jonów. Znacznie więcej obliczeń, w porównaniu do metody PM7, kończyło się również niepowodzeniem, w związku z czym zdecydowano się przeprowadzić optymalizację wykorzystując jedynie metodę PM7. Otrzymano 73 cieczy jonowe, dla których obliczono następnie deskryptory molekularne wykorzystując oprogramowanie DRAGON 7. Tak, jak w przypadku wcześniej omawianych badań nad stałą Henry’ego, wykonano szeregu operacji zawężających liczbę deskryptorów z wejściowych 10540. Po odrzuceniu wartości stałych i niemal stałych oraz po analizie korelacji i współliniowości, otrzymano 26 deskryptorów kationowych i 9 anionowych. Należy zauważyć, że w przypadku poprzednich badań, na tym etapie zbiór danych zawierał 612 deskryptorów. Wspólna optymalizacja kationów i anionów pozwoliła więc na redukcję liczby zmiennych na początku procesu modelowania o 94%. Ze względu na niewielki rozmiar zbioru danych, stwierdzono, że na tym etapie nie ma potrzeby przeprowadzenia redukcji z użyciem algorytmu genetycznego.

### 3.2.3. Analiza metod optymalizacji

Ciecze jonowe optymalizowane były dotychczas osobno lub wspólnie, natomiast zawsze proces ten przeprowadzany był w próżni. Optymalizacja ta nie uwzględniała wpływu otoczenia na interakcji pomiędzy jonami, a który może mieć miejsce w cieczach. Przebadanie i analiza interakcji wielu cząsteczek w danej objętości jest domeną dynamiki molekularnej, natomiast aby wziąć pod uwagę ten prawdopodobnie istotny czynnik również podczas optymalizacji geometrii, w literaturze zaproponowano optymalizację struktur związków chemicznych w określonym środowisku.[226] Dzięki narzuceniu konkretnej stałej dielektrycznej środowiska, jego wpływ może oddziaływać na interakcje sobą pomiędzy jonami. W przypadku cieczy jonowych zaproponowano w tym celu etanol, jako że

wartość jego stałej dielektrycznej jest w podobnym zakresie, co stała dielektryczna wielu dobrze przebadanych cieczy.[227]

W omawianym badaniu wykorzystano dwa dotychczasowe zestawy danych uzupełnione o warianty optymalizacji przeprowadzonej w środowisku etanolu. Wykorzystano w tym celu dane z poprzednich badań – wzory poszczególnych jonów oraz cieczy jonowych przygotowane w programie Avogadro i wstępnie zoptymalizowane metodą MMFF94. Następnie używając oprogramowania GAUSSIAN oraz metody PM7 cząsteczki zoptymalizowano w środowisku etanolu. W rezultacie otrzymano cztery zestawy zoptymalizowanych cieczy jonowych:

1. optymalizacja jonów osobno w próżni
2. optymalizacja jonów osobno w etanolu
3. optymalizacja jonów wspólnie w próżni
4. optymalizacja jonów wspólnie w etanolu

W tym momencie warto zauważyć, że zastąpienie optymalizacji w próżni środowiskiem o określonej stałej dielektrycznej miało na tym etapie zauważalny bardzo pozytywny wpływ na proces przygotowywania danych. Ciecze jonowe optymalizowane razem w próżni bardzo często kończyły proces jako dwa niezależne związki pozbawione ładunku ze względu na błędne przypisanie atomów wodoru pomiędzy jonami. Podczas optymalizacji w etanolu problem ten nie wystąpił ani razu. Optymalizacja zajmowała również mniej czasu oraz rzadziej kończyła się niepowodzeniem i nie wymagała ponownego uruchamiania od punktów kontrolnych.

W oparciu o otrzymane dane obliczono deskryptory molekularne z wykorzystaniem oprogramowania DRAGON 7, a następnie zawężono ich liczbę poprzez odrzucenie wartości stałych, niemal stałych (powyżej 90%), a także analizę korelacji oraz współliniowości. W rezultacie otrzymano zestawy danych zawierające odpowiednio:

1. 40 deskryptorów dla jonów optymalizowanych osobno w próżni
2. 36 deskryptorów dla jonów optymalizowanych osobno w etanolu
3. 36 deskryptorów dla jonów optymalizowanych wspólnie w próżni
4. 57 deskryptorów dla jonów optymalizowanych wspólnie w etanolu

### **3.3. Podział zbioru danych**

Kolejnym krokiem w procesie modelowania był podział zbioru danych na dwa podzbiory, z których jeden, znany jako zbiór uczący, posłużyć miał do zbudowania modelu, a drugi, zbiór testowy, do jego walidacji. Podziału zbioru jest jednym z najważniejszych etapów w procesie modelowania QSPR, gdyż

odpowiednio dobrana metoda może zaważyć na pozytywnych wynikach walidacji modelu, a przez to zwiększyć jego rzetelność i wiarygodność.[228] W środowisku naukowym nie ma jednak konsensusu co do uniwersalnego sposobu, w jaki należy dzielić zbiór danych, wiadomo jednak, że podział w sposób zorganizowany i przemyślany daje lepsze wyniki, niż podział losowy.[229]

Podziału na podzbiory można dokonać na podstawie zmiennej zależnej (Y) lub zmiennych niezależnych (X).[229] W przypadku tego pierwszego, związki przeważnie szeregowane są w kolejności rosnącej pod względem wartości analizowanej właściwości fizykochemicznej, a następnie dzielone w taki sposób, aby w zbiorze uczącym i testowym znalazły się związki o równomiernie rozłożonych wartościach zmiennej Y i aby uniknąć sytuacji, w której skrajne wartości znalazły się tylko w jednym ze zbiorów. Druga metoda podziału wykorzystuje algorytmy opierających się na analizie danych zmiennych niezależnych, czyli wartości obliczonych deskryptorów. W ten sposób zapewnia się podział zbioru danych w sposób, w którym w obu podzbiorach znajdują się związki o odpowiadających sobie cechach. Niewątpliwą zaletą tej metody jest uniknięcie sytuacji, w której w jednym ze zbiorów znajdują się związki nieposiadające reprezentacji w drugim, co w konsekwencji mogłoby się negatywnie przekładać na dziedzinę stosowalności modelu.[230]

#### 3.3.1. Modelowanie stałej Henry'ego

W opisywanym w rozprawie badaniu, ciecze jonowe zostały posegregowane od najwyższej wartości  $H_{CO_2}$  do najniższej. Po przeanalizowaniu powstałej w ten sposób listy zgodnie z przewidywaniami zauważono, że pewne makroskopowe cechy budowy cząsteczkowej cieczy jonowych, takie jak obecność atomów fluoru lub długość łańcucha alkilowego, korespondują z odpowiednio wyższymi zaobserwowanymi wartościami  $H_{CO_2}$ . W związku z tymi obserwacjami podjęto decyzję o podziale zbioru danych na podzbiory na podstawie wartości zmiennej Y w proporcjach 2:1 dla zbioru uczącego do zbioru testowego.[82] Ciecze jonowe uszeregowane według wzrastających wartości  $H_{CO_2}$  podzielono najpierw na grupy zawierające po trzy sąsiadujące ze sobą związki, a z każdej grupy wylosowano jeden z nich, który trafił do zbioru testowego. W ten sposób zapewniono odpowiednią reprezentację zróżnicowanego strukturalnie zbioru cieczy jonowych w obu podzbiorach. Szczegółowy podział ujęto w Tabeli A2 w Materiałach uzupełniających.

#### 3.3.2. Analiza frakcyjnej wolnej objętości

Podziału danych na zbiór uczący i testowy dokonano na podstawie zmiennych niezależnych X. Wykorzystano algorytm Kennarda-Stone'a oraz algorytm odległości euklidesowej.[229,231] Ze względu na mniejszą liczbę obserwacji odstających wszystkie kolejne obliczenia oraz modelowanie zostały przeprowadzone na zbiorach danych zawierających łącznie 65 cieczy jonowych i uzyskanych z wykorzystaniem algorytmu odległości euklidesowej z wykorzystaniem oprogramowania Dataset

Division 1.2 autorstwa dr Kunała Roya [232] (dostęp 12 września 2025). Informacje dotyczące szczegółowego podziału danych znajdują się w Tabeli A2 w Materiałach uzupełniających.

### 3.3.3. Analiza metod optymalizacji struktury cząsteczkowej

Badanie miało na celu analizę wpływu sposobów optymalizacji struktury jonów na modelowanie QSPR, więc za najistotniejszy czynnik uznano w nim różnice pomiędzy wartościami deskryptorów obliczonych na cztery różne sposoby tak jak to zdefiniowano w p. 3.2.3. Z tego powodu, podobnie jak w przypadku poprzedniego badania, zdecydowano się na podział zestawu danych na podstawie zmiennych niezależnych X. Do podziału wykorzystano jak poprzednio algorytm odległości euklidesowej.

## 3.4. Modelowanie QSPR

### 3.4.1. Modelowanie stałej Henry'ego

Deskryptory pozostałe po optymalizacji z zastosowaniem algorytmu genetycznego zostały następnie użyte w procesie modelowania QSPR wykonanego z użyciem oprogramowania ADMEWORKS ModelBuilder. Zbudowano trzy modele o akceptowalnych statystycznie parametrach wykorzystując wielokrotną regresję liniową [MLR], regresję logistyczną [LR] oraz metodę cząstkowych najmniejszych kwadratów [PLS]:

a) Wielokrotna regresja liniowa jest jedną z metod, które są najczęściej wykorzystywane w modelowaniu QSPR ze względu na łatwość jej zastosowania oraz późniejszej interpretacji.[142] Ogólne równanie modelu przedstawia Równanie 3.1:

$$Y_{MLR} = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (3.1)$$

gdzie  $Y_{MLR}$  oznacza wartość przewidywanej właściwości,  $X_1$ ,  $X_2$  oraz  $X_n$  są kolejnymi wartościami deskryptorów (zmiennie niezależne),  $a_0$  jest wyrazem stałym wyrażenia, natomiast  $a_1$ ,  $a_2$  oraz  $a_n$  są współczynnikami regresji odpowiadających im deskryptorów.

b) Regresja logistyczna jest oparta na nieliniowej funkcji sigmoidalnej i najczęściej wykorzystywana jest do rozwiązywania przypadków opartych na klasyfikatorach binarnych, gdyż jej wartość zawiera się pomiędzy dwoma ustalonymi ekstremami.[233] Model regresji logistycznej oparty jest na Równaniu 3.2 w następującej postaci:

$$Y_{LR} = \frac{1}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n)}} \quad (3.2)$$

gdzie  $Y_{LR}$  oznacza wartość przewidywanej właściwości,  $X_1$ ,  $X_2$  oraz  $X_n$  są kolejnymi wartościami deskryptorów,  $b_0$  jest wyrazem stałym wyrażenia, natomiast  $b_1$ ,  $b_2$  oraz  $b_n$  są współczynnikami regresji

odpowiadających im deskryptorów. W przypadku opisywanego badania za wartości ekstremów ustalono najniższą oraz najwyższą raportowaną wartość  $H_{CO_2}$ , a pozostałe wartości zostały dopasowane do liniowego fragmentu funkcji leżącej pomiędzy nimi na podstawie interpolacji wyznaczonej funkcji.

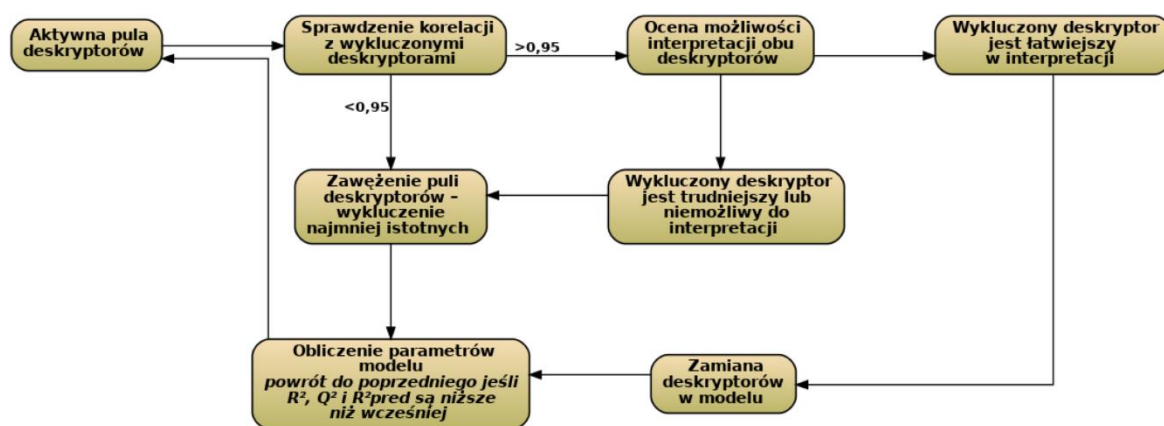
c) Metoda cząstkowych najmniejszych kwadratów jest podobna do wielokrotnej regresji liniowej, lecz lepiej radzi sobie w zbiorach danych zawierających silnie skorelowane ze sobą deskryptory oraz jest skuteczna w przypadku wartości ograniczonej liczebności zbioru danych eksperymentalnych. W metodzie tej zmienne zależne i niezależne zostają ustandaryzowane do wartości średniej równej 0 oraz wariancji równej 1, utworzony zostaje zbiór zmiennych utajonych, który zawiera informacje na temat zależności pomiędzy zmiennymi zależnymi i niezależnymi, a model zostaje zbudowany ich podstawie.[234]

Podczas badań budowano również model oparty o metodę maszyny wektorów nośnych (*ang. Support Vector Machine, SVM*), ale żadna z prób nie doprowadziła do otrzymania modelu, który przeszedłby pomyślnie narzuconą procedurę walidacji statystycznej. Zbiór danych został również zmniejszony o 11 cieczy jonowych, które zostały zidentyfikowane jako obserwacje odstające i wykluczone z dalszych obliczeń. Związki te wyszczególniono w Tabeli A2 w Materiałach uzupełniających.

Podczas procesu modelowania liczba deskryptorów była sukcesywnie zmniejszana poprzez odrzucanie najmniej istotnych statystycznie parametrów, zachowując jednocześnie akceptowalne parametry walidacji statystycznej. Na potrzeby tego badania do procesu wprowadzono dodatkowy, autorski algorytm. Zauważono, że znaczna część deskryptorów w zbiorze danych była ze sobą znacząco skorelowana, często ze współczynnikiem korelacji przekraczającym 0,95. Podczas redukcji zbioru działanie programu ModelBuilder polegało na losowym odrzuceniu jednego ze skorelowanych deskryptorów. W przypadku bardzo wysokiej korelacji nie ma statystycznego znaczenia dla jakości modelu, który z deskryptorów pozostanie w puli, tym nie mniej wpływ tej decyzji na ostatni z etapów modelowania QSPR, czyli interpretację wyników, jest znaczący. Niewykluczona jest sytuacja, w której istotne informacje zostają pominięte lub przeoczone ponieważ deskryptor z końcowego modelu okazuje się być niezwykle trudny do zinterpretowania, a silnie skorelowany z nim deskryptor, który jest łatwiejszy w interpretacji, został odrzucony na początkowym etapie redukcji zbioru. Z tego powodu, aby zapewnić odpowiednią jakość wniosków płynących z prowadzonych badań, autor zdecydował się na wprowadzenie modyfikacji do procesu modelowania, która miała zapewnić obecność interpretowalnych deskryptorów w końcowych modelach matematycznych, bez negatywnego wpływu na ich parametry statystyczne.

Po wstępnym odrzuceniu najmniej istotnych statystycznie deskryptorów, po każdym kolejnym zawężeniu puli pozostałe w niej istotne statystycznie deskryptory sprawdzano pod kątem skorelowania

z innymi deskryptorami odrzuconymi na etapie redukcji zbioru. Jeśli istniały deskryptory skorelowane z nimi współczynnikiem korelacji równym lub wyższym niż 0,95, dokonywano ich oceny pod kątem interpretowalności. Jeśli któryś z odrzuconych deskryptorów charakteryzował się lepszą możliwością wyciągnięcia istotnych wniosków, zbiór danych do modelowania zostawał zmodyfikowany – deskryptory zostawały zamieniane, a model był poddawany ponownej walidacji statystycznej. Jeśli parametry nie ulegały pogorszeniu, zmodyfikowany zbiór danych trafiał do kolejnego etapu modelowania. Jeśli jednak walidacja wskazywała na pogorszenie parametrów, zamiana deskryptorów zostawała cofnięta. W tym momencie warto wspomnieć, że parametry statystyczne w niektórych przypadkach ulegały poprawie. W ten sposób, zgodnie z założeniami, uzyskano modele o satysfakcjonujących parametrach statystycznych oraz charakteryzujące się interpretowalnymi deskryptorami. Dla lepszego zrozumienia procedury, jej schemat zamieszczono poniżej na Rysunku 3.1.



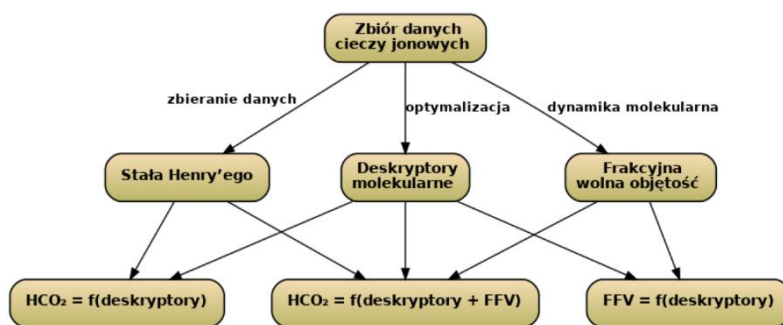
**Rysunek 3. 1.** Schemat postępowania w przypadku wystąpienia skorelowanych ze sobą deskryptorów

### 3.4.2. Analiza frakcyjnej wolnej objętości

Kierując się wynikami uzyskanymi z wcześniejszych badań, w modelowaniu zdecydowano się wykorzystać metodę wielokrotnej regresji liniowej, ze względu na akceptowalne parametry statystyczne zbudowanego wcześniej metodą MLR modelu oraz na łatwość interpretacji deskryptorów molekularnych.

Modelowanie QSPR przeprowadzono na zasadach opisanych w rozdziale trzecim. Zestaw deskryptorów zawężano usuwając z puli, w jednym kroku, jeden lub dwa deskryptory, z każdym krokiem dokonując jednocześnie walidacji. Aby uniknąć pominięcia istotnych informacji, za punkt końcowy uznano otrzymanie sześciu deskryptorów w modelu lub zauważalne pogorszenie parametrów statystycznych modelu. Drugim z warunków zakończenia modelowania była weryfikacja, czy każdy z pozostałych w modelu deskryptorów jest istotny statystycznie. Podczas tego etapu modelowania nie wykorzystano algorytmu mającego zapewnić otrzymanie w pełni interpretowalnych deskryptorów.

W pierwszym etapie zbudowano model, w którym przewidywaną właściwością była frakcyjna wolna objętość. Obliczenie FFV 73 cieczy jonowych z wykorzystaniem dynamiki molekularnej zajęło prawie 1,5 roku, więc celowym wydawało się sprawdzenie, czy na podstawie dokonanych już obliczeń możliwym jest stworzenie modelu QSPR, który z powodzeniem przewidzi wartości FFV dla nieprzebadanych dotąd cieczy. Model ten oparto jedynie na wynikach obliczeń MD oraz deskryptorach molekularnych. W drugiej kolejności zbudowano dwa modele przewidujące wartość  $H_{CO_2}$ , analogicznie jak w przypadku poprzednich badań. Pierwszy z modeli, nazywany w dalszej części rozprawy „modelem standardowym”, posłużyć miał do porównania wpływu sposobu optymalizacji cząsteczek cieczy jonowej (osobna lub wspólna optymalizacja jonów) na zdolność predykcyjną modelu. W drugim modelu, nazywanym „modelem FFV” jako jeden z parametrów wejściowych wykorzystano obliczone wartości wolnej objętości, dodając FFV jako dodatkową zmienną niezależną podczas modelowania. Na Rysunku 3.2 znajduje się graficzne przedstawienie procesu przygotowania danych oraz procesu modelowania.



**Rysunek 3. 2.** Graficzna reprezentacja procesu przygotowania danych oraz procesu modelowania

### 3.4.3. Analiza metod optymalizacji

Modelowanie QSPR przeprowadzono na uzyskanych czterech zbiorach danych, a dla każdego zbioru skonstruowano cztery modele z wykorzystaniem metod MLR, PLS, LR oraz SVM. Proces przeprowadzono wykorzystując jako zmienną Y wartość  $H_{CO_2}$  oraz wartość FFV. W rezultacie planowano otrzymać 32 modele QSPR. W rzeczywistości podczas prac niemożliwym okazało się uzyskanie satysfakcjonujących modeli LR dla  $H_{CO_2}$  oraz SVM dla FFV, a także modele MLR, PLS oraz LR dla FFV w przypadku cząsteczek optymalizowanych osobno w środowisku etanolu. W konsekwencji zbudowano 24 modele, natomiast analizie ostatecznie poddano 21 spośród nich.

W celu zminimalizowania wpływu sposobu modelowania na wyniki, każdy proces przeprowadzono w identyczny sposób, tak, aby parametry modeli różniły się jedynie ze względu na sposób optymalizacji. Modelowanie przeprowadzono niemal na identycznych zasadach, jakie opisane zostały w podpunkcie 3.4.2. Wyjątek stanowi przyjęcie za punkt końcowy modelowania jedynie momentu wyraźnego

pogorszenia się jakości parametrów statystycznych modelu – proces prowadzony był dalej po uzyskaniu sześciu deskryptorów w modelu. W wyniku tej decyzji w każdym z modeli służących do przewidywania  $H_{CO_2}$  znalazło się jedynie pięć deskryptorów molekularnych. Należy dodać również, że podczas modelowania  $H_{CO_2}$  w każdym z przypadków parametr FFV został dodany do pełnego zestawu deskryptorów molekularnych jako jedna ze zmiennych niezależnych.

### 3.5. Walidacja modeli

Współczynnik determinacji  $R^2$  jest podstawowym parametrem statystycznym, którym opisać można model. Określa on, jaka część zmienności w zbiorze danych została wyjaśniona przez model regresyjny. Im wyższa jest jego wartość, tym lepszy jest model.  $R^2$  obliczany jest za pomocą Równania 3.3:

$$R^2 = 1 - \frac{\sum(Y_{obs} - Y_{pred})^2}{\sum(Y_{obs} - \bar{Y}_{train})^2} \quad (3.3)$$

Mimo, że w symbolu współczynnika determinacji występuje wartość kwadratowa „<sup>2</sup>”, współczynnik ten może przyjmować również wartości ujemne w przypadku, gdy predykcje modelu nie wpisują się w trend obserwowanych danych lub gdy rzeczywista zależność pomiędzy nimi jest odwrotna. Prawidłowo zbudowany model powinien charakteryzować się wartością  $R^2$  jak najbardziej zbliżoną do 1.[142,235]

Ocenę dopasowania modelu przeprowadza się na zasadzie walidacji wewnętrznej. Dla małych zbiorów danych często wykorzystywaną metodą jest walidacja krzyżowa metodą *leave-one-out* [LOO]. W metodzie tej jeden ze związków usuwany jest ze zbioru danych, a model QSPR jest budowany ponownie wykorzystując wszystkie pozostałe związki. Deskryptory powiązane z usuniętym związkiem są następnie wykorzystywane, aby obliczyć przewidywaną właściwość, którą jest w tym przypadku wartość stałej Henry’ego. Wartość ta jest następnie porównywana z wynikiem eksperymentalnym, a proces powtarzany jest dla każdej cząsteczki. Suma obliczonych różnic wykorzystywana jest do obliczenia parametru  $Q^2$ , czyli współczynnika determinacji  $R^2$  dla walidacji krzyżowej na podstawie Równania 3.4. W prezentowanym badaniu wartość graniczną współczynnika  $Q^2$  decydującą o pomyślnym przejściu walidacji wewnętrznej ustalono jako 0,5.

$$Q^2 = 1 - \frac{\sum(Y_{obs} - Y_{pred(CV)})^2}{\sum(Y_{obs} - \bar{Y}_{train})^2} \quad (3.4)$$

W równaniach 3.3 oraz 3.4  $Y_{obs}$  oraz  $Y_{pred}$  odnoszą się odpowiednio do eksperymentalnej wartości obserwowanej zmiennej oraz jej wartości obliczonej na podstawie zbioru uczącego,  $Y_{pred(CV)}$  oznacza wartość obliczoną podczas walidacji krzyżowej, natomiast  $\bar{Y}_{train}$  oznacza średnią wartość obserwowanej wartości ze zbioru uczącego.[142]

Walidacja ta jest w stanie określić jedynie stopień dopasowania modelu i jego ewentualne nadmierne dopasowanie, nie jest w stanie jednak zweryfikować, jak model sprawdza się w stosunku do danych, które nie zostały uwzględnione w zbiorze uczącym. W tym celu dokonuje się walidacji zewnętrznej, która służy do oceny tego, z jaką dokładnością model jest w stanie przewidzieć właściwości nieznanymi sobie związków. Walidacja ta wykorzystuje drugi zbiór danych nazywany zbiorem testowym i wydzielonym z danych eksperymentalnych na etapie przygotowywania danych. Zbudowany model QSPR używany jest do obliczenia wartości obserwowanej zmiennej, a wartość ta jest porównywana z danymi eksperymentalnymi. Ostatecznie, współczynnik determinacji  $R^2$  dla walidacji zewnętrznej ( $R^2_{pred}$ ) jest obliczany na podstawie Równania 3.5.

$$R^2_{pred} = 1 - \frac{\sum(Y_{obs(test)} - Y_{pred(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2} \quad (3.5)$$

Równanie to jest bardzo podobne do Równań 3.3 oraz 3.4, a jedyna różnica to  $Y_{obs}$  oraz  $Y_{pred}$  odnoszące się do zbioru testowego i opisywane dla rozróżnienia jako  $Y_{obs(test)}$  oraz  $Y_{pred(test)}$ . Analogicznie jak w przypadku parametru  $Q^2$ , wartość graniczną dla  $R^2_{pred}$  ustalono jako 0,5. Każdy model, który wykazywał wyniki walidacji statystycznej niższe, niż wspomniane wartości, był budowany ponownie używając innych deskryptorów i poddawany reewaluacji.[142]

Do innych stosowanych w walidacji parametrów należą m.in. pierwiastek błędu średniokwadratowego (*ang. Root Mean Squared Error, RMSE*), skorygowany współczynnik determinacji (*ang. Adjusted  $R^2$ ,  $R^2_{adj}$* ) oraz  $y$ -Randomizacja (*ang.  $y$ -Randomization,  $Y_{rand}$* ).

RMSE służy do określenia dokładności predykcji modelu. Parametr ten wrażliwy jest na obecność obserwacji odstających oraz błędnych obserwacji w zbiorze danych, przez co nie jest wskazany w przypadku zbiorów o znacznych rozmiarach. W opisywanych badaniach na obiektywnie niewielkim zbiorze 73 cieczy jonowych może być stosowany z powodzeniem. Im większa jest jego wartość, tym mniej dokładne są przewidywania modelu. RMSE oblicza się na podstawie Równania 3.6.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(Y_{pred i} - Y_{obs i})^2}{n}} \quad (3.6)$$

gdzie  $n$  oznacza liczbę obserwacji, a  $Y_{obs}$  oraz  $Y_{pred}$  odnoszą się odpowiednio do eksperymentalnej wartości obserwowanej zmiennej oraz jej wartości obliczonej.[236]

Kolejnym analizowanym parametrem jest  $R^2_{adj}$ , który pełni funkcję podobną do  $R^2$ , ale uwzględnia dodatkowo liczbę parametrów w modelu. Wzrost liczby parametrów prowadzi do podwyższenia wartości  $R^2$ , co w rezultacie skutkować może nadmiernym dopasowaniem modelu i błędną interpretacją wysokiego współczynnika determinacji. Skorygowany współczynnik

determinancji określa, jaka zmienność zbioru danych została wyjaśniona przez model w odniesieniu do liczby parametrów, które się na niego składają. Obliczany jest z Równania 3.7:

$$R_{adj}^2 = 1 - \left( \frac{(1-R^2)(n-1)}{n-p-1} \right) \quad (3.7)$$

gdzie  $R^2$  oznacza współczynnik determinancji,  $n$  oznacza liczbę obserwacji, a  $p$  oznacza liczbę parametrów w modelu.[237,238]

y-Randomizacja to parametr służący do porównania, jak wartość  $R^2$  oryginalnego modelu wypada na tle modelu zbudowanego w ten sam sposób, ale zmienionym zbiorze uczącym. Podczas obliczania parametru  $Y_{rand}$  dane wejściowe zostają losowo zmodyfikowane – wartości  $Y_{obs}$  w zbiorze uczącym zostają wymieszane i dodatkowo losowo zamienione na wartości  $Y_{pred}$  z modelu. Następnie konstruowany jest nowy model, dla którego oblicza się  $R^2$ . Proces powtarzany jest wielokrotnie, a końcowa wartość współczynnika  $Y_{rand}$  stanowi średnią arytmetyczną wszystkich współczynników determinancji losowo utworzonych modeli. Im niższa jest wartość  $Y_{rand}$  w porównaniu do oryginalnego  $R^2$ , tym lepiej i świadczy to o tym, że model jest selektywny.[239]

Nawet najlepiej dopasowany model o perfekcyjnych parametrach statystycznych jest w stanie przewidzieć jedynie właściwości związków, które znajdują się w granicach dziedziny stosowalności modelu (*ang. Applicability Domain, AD*). Właściwości związków chemicznych charakteryzujących się cechami budowy, które znacznie odbiegają swoją strukturą od pozostałych związków chemicznych użytych do budowy modelu nie będą przewidywalne z akceptowalną dokładnością.[240]

W celu odnalezienia elementów odstających w procesie walidacji krzyżowej oblicza się reszty standaryzowane dla każdej obserwacji i porównuje ich wartości z odchyleniem standardowym. Jeśli reszta standaryzowana jest większa od 2,5- lub 3-krotności odchylenia standardowego, odpowiadająca jej obserwacja zostaje zakwalifikowana jako odstająca (*ang. outlier*), a poprawność predykcji analizowanej wartości dla tego związku może zostać poddana w wątpliwość. Liczba elementów odstających zidentyfikowana w ten sposób w zbiorze testowym jest dobrym wyznacznikiem rzetelności i dokładności zbudowanego modelu matematycznego.

Oprócz opisanych obserwacji odstających, wyróżnić można również ich drugi rodzaj – elementy, które wywierają zdecydowanie większy wpływ na model od jego pozostałych składowych. Zróżnicowany wpływ poszczególnych związków użytych do budowy modelu na jego zdolności predykcyjne jest spodziewaną cechą, lecz niektóre z nich, mimo znajdowania się w zakresie poniżej 3-krotności odchylenia standardowego, mogą charakteryzować się zbyt odbiegającymi od pozostałych cechami strukturalnymi. W przełożeniu na wartości obliczonych deskryptorów oraz – w przypadku analizowanej w omawianych badaniach regresji liniowej – przyporządkowanym im wartościom

współczynników regresji może okazać się, że dany związek wywiera, statystycznie, istotnie większy od pozostałych wpływ na model. Wartością, która opisuje ten wpływ i pozwala go zidentyfikować na etapie walidacji modelu jest dźwignia  $h$  (*ang. leverage*), która jest określana jako odległość od środka zbioru danych. Im większa jest dźwignia danej obserwacji, tym większy wpływ wywiera ona przebieg krzywej regresji. Graniczna wartość dźwigni  $h^*$ , jest obliczana na podstawie Równania 3.8, gdzie  $k$  oznacza liczbę parametrów w modelu, a  $n$  liczbę związków w zbiorze danych.

$$h^* = \frac{3(k+1)}{n} \quad (3.8)$$

Jeśli wartość dźwigni obliczona dla danego związku jest większa od wartości granicznej, związek ten uznawany jest za element odstający. Jeśli znajduje się w zbiorze uczącym, powinien zostać z niego usunięty, natomiast jego obecność w zbiorze testowym świadczy o tym, że obliczona wartość otrzymana została poprzez znaczną ekstrapolację modelu i powinna być traktowana z ograniczonym zaufaniem.

Obie opisane powyżej metody walidacji można przedstawić w formie graficznej na wykresie Williamsa.[241] Na osi X opisane zostają wartości dźwigni, a na osi Y wartości reszt standaryzowanych. Liniami zostają również zaznaczone wartości graniczne. Wykres ten pozwala w przystępny sposób wizualnie ocenić jakość otrzymanego modelu oraz sprawnie zidentyfikować obserwacje odstające.[83]

## 4. WYNIKI

### 4.1. Modelowanie stałej Henry'ego

Stała Henry'ego jest właściwością zależną od wielu czynników, takich jak jej polarność cieczy jonowych oraz złożoność ich budowy cząsteczkowej. Spodziewano się, że uzyskane modele oparte będą na wielu zróżnicowanych deskryptorach zamiast pojawiających się przeważnie w literaturze modeli QSPR zbudowanych z ich niewielkiej liczby. Zgodnie z przewidywaniami, końcowy zbiór danych zawierał sześć najważniejszych deskryptorów, na których podstawie zbudowano modele predykcyjne. Pierwsze trzy deskryptory (GATS2s, P\_VSA\_PPP\_A oraz Chi\_G/D) dotyczyły kationów, a trzy pozostałe (B02[F-F], Hy oraz CATS3D\_04\_AA) opisywały aniony. W Tabeli 4.1 znajduje się zestawienie definicji analizowanych deskryptorów wraz z kategoriami, do których należą.

| Deskryptor    | Definicja                                                                                            | Kategoria                      |
|---------------|------------------------------------------------------------------------------------------------------|--------------------------------|
| <b>KATION</b> |                                                                                                      |                                |
| GATS2s        | autokorelacja Gearego w odległości 2, ważona po stanie wewnętrznym                                   | Autokorelacje 2D               |
| P_VSA_ppp_A   | obecność w strukturze cząsteczki potencjalnych farmakoforów w postaci akceptorów wiązania wodorowego | Deskryptory P_VSA              |
| Chi_G/D       | indeks Randica obliczany z macierzy odległości geometryczna/odległość topologiczna                   | Deskryptory 3D na bazie matryc |
| <b>ANION</b>  |                                                                                                      |                                |
| B02[F-F]      | obecność/brak dwóch atomów fluoru w odległości topologicznej 2 od siebie                             | Pary atomów 2D                 |
| Hy            | Indeks hydrofilowości                                                                                | Właściwości cząsteczki         |
| CATS3D_04_AA  | dwa potencjalne miejsca akceptorowe w odległości 4-5 Å od siebie                                     | CATS 3D                        |

**Tabela 4. 1.** Zestawienie definicji analizowanych deskryptorów wraz z kategoriami, do których należą

Współczynniki regresji przyporządkowane sześciu deskryptorom w każdym z trzech zbudowanych modeli znajdują się poniżej w Tabeli 4.2. Wartości obliczonych deskryptorów zestawiono w Tabeli A4 w Materiałach uzupełniających.

| Deskryptor   | MLR      | LR        | PLS      |
|--------------|----------|-----------|----------|
| GATS2s       | -3240,48 | -0,161256 | -432,025 |
| P_VSA_ppp_A  | 83,68    | 0,748779  | 2219,77  |
| Chi_G/D      | 20173,94 | 0,222386  | 438,896  |
| B02[F-F]     | -5013,51 | -0,781993 | -2133,9  |
| Hy           | 7538,04  | 0,499704  | 1113,84  |
| CATS3D_04_AA | -835,30  | -0,330897 | -795,364 |
| Stała        | -6628,05 | -1,23842  | 6643,79  |

**Tabela 4. 2.** Współczynniki regresji deskryptorów w analizowanych modelach

Do modeli trafiły zarówno deskryptory należące do kategorii 2D, jak i kategorii 3D. Dwa z trzech deskryptorów kationów (P\_VSA\_PPP\_A oraz Chi\_G/D) są deskryptorami trójwymiarowymi. Przeciwna sytuacja ma miejsce w przypadku deskryptorów anionów, gdzie jedynie CATS3D\_04\_AA jest deskryptorem 3D. W cieczach jonowych kationy są przeważnie większe od anionów, co pozwala przypuszczać, że deskryptory uwzględniające ich objętość będą miały większy wpływ na rozpuszczalność ditlenku węgla niż analogiczne deskryptory dotyczące anionów.

Zestawienie wartości eksperymentalnie zmierzonych wartości  $H_{CO_2}$  z wartościami przewidzianymi przez modele MLR, PLS oraz LR znajduje się w Tabeli A3 w Materiałach uzupełniających.

#### 4.1.1. Walidacja

Parametry statystyczne zbudowanych modeli przedstawiono poniżej w Tabeli 4.3.

| Model | $R^2$  | $Q^2$  | $R^2_{pred}$ | RMSE (uczący) | RMSE (testowy) |
|-------|--------|--------|--------------|---------------|----------------|
| MLR   | 0,9241 | 0,8780 | 0,9104       | 1265          | 1185           |
| LR    | 0,9079 | 0,8916 | 0,9035       | 1134          | 1230           |
| PLS   | 0,9240 | 0,8803 | 0,9276       | 1309          | 1065           |

**Tabela 4. 3.** Parametry statystyczne modeli

Wszystkie trzy zbudowane modele osiągnęły zadowalające parametry statystyczne podczas walidacji. Współczynnik determinacji  $R^2$ , który określa jaką część wariancji zmiennej w danych wejściowych jest opisywana przez zmienne w równaniu modelu, dla każdego z analizowanych przypadków przekroczył wartość 0,9. Modele uszeregowane w kolejności malejącej pod względem wartości  $R^2$  to: MLR > PLS > LR. Z wewnętrzną walidacją najlepiej poradził sobie model LR, jednakże wartości  $Q^2$  dla pozostałych modeli były bardzo zbliżone. Modele ułożone w kolejności malejącej pod względem wartości  $Q^2$  prezentują się następująco: LR > PLS > MLR. Najbardziej homogenicznym współczynnikiem okazał się być współczynnik walidacji zewnętrznej,  $R^2_{pred}$ , który zmniejszał się w kolejności: PLS > MLR > LR. Obliczone wartości pierwiastka z błędu średniokwadratowego (*ang. Root Mean Square Error*, RMSE) dla zbioru uczącego i testowego były do siebie zbliżone.

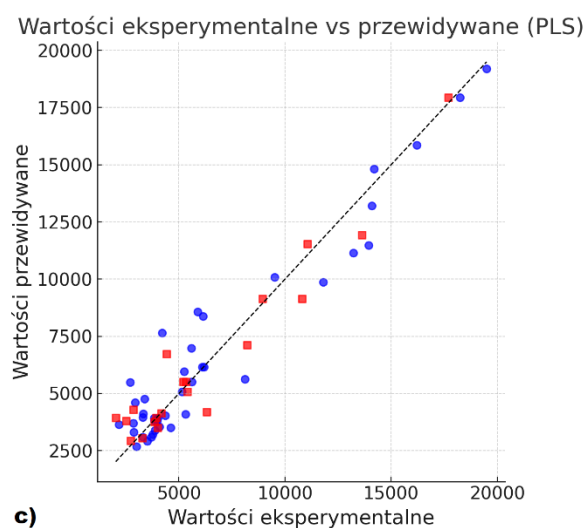
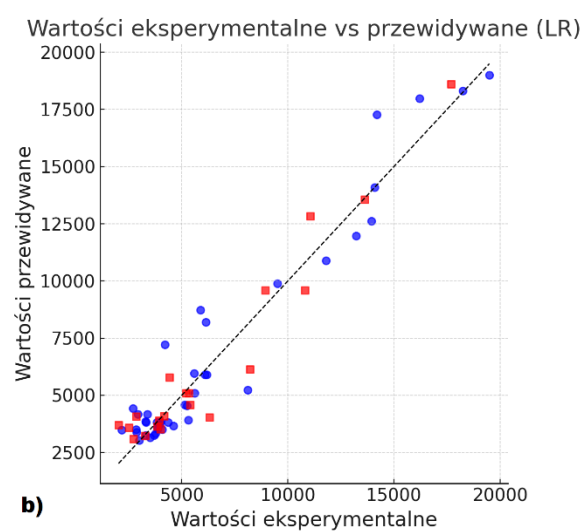
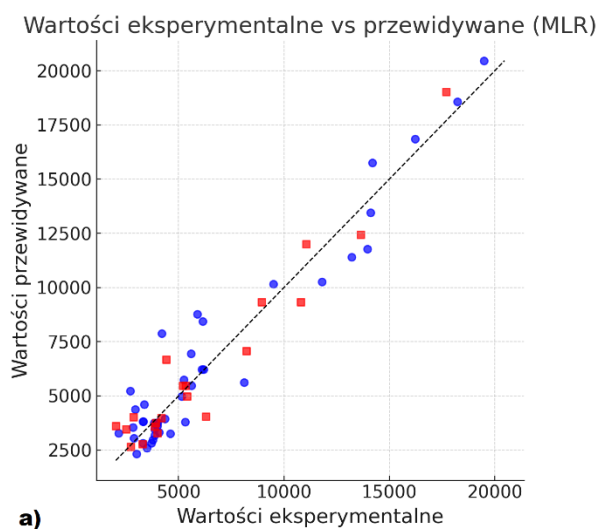
Model MLR charakteryzował się najlepszym wyjaśnieniem wariancji, ale wypadł najgorzej podczas walidacji wewnętrznej. Należy jednak zauważyć, że wartość  $Q^2$  była bardzo zbliżona do wartości  $Q^2$  modelu PLS. Modelem, który wypadł najlepiej w walidacji wewnętrznej był za to model LR, który z kolei charakteryzował się najniższą wartością  $R^2$  oraz poradził sobie najgorzej w walidacji zewnętrznej. Pośrednimi wartościami  $R^2$  i  $Q^2$  wykazał się model PLS, który wypadł najlepiej w walidacji zewnętrznej.

Aby ocenić i porównać skonstruowane modele, trzy analizowane parametry potraktowano jako równie istotne, a modelom przyznano punkty w zależności od miejsca zajmowanego po uszeregowaniu trzech parametrów w kolejności malejącej. Za pierwsze miejsce, czyli najwyższą wartość parametru przyznano jeden punkt, za drugie dwa, a za trzecie – analogicznie – trzy, a za najlepszy uznano model, który zdobył najmniejszą liczbę punktów. Modele uszeregowano w następującej kolejności: PLS > MLR > LR, co odpowiadało odpowiednio zdobyciu 5, 6 i 7 punktów. Warto zwrócić jednak uwagę, że wszystkie modele charakteryzowały się bardzo dobrymi parametrami statystycznymi, a różnice pomiędzy nimi były w większości przypadków zanedbywalne.

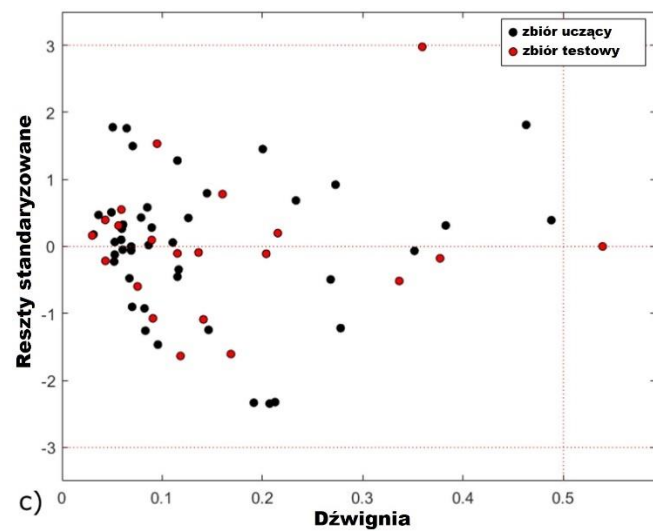
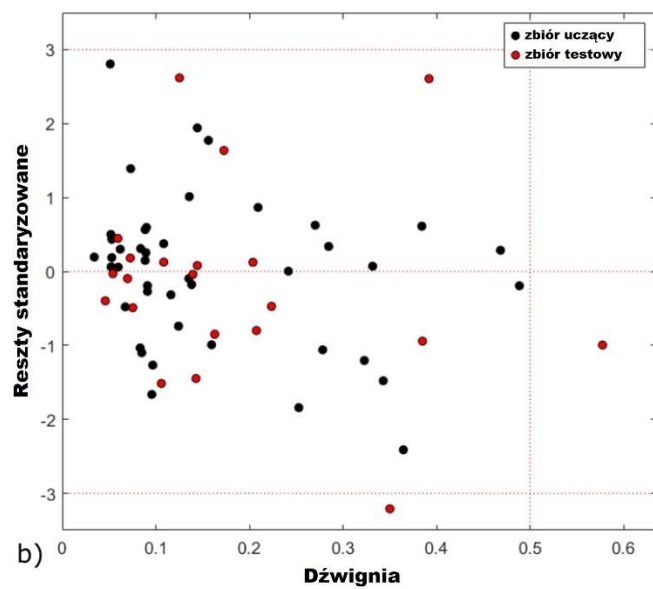
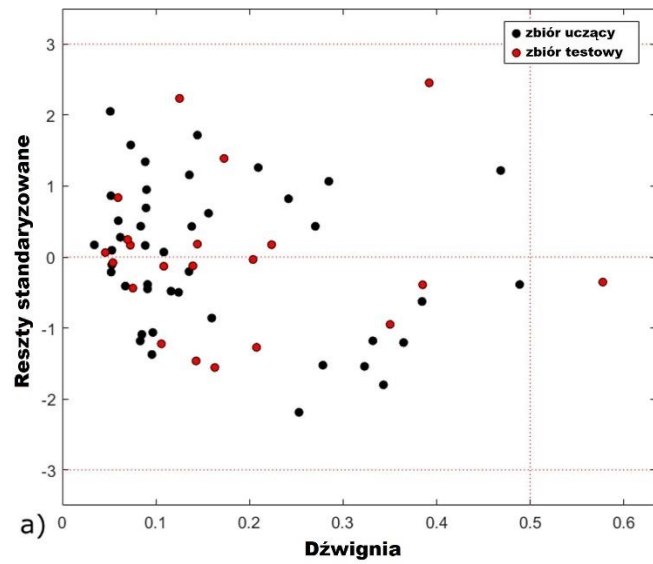
Przewidziane we wszystkich trzech modelach wartości  $H_{CO_2}$  z dużą dozą dokładności odpowiadały pomiarom eksperymentalnym, co widać na Rysunku 4.1.

Na Rysunku 4.2 przedstawiono wykresy Williamsa dla każdego z trzech modeli wykonane za pomocą programu Applicability domain toolbox.[242] W każdym z nich jeden ze związków ze zbioru testowego przekroczył wartość graniczną dźwigni. W modelu LR pojawił się dodatkowy element odstający w postaci związku ze zbioru testowego, który przekroczył granicę trzech odchyłeń standardowych. Liczba obserwacji odstających we wszystkich modelach jest niewielka w porównaniu do liczby wszystkich obserwacji, a w żadnym z nich nie pojawił się outlier ze zbioru uczącego.

Wysokie wartości parametrów statystycznych  $R^2$ ,  $Q^2$  oraz  $R^2_{pred}$ , a także analiza wykresów Williamsa udowadnia, że modele są prawidłowo skonstruowane i są w stanie dokładnie przewidzieć wartości  $H_{CO_2}$  dla związków w swojej dziedzinie stosowności z dużą dokładnością. Mimo że model PLS wypadł najlepiej przy założeniu równej istotności wszystkich trzech parametrów, różnica pomiędzy nim, a modelem MLR jest zanedbywalnie mała. Każdy z analizowanych modeli może zostać z powodzeniem użyty do przewidywania  $H_{CO_2}$  ditlenku węgla w cieczach jonowych, jednak ze względu na brak istotnych różnic pomiędzy nimi, a także na łatwość interpretacji modelu opartego na wielokrotnej regresji liniowej, podjęto decyzję o stosowaniu tej metody w kolejnych badaniach.



**Rysunek 4. 1.** Wykresy zależności wartości przewidywanej  $H_{CO_2}$  od wartości eksperymentalnej dla trzech modeli. a) MLR, b) LR, c) PLS. Kolor niebieski – zbiór uczący, kolor czerwony – zbiór testowy



Rysunek 4. 2. Wykresy Williamsa dla modeli a) MLR, b) LR, c) PLS

#### 4.1.2. Interpretacja deskryptorów

Jak wspomniano w poprzednim akapicie, model oparty na MLR jest najłatwiejszy w interpretacji. Każdy z deskryptorów w modelu charakteryzuje się określoną wagą, a krzywa regresji jest wynikiem liniowej kombinacji wkładu każdego z nich. Ze względu na bardzo zbliżone wyniki analizy statystycznej modeli, założono, że interpretacja wpływu deskryptorów oparta na istotności statystycznej współczynników regresji liniowej może zostać z powodzeniem przełożona na modele PLS oraz LR.

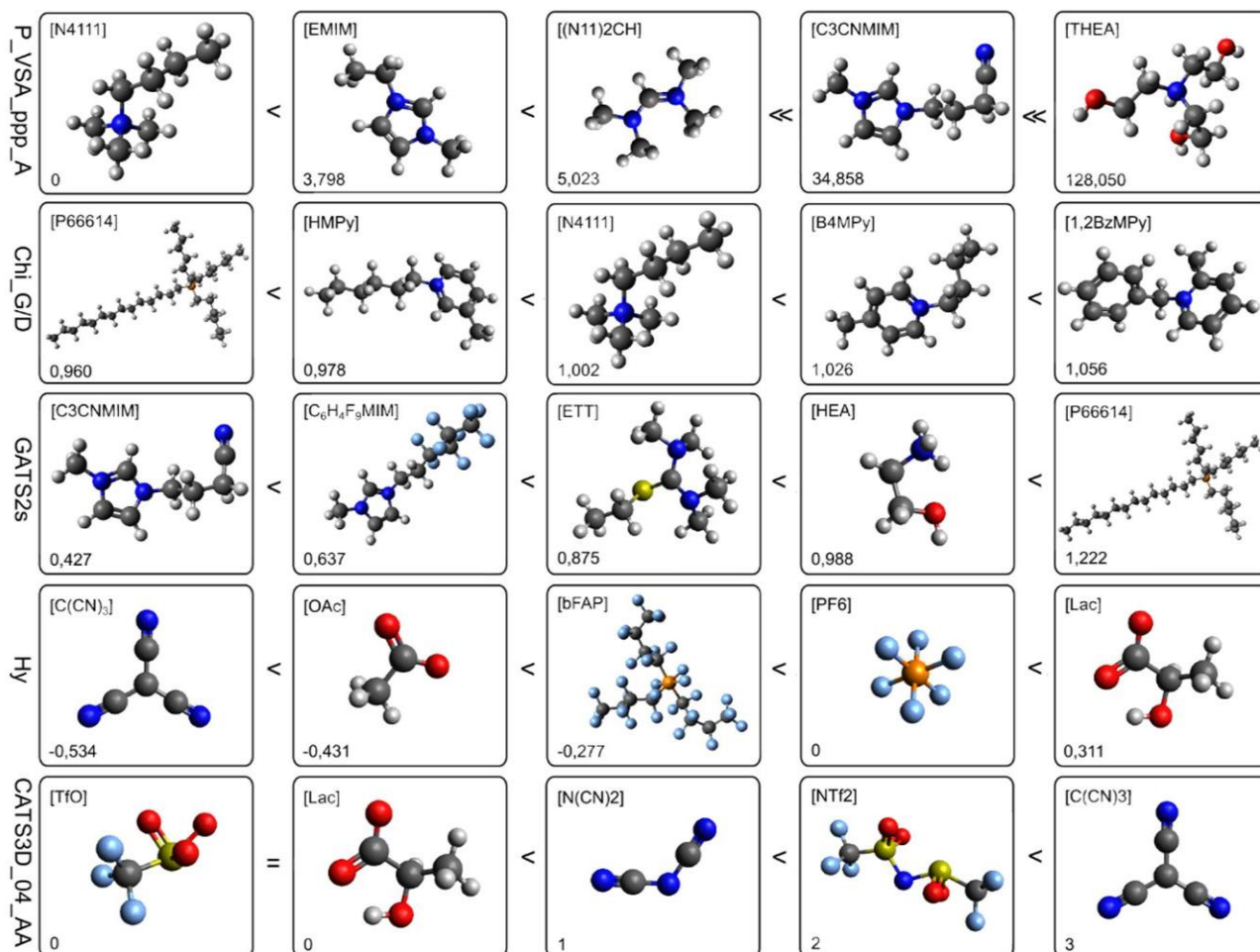
Przeprowadzenie analizy oraz interpretacja deskryptorów w relacji do cech strukturalnych cząsteczek dostarcza wielu wartościowych informacji dotyczących mechanizmów odpowiedzialnych za zdolność cieczy jonowych do rozpuszczania ditlenku węgla, ale sama w sobie nie wystarczy – każdy z deskryptorów oddziałuje przecież z różną siłą. W Tabeli 4.4 przedstawiono deskryptory uszeregowane pod względem ich istotności statystycznej. Dla modelu MLR obliczono parametr  $t$ , natomiast dla modelu LR parametr Walda. Wskaźniki te służą do określenia i porównania wpływu poszczególnych zmiennych na ostateczny wynik modelowania. Im wyższa jest ich wartość, tym wpływ jest istotniejszy. W obu modelach deskryptory uszeregowane od najważniejszego do najmniej ważnego ułożone są w niemal identycznej kolejności – za wyjątkiem dwóch ostatnich – kationowych GATS2s oraz Chi\_G/D, których miejsca są zamienione, lecz obliczone wartości są bardzo zbliżone do siebie. Analizując modele całościowo widać, że mimo, że najważniejszym deskryptorem jest kationowy P\_VSA\_PPP\_A, to wpływ anionów na rozpuszczalność CO<sub>2</sub> w cieczach jonowych jest większy, a różnica we wpływie poszczególnych deskryptorów jest znacząca. Nie oznacza to jednak, że wpływ GATS2s oraz Chi\_G/D powinien zostać zignorowany. Warto nadmienić, że podczas modelowania usunięcie obu deskryptorów z modeli prowadziło do pogorszenia ich jakości.

| Deskryptor      | Parametr Walda | Parametr $t$ |
|-----------------|----------------|--------------|
| P_VSA_ppp_A_cat | 71,802         | 7,906        |
| B02[F-F]_an     | 51,150         | 7,805        |
| Hy_an           | 23,115         | 5,525        |
| CATS3D_04_AA_an | 9,754          | 3,225        |
| Chi_G/D_cat     | 3,868          | 1,949        |
| GATS2s_cat      | 3,585          | 2,114        |

**Tabela 4. 4.** Zestawienie parametrów istotności statystycznej deskryptorów

W poniższym akapicie najpierw zostaną omówione deskryptory powiązane z kationami, a następnie deskryptory związane z anionami. Graficzne przedstawienie wybranych cieczy jonowych

uszeregowanych w kolejności odpowiadającej zmianom wartości analizowanych deskryptorów (rosnąco) przedstawiono na Rysunku 4.3.



**Rysunek 4. 3.** Graficzna reprezentacja struktur jonów wraz z odpowiadającymi im wartościami analizowanych deskryptorów

#### 4.1.2.1. Deskryptory kationowe

GATS2s to deskryptor 2D definiowany jako „autokorelacja Gearego w odległości 2, ważona po stanie wewnętrznym”. Wartość stanu wewnętrznego atomu (*ang. intrinsic state, I-state*) zależy od jego elektroujemności, liczby elektronów walencyjnych oraz wiązań  $\sigma$  i  $\pi$  łączących go z sąsiadującymi atomami.[243] Deskryptor ten jest miarą autokorelacji przestrzennej i przyjmuje wartości od zera do nieskończoności. Wartości w przedziale ( $<0:1$ ) oznaczają pozytywną korelację, wartość 1 oznacza brak korelacji, a wartości ( $1:\infty$ ) oznaczają korelację ujemną.[84] Pozytywna autokorelacja Gearego oznacza, że obszary o podobnej elektroujemności występują w cząsteczce obok siebie i w większych skupieniach, natomiast ujemna autokorelacja wskazuje, że obszary te występują naprzemiennie. W analizowanym badaniu autokorelacja obliczana jest dla obszarów w odległości dwóch atomów od siebie po zapisaniu wzoru cząsteczki w formie grafu. Wyższe wartości tego deskryptora odpowiadają niższym wartościom  $H_{CO_2}$ , więc w celu uzyskania wysokiej rozpuszczalności ditlenku węgla w cieczy jonowej preferowane są

kationy charakteryzujące się ujemną korelacją stanu wewnętrznego atomów. Spośród wszystkich analizowanych cieczy, najniższe wartości deskryptora GATS2s wykazywały cząsteczki zawierające grupy funkcyjne -CN, takie jak kation 1-(3-cyjanopropyl)-3-metylimidazolowy o wartości 0,427. Wpływ grypu -CN jest znaczący, co widać na przykładzie analogu strukturalnego omawianego kationu, który jest pozbawiony grupy nitylowej (1-metyl-3-propylimidazol) i wykazuje wartość GATS2s na poziomie 0,919, wskazującą na bardziej losowe rozmieszczenie atomów o zbliżonej elektroujemności. Rodzaj pierścienia w kationie ma niewielki wpływ na wartość deskryptora: dla kationów 3-metylimidazolowych, 3-metylpirydynowych i 3-metylpirolidynowych z podstawnikiem butylowych obliczone wartości to odpowiednio: 0,919, 0,863 oraz 0,949. Spośród całego zbioru cieczy jonowych jedynie trzy kationy charakteryzowały się ujemną autokorelacją i posiadały wartość deskryptora wskazującą na jego istotność dla procesu rozpuszczania CO<sub>2</sub> (GATS2s > 1), a najwyższa obliczona wynosiła 1,222 dla kationu triheksylo(tetradecylo)-fosfoniowego.

Kolejnym omawianym deskryptorem jest deskryptor 3D Chi\_G/D, definiowany jako indeks Randica obliczany z macierzy odległość geometryczna/odległość topologiczna. Macierz ta jest obliczana na podstawie stosunku odległości geometrycznej do topologicznej. Odległość geometryczna wyznaczana jest na podstawie rzeczywistych odległości pomiędzy atomami w zoptymalizowanej trójwymiarowej strukturze cząsteczki, natomiast odległość topologiczna obliczana jest na podstawie odległości pomiędzy atomami stanowiącymi węzły w strukturze cząsteczki przedstawionej w formie dwuwymiarowego grafu. Wartość Chi\_G/D zależy również od przestrzennego rozkładu atomów, gdzie związki cykliczne charakteryzują się najwyższymi wartościami, a związki bardziej rozgałęzione niższymi. Deskryptor ten zawiera informacje o stopniu rozgałęzienia cząsteczki, a także o jej pofałdowaniu (np. w przypadku długich łańcuchów alkilowych).[244] Można założyć, że bardziej rozgałęzione związki będą mniej upakowane od mniej rozległych przestrzennie związków cyklicznych. Wraz ze wzrostem upakowania cząsteczek zmniejsza się wartość wolnej objętości pomiędzy nimi, która mogłaby zostać zajęta przez cząsteczki ditlenku węgla, co efektywnie przekłada się na spadek rozpuszczalności CO<sub>2</sub>. Teza ta znajduje swoje potwierdzenie w obserwacjach, gdyż wyższe wartości H<sub>CO2</sub> w analizowanych modelach są powiązane z wyższymi wartościami tego deskryptora. Zakres jego wartości jest jednak dość wąski, od 0,9417 dla kationu [C8H4F13MIM] do 1,0559 dla [1,2BzMPY]. Związki chemiczne o najdłuższych łańcuchach alkilowych, takie jak [C8H4F13MIM], [OMIM] oraz [P66614] charakteryzują się najniższymi wartościami, a więc najlepszą rozpuszczalnością ditlenku węgla. Warto nadmienić, że niska wartość deskryptora została również obliczona dla stosunkowo małego i liniowego kationu [BHEA]. Deskryptor Chi\_G/D jest silnie zależny od obecności i rodzaju pierścienia w strukturze kationu. Dla cząsteczki [BMIM] przyjmuje wartość 0,980, natomiast dla analogu pirydynowego [B3MPy], gdzie jedyną różnicą w strukturze jest obecność dodatkowego atomu węgla w pierścieniu, wartość ta rośnie

do 1,028. Co więcej, wzrost liczby atomów węgla w przypadku pierścienia niearomatycznego ma mniejszy wpływ na wzrost wartości deskryptora, co widać na przykładzie cząsteczki [BMPyrr], dla której  $\text{Chi}_G/D$  równy jest 1,017. Zgodnie z powyższymi obserwacjami, najwyższymi wartościami deskryptora cechują się kationy zawierające aromatyczne pierścienie – pirydynowe i benzylowe.

Ostatnim kationowym deskryptorem jest  $P\_VSA\_PPP\_A$ , czyli deskryptor 3D opisujący obecność w strukturze cząsteczki potencjalnych farmakoforów w postaci akceptorów wiązania wodorowego. Związany jest z właściwościami hydrofobowymi cząsteczki i obliczany jest na podstawie jej parametrów sterycznych i elektronowych, od których zależy jej zdolność do bycia akceptorem wiązania wodorowego (H-B). Im więcej potencjalnych miejsc akceptorowych zostaje zidentyfikowanych w strukturze cząsteczki, tym wyższa jest wartość obliczanego deskryptora.[245] Wraz z jego wzrostem rośnie stała Henry’ego, co przekłada się na spadek rozpuszczalności  $\text{CO}_2$ . W analizowanym zbiorze kationów, najniższe wartości  $P\_VSA\_PPP\_A$  przyporządkowano trzecio- i czwartorzędowym aminom. Dla tych ostatnich zdolność do akceptowania wiązań wodorowych została obliczona jako równa zeru. Niewielkim potencjałem akceptorowym cechują się z kolei zdelokalizowane wiązania azot-węgiel, a wartości deskryptora są wyższe dla związków liniowych, takich jak [(N11)2CH], niż dla związków cyklicznych zawierających atom azotu wewnątrz pierścienia aromatycznego (np. kationy pirydynowe i imidazolowe). Obecność grup nitrylowych i hydroksylowych związana jest z drastycznym wzrostem wartości deskryptora (ok. 30-40 jednostek dla każdej z grup funkcyjnych), co zgodnie z analizą współczynników regresji modelu, przekłada się na spadek rozpuszczalności  $\text{CO}_2$ . We wszystkich przypadkach wartość  $P\_VSA\_PPP\_A$  była związana z obecnością heteroatomów w strukturze cząsteczki, a także z rzędowością ich wiązań. Organiczny szkielet pozostawał niemalże bez wpływu na wartość deskryptora, co doskonale ilustruje szereg analogów imidazolowych różniących się długością łańcuchów alkilowych, dla których wartość  $P\_VSA\_PPP\_A$  pozostawała stała i wynosiła ok. 3,8 bez względu na pozycję i długość łańcucha.

#### 4.1.2.2. Deskryptory anionowe

Spośród trzech obecnych w modelach deskryptorów anionów,  $B02[F-F]$  jest deskryptorem najistotniejszym statystycznie ( $t$ -statystyka równa 7,81 oraz statystyka Walda równa 51,15), a zarazem najłatwiejszym do interpretacji. Jest jedynym deskryptorem jakościowym w omawianym badaniu i dotyczy obecności lub braku dwóch atomów fluoru w topologicznej odległości dwóch atomów od siebie. Przyjmuje wartości 1 lub 0, odpowiednio w zależności od występowania pary atomów F-F lub ich braku. W rezultacie, w zbiorze anionów analizowanych w badaniu, istnieją tylko dwa podzbiory – zawierające fluor lub nie. Wartość 1 jest związana ze zmniejszeniem  $H_{\text{CO}_2}$ . Deskryptor ten ma wiele ograniczeń. Związki zawierające pojedynczy atom fluoru zostają zakwalifikowane jako wartość 0, a w

przypadku częstszego występowania atomów fluoru ich dokładna liczba nie ma znaczenia.[84] W rezultacie anion [bFAP] charakteryzuje się taką samą wartością B02[F-F], co [TA], mimo że trifluorooctan zawiera w swojej cząsteczce dziesięć razy mniej atomów fluoru. Zaskakujący jest fakt, że zamiana deskryptora na dużo dokładniejszy deskryptor nF, który zawiera informacje o dokładnej liczbie atomów, poskutkowało znaczącym pogorszeniem parametrów statystycznych modeli. Nie zmienia to jednak faktu, że obecność wielu atomów fluoru w cząsteczce anionu cieczy jonowej zwiększa jej zdolność do rozpuszczania ditlenku węgla.

Kolejnym deskrytorem jest Hy, definiowany jako indeks hydrofilowości i oparty na metodzie liczenia atomów oraz grup funkcyjnych. Jest deskrytorem dwuwymiarowym, a wzajemne położenie analizowanych grup nie ma wpływu na jego wartość. Obliczany jest za pomocą Równania 4.1, w którym  $N_{Hy}$  oznacza liczbę grup hydrofilowych (-OH, -SH oraz -NH),  $N_c$  jest liczbą atomów węgla, natomiast A określa łączną liczbę atomów w cząsteczce nie uwzględniając atomów wodoru.[84]

$$Hy = \frac{(1+N_{Hy})\log_2(1+N_{Hy})+N_c(\frac{1}{A}\log_2\frac{1}{A})+\sqrt{\frac{N_{Hy}}{A^2}}}{\log_2(1+A)} \quad (4.1)$$

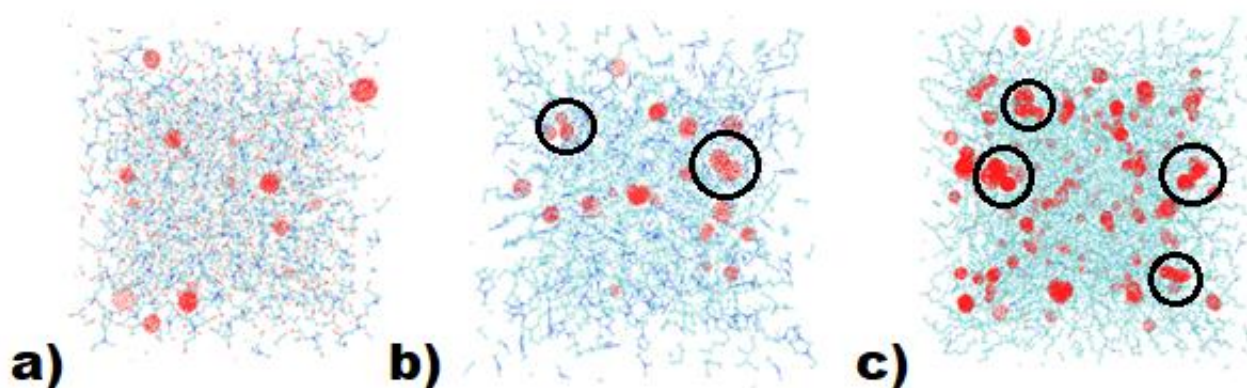
Pod względem istotności statystycznej jest to drugi najważniejszy deskryptor anionowy. Jak wynika ze wzoru, jego wartość jest zależna od liczby grup hydrofilowych oraz ich stosunku do liczby pozostałych atomów. Wartość Hy rośnie wraz ze wzrostem hydrofilowości cząsteczki. Brak jednej lub więcej z wymienionych grup hydrofilowych obniża wartość deskryptora, prowadząc do uzyskania wartości ujemnych. W analizowanym zbiorze danych jedynie anion mleczanowy cechował się dodatnią wartością Hy (0,311), co w porównaniu do klasycznych rozpuszczalników stawia go obok propanolu (Hy = 0,37). Dla porównania cząsteczka wody charakteryzuje się wartością Hy na poziomie 3,44. Wyraźnie wskazuje to, że na deskryptor wpływ ma również stosunek liczby heteroatomów do liczby atomów węgla. Nieznacznie ujemne wartości Hy obliczone dla perfluorowanych anionów fosfoniowych ([bFAP], [eFAP], [pFAP]) są konsekwencją wysokiego stosunku atomów fluoru do atomów węgla w ich łańcuchach alkilowych. Biorąc pod uwagę dodatni współczynnik regresji deskryptora w modelu, można sformułować ogólną zależność: obecność hydrofilowych grup -OH, -SH oraz -NH zwiększa wartość  $H_{CO_2}$ , a wysoki stosunek heteroatomów do atomów węgla obniża ją, co przekłada się na zwiększenie rozpuszczalności  $CO_2$ . Nieorganiczne aniony nie mają na nią wpływu, gdyż obliczona dla nich wartość Hy równa jest 0. Im mniej hydrofilowy jest anion cieczy jonowej, tym większa jest jej rozpuszczalność ditlenku węgla. Najwyższe wartości deskryptora zostały obliczone dla anionu [Lac], zaś najniższa dla [OAc] oraz  $[C(CN)_3]$  i wynosiły odpowiednio 0,311391, -0,430677 oraz -0,534734. Mleczan i octan różnią się strukturalnie obecnością grupy -OH w mleczanie, co prowadzi do wniosku, że obecność grup hydroksylowych w anionach jest niepożądana.

Ostatnim omawianym deskryptorem anionowym jest jedyny deskryptor 3D w tym zestawieniu. CATS3D\_04\_AA łączy podobieństwo z omawianym w przypadku kationów deskryptorem P\_PVS\_PPP\_A, gdyż on również, w sposób ilościowy, reprezentuje obecności akceptorów wiązań wodorowych w cząsteczce. Różnica polega na tym, że w przypadku obliczania CATS3D\_04\_AA brane są pod uwagę jedynie dwa potencjalne miejsca akceptorowe w odległości 4-5 Å od siebie.[84] Podobnie jak w przypadku B02[F-F] rozdzielczość tego deskryptora jest niewielka. W całym zbiorze danych występują jedynie wartości: 3, 2, 1 przyporządkowane odpowiednio anionom  $[C(CN)_3]$ ,  $[NTf_2]$ ,  $[N(CN)_2]$  oraz 0 dla wszystkich pozostałych cząsteczek. Zamiana tego deskryptora na analogiczny, ale o większej rozdzielczości (CATS3D\_00\_AA, który analizuje zakres odległości od 0 do 6 Å), podobnie jak w przypadku deskryptora opisującego liczbę atomów fluoru, doprowadziła do znacznego pogorszenia parametrów statystycznych modeli. Ujemny współczynnik regresji wskazuje, że obecność akceptorów wiązań wodorowych w anionach obniża stałą Henry'ego, a w rezultacie zwiększa rozpuszczalność ditlenku węgla. Zależność ta jest więc odwrotna, niż w przypadku analizy kationów.

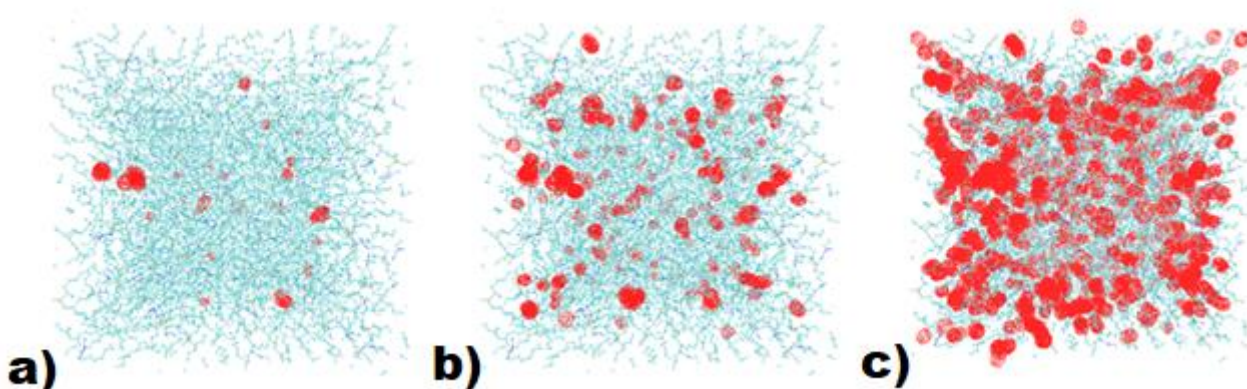
## **4.2. Analiza wolnej objętości**

### **4.2.1. Frakcyjna wolna objętość**

Kluczową rzeczą podczas analizy wolnej objętości związku chemicznego jest zrozumienie powstawania i rozkładu luk w jego strukturze. Jak widać na Rysunku 4.4 ciecze jonowe cechujące się wyższymi wartościami FFV zwykle posiadają w swojej strukturze więcej luk o średnicy 170 pm, czyli niewiele mniejszej niż z rozmiar cząsteczki  $CO_2$  wynoszący ok. 200pm (luki zaznaczono na rysunkach kolorem czerwonym). Mimo, że w większości cieczy liczba luk o tak znacznych rozmiarach (200 pm średnicy) jest niewielka, licznie występują w nich mniejsze (170 pm) luki zlokalizowane w swojej bezpośredniej okolicy. Można założyć, że wspomniane skupiska luk są w stanie zmienić ułożenie atomów w wyniku działania sił intermolekularnych, a w rezultacie połączyć się i wytworzyć miejsce, które będzie w stanie zająć cząsteczka gazu. Co więcej, Rysunek 4.5 pokazuje, że liczba luk w strukturze cieczy drastycznie zwiększa się wraz ze spadkiem ich średnicy. I tak, na rysunku 4.5a widać, że liczba dużych luk, o rozmiarze 200 pm, jest niewielka, natomiast na rysunkach b oraz c widać, że liczba luk zauważalnie zwiększa się wraz ze spadkiem ich średnicy.

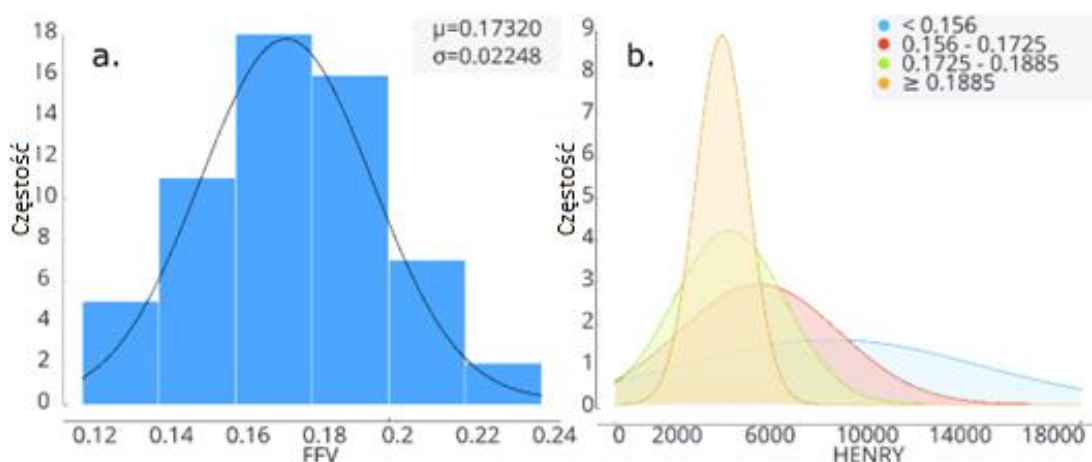


**Rysunek 4. 4.** Luki w cieczech o niskim, średnim i wysokim FFV. a) [THE][OAc], b) [HMIM][Cl], c) [P66614][N(CN)<sub>2</sub>]. Na rysunkach przedstawiono luki o średnicach 170 pm. Na rysunku zaznaczono potencjalne miejsca tworzenia się skupisk wolnych przestrzeni.



**Rysunek 4. 5.** Porównanie liczby luk o różnych rozmiarach w cieczy [P66614][DCA]. a) 200 pm, b) 170 pm, c) 140 pm

Obliczona frakcyjna wolna objętość analizowanych cieczy jonowych wahała się w zakresie od 0,133 do 0,232 ze współczynnikiem zmienności na poziomie 13%. Na przykładzie Rysunku 4.6a widać, że rozkład wartości FFV w zbiorze danych w znacznym stopniu odpowiada rozkładowi normalnemu. Ciekawsze wnioski można wyciągnąć z Rysunku 4.6b czyli porównania rozkładu wartości  $H_{CO_2}$  w odniesieniu do FFV. Zakres zmienności wartości frakcyjnych wolnych objętości podzielono na cztery kwartyle i nadano im kolory - niebieski, czerwony, zielony i pomarańczowy – które w rosnącej kolejności odpowiadają wzrostowi FFV i przyporządkowano im korespondującą wartość zmierzonej eksperymentalnie  $H_{CO_2}$ . Z analizy histogramu wynika, że dystrybucja FFV traci cechy rozkładu normalnego, a średnia przesuwana się w kierunku niższych wartości  $H_{CO_2}$  wraz ze wzrostem wolnej objętości, co oznacza, że im większa jest wartość FFV, tym większa jest także rozpuszczalność ditlenku węgla. Wnioski te potwierdzają założenia teoretyczne i mogą sugerować, że rozpuszczanie  $CO_2$  przez ciecze jonowe w dużej mierze odbywa się na zasadzie absorpcji cząsteczek gazu i wypełniania nim wolnych luk w ich strukturze.



**Rysunek 4. 6.** a) rozkład wartości FFV, b) rozkład wartości FFV w odniesieniu do  $H_{CO_2}$

Należy odnotować również, że w przypadku niektórych cieczy występują rozbieżności pomiędzy zaobserwowanymi zasadami a zależnością wartości  $H_{CO_2}$  a wartością FFV. Szczególnym przypadkiem są ciecze zawierające perfluorowane podstawniki w anionach, takich jak [HMIM][eFAP], gdzie obliczona wolna objętość jest stosunkowo niewielka, a mimo to stała Henry'ego ma małą wartość. Przypadki te wskazują najprawdopodobniej na inny mechanizm absorpcji gazu lub interakcje pomiędzy cząsteczkami  $CO_2$  a licznymi atomami fluoru.

Analiza porównawcza cieczy jonowych o wspólnych anionach i różniących się kationami dostarcza informacji na temat wpływu kationów na wartość FFV. Podstawową zależnością jest wpływ rozmiaru łańcucha alkilowego na wolną objętość. Porównując ciecze [HMIM][eFAP] oraz [EMIM][eFAP] różniące się długością łańcucha alkilowego o cztery atomy węgla okazuje się, że wartość FFV dla dłuższego łańcucha jest większa o 6%. Analogiczną zależność można zaobserwować porównując [BMIM][N(CN)<sub>2</sub>] oraz [EMIM][N(CN)<sub>2</sub>] różniące się tylko dwoma atomami węgla i 4% różnicą w wartości FFV. Efekt ten jest jeszcze bardziej zauważalny w przypadku obecności podstawnika w łańcuchu alkilowym, gdzie pojawienie się dwóch grup fluorometylenowych odpowiada za wzrost wartości FFV o 23%, z 0,166 do 0,204. Szeregując malejąco cztery ciecze jonowe zawierające anion [BF<sub>4</sub>] uzyskujemy kolejność [OMIM] > [BMIM] > [EMIM] = [MMIM], a szeregując ciecze zawierające anion [PF<sub>6</sub>] uzyskujemy kolejność [OMIM] > [BMIM] = [PMIM]. Obserwacje te wskazują, że większy wpływ na FFV uzyskuje się przy dłuższych łańcuchach alkilowych, natomiast nieznaczne wydłużenie krótkiego łańcucha ma zanedbywalny wpływ na wolną objętość. Krótsze łańcuchy alkilowe są w stanie lepiej ułożyć się w przestrzeni zajmowanej przez ciecz, co skutkuje większym upakowaniem i występowaniem mniejszej liczby luk. W przypadku amoniowych cieczy jonowych zawierających wspólny anion mrówczanowy zauważyć można, że kationy amoniowe o wyższej rzędowości charakteryzują się wyższymi wartościami FFV. Zależności tej nie można jednak zaobserwować w przypadku cieczy amoniowych z anionem mleczanowym. Może być to związane z obecnością grupy hydroksylowej w

anionie lub ze zmniejszonym stosunkiem rozmiaru kationu do anionu, gdzie stosunkowo większy niż zazwyczaj anion ma większą szansę wpłynąć na przestrzenne ułożenie cząsteczek cieczy jonowej.

Podobną analizę można przeprowadzić porównując ciecze jonowe o wspólnych kationach i różniące się anionami. Analiza ta jest jednak znacznie utrudniona, gdyż w odróżnieniu od kationów, przeważająca część anionów, w zbiorze danych, nie tworzy szeregów homologicznych. Ciecze [HMIM][pFAP] oraz [HMIM][eFAP] znacznie różnią się wartością FFV, które wynoszą odpowiednio 0,182 oraz 0,228. W tym przypadku wzrost rozmiarów łańcucha alkilowego z podstawnikami prowadzi do zmniejszenia wartości FFV. Szeregując malejąco ciecze o wspólnym anionie [BMIM] otrzymujemy kolejność [TfO] > [PF6] = [BF4] > [NTf<sub>2</sub>] > [TA] > [N(CN)<sub>2</sub>] i wartości FFV równe odpowiednio 0,193 > 0,191 = 0,191 > 0,184 > 0,182 > 0,178. Mimo, że aniony te znacznie różnią się strukturalnie, wartości FFV są do siebie zbliżone i zawierają wartości stanowiące 15% całkowitego zakresu. Wyniki te wyraźnie wskazują, że w większości przypadków wpływ anionu na wolną objętość cieczy jonowej jest znikomy.

#### 4.2.2. Modele QSPR

W Tabeli 4.5 umieszczono uszeregowaną według istotności statystycznej szczegółową listę deskryptorów na bazie których zbudowano trzy analizowane modele QSPR wraz z nich definicjami, kategoriami, do których należą oraz wartością parametru *t*, natomiast w Tabeli 4.6 umieszczono współczynniki regresji deskryptorów.

| Model                           | Deskryptor  | Definicja                                                                 | Kategoria              | Parametr <i>t</i> |
|---------------------------------|-------------|---------------------------------------------------------------------------|------------------------|-------------------|
| <b>FFV jako zmienna zależna</b> | nN+_cat     | Liczba dodatnio naładowanych atomów N                                     | Grupy funkcyjne        | -7.7              |
|                                 | R4i_cat     | Autokorelacja R w odległości 4/ ważona po potencjale jonizacji            | Deskryptory GETAWAY    | 4.21              |
|                                 | DLS_04_cat  | Zmodyfikowany indeks lekopodobności Chena (7 zasad)                       | Lekopodobność          | 3.17              |
|                                 | C-042_cat   | X--CH..X                                                                  | Fragmenty strukturalne | 2.45              |
|                                 | nHAcc_cat   | Liczba akceptorów wiązania wodorowego (atomy N, O, F)                     | Grupy funkcyjne        | -2.17             |
|                                 | R7s+_cat    | Maksymalna autokorelacja R w odległości 7/ ważona po stanie wewnętrznym   | Deskryptory GETAWAY    | 1.55              |
|                                 |             |                                                                           |                        |                   |
| <b>Model standardowy</b>        | B06[F-F]_an | Obecność/brak dwóch atomów fluoru w odległości topologicznej 6 od siebie  | Pary atomów 2D         | -7.13             |
|                                 | R7i_cat     | Autokorelacja R w odległości 7/ ważona po potencjale jonizacji            | Deskryptory GETAWAY    | -6.66             |
|                                 | R7i+_cat    | Maksymalna autokorelacja R w odległości 7/ ważona po potencjale jonizacji | Deskryptory GETAWAY    | 5.23              |
|                                 | R3i_cat     | Autokorelacja R w odległości 3/ ważona po potencjale jonizacji            | Deskryptory GETAWAY    | 4.98              |
|                                 | C-042_cat   | X--CH..X                                                                  | Fragmenty strukturalne | -4.86             |

|                  |                  |                                                                              |                     |       |
|------------------|------------------|------------------------------------------------------------------------------|---------------------|-------|
|                  | CATS3D_06_AL_cat | Miejsce akceptorowe oraz lipofilowe w odległości 6-7 Å od siebie             | CATS 3D             | 1.14  |
|                  |                  |                                                                              |                     |       |
| <b>Model FFV</b> | F01[C-O]_an      | Częstość występowania pary atomów C-O w odległości topologicznej 1 od siebie | Pary atomów 2D      | 6.88  |
|                  | R7i+_cat         | Maksymalna autokorelacja R w odległości 7/ ważona po potencjale jonizacji    | Deskryptory GETAWAY | 3.99  |
|                  | B06[F-F]_an      | Obecność/brak dwóch atomów fluoru w odległości topologicznej 6 od siebie     | Pary atomów 2D      | -3.71 |
|                  | F02[C-C]_an      | Częstość występowania pary atomów C-C w odległości topologicznej 2 od siebie | Pary atomów 2D      | -2.41 |
|                  | FFV              | Frakcyjna wolna objętość                                                     | -                   | -2.12 |
|                  | DLS_04_an        | Zmodyfikowany indeks lekopodobności Chena (7 zasad)                          | Lekopodobność       | -1.73 |

**Tabela 4. 5.** Deskryptory w poszczególnych modelach wraz z ich definicjami, kategorią, do której należą oraz wartością parametru t

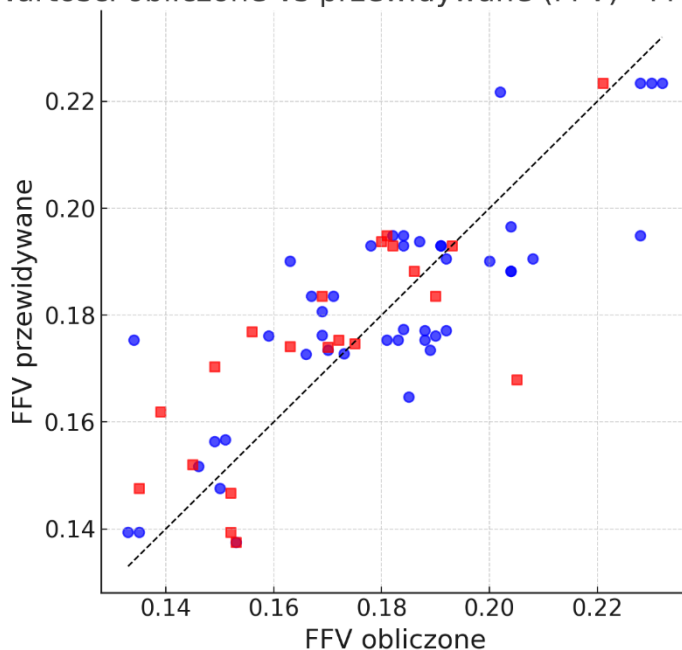
| Model                           | Deskryptor       | Współczynnik regresji |
|---------------------------------|------------------|-----------------------|
| <b>FFV jako zmienna zależna</b> | nN+_cat          | -0,074901             |
|                                 | R4i_cat          | 0,019936              |
|                                 | DLS_04_cat       | 0,041305              |
|                                 | C-042_cat        | 0,014517              |
|                                 | nHAcc_cat        | -0,004150             |
|                                 | R7s+_cat         | 0,033659              |
|                                 | Stała            | 0,163766              |
|                                 |                  |                       |
| <b>Model standardowy</b>        | B06[F-F]_an      | -4038,33              |
|                                 | R7i_cat          | -6551,86              |
|                                 | R7i+_cat         | 69035,37              |
|                                 | R3i_cat          | 2649,94               |
|                                 | C-042_cat        | -3336,66              |
|                                 | CATS3D_06_AL_cat | 534,89                |
|                                 | Stała            | 4451,56               |
|                                 |                  |                       |
| <b>Model FFV</b>                | F01[C-O]_an      | 3408,3                |
|                                 | R7i+_cat         | 38863,5               |
|                                 | B06[F-F]_an      | -2498,2               |
|                                 | F02[C-C]_an      | -400,4                |
|                                 | FFV              | -21886,5              |
|                                 | DLS_04_an        | -2864,1               |
|                                 | Stała            | 9838,2                |
|                                 |                  |                       |

**Tabela 4. 6.** Współczynniki regresji deskryptorów w modelach

#### 4.2.2.1. FFV jako zmienna zależna

Pierwszy zbudowany model miał, jak już wspomniano, służyć do przewidywania frakcyjnej wolnej objętości. Ze względu na długi czas obliczeń z wykorzystaniem dynamiki molekularnej wymaganych do wyznaczenia wartości FFV dla jednej cieczy, wynoszący nawet do 36 godzin na superkomputerze Tryton, zdecydowano się stworzyć model, który na podstawie wyników obliczonych dla 73 cieczy jonowych pozwoli z dużą dokładnością przewidywać wartości FFV dla nieuwzględnionych w nim cieczy. Model zbudowano z sześciu deskryptorów. Co ciekawe, wszystkie należały do deskryptorów powiązanych z kationem, co stanowi kolejne potwierdzenie wyników analizy przedstawionej w poprzednim podrozdziale, w którym stwierdzono, że wpływ anionów na wolną objętość cieczy jonowych jest zaniedbywalny. Najbardziej istotny statystycznie deskryptor anionowy znajdował się podczas modelowania na jedenastym miejscu, co oznacza, że dziesięć deskryptorów kationowych miało większy wpływ na wartość FFV od niego. Parametry statystyczne uzyskanego modelu spełniają założone minimalne wymagania co do wartości, jednak na tle pozostałych modeli nie są satysfakcjonujące. Współczynnik determinacji  $R^2$  oraz oba parametry walidacji  $Q^2$  oraz  $R^2_{pred}$  są najniższe spośród wszystkich zbudowanych na potrzeby rozprawy doktorskiej modeli i wynoszą odpowiednio 0,683, 0,501 oraz 0,560. Zdolność do wiarygodnego przewidywania wartości FFV cieczy spoza zbioru uczącego jest zdecydowanie poniżej oczekiwań, a jego zastosowanie jest ograniczone. Model ten wyjaśnia jednak niemal 70% zmienności w zbiorze uczącym, dzięki czemu może zostać z pewną dozą ostrożności zinterpretowany i wykorzystany do analizy i opisanie czynników wpływających na wartość wolnej objętości. Na Rysunku 4.7 przedstawiono zależność pomiędzy przewidywanymi a eksperymentalnymi (obliczonymi z wykorzystaniem dynamiki molekularnej) wartościami FFV. Dane znajdują się Tabeli A5 w Materiałach uzupełniających.

Wartości obliczone vs przewidywane (FFV) - FFV jako Y

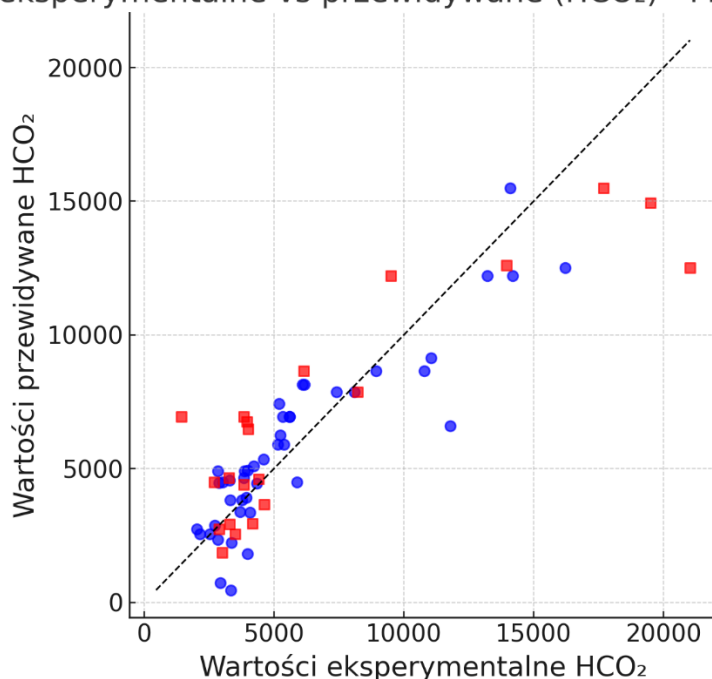


**Rysunek 4. 7.** Wykresy zależności wartości przewidywanej FFV od wartości eksperymentalnej dla modelu FFV jako zmienna zależna. Kolor niebieski – zbiór uczący, kolor czerwony – zbiór testowy

#### 4.2.2.2. Model standardowy

W modelu służącym do przewidywania wartości  $H_{CO_2}$  bez uwzględniania parametru FFV znalazło się sześć deskryptorów. Zdecydowana większość z nich była powiązana z kationami, a jedynym deskryptorem anionowym był dwuwymiarowy B06[F-F], który okazał się być najistotniejszym statystycznie spośród wszystkich. Pojedyncze deskryptory 3D okazały się być mniej wpływowe, co jest zaskakujące biorąc pod uwagę ich większą dokładność związaną ze wspólną optymalizacją jonów. Z drugiej strony, jednak, w modelu znajdują się cztery deskryptory 3D i jedynie dwa 2D. Parametry statystyczne modelu wynosiły odpowiednio 0,795, 0,697 oraz 0,747 dla  $R^2$ ,  $Q^2$  oraz  $R^2_{pred}$ . Warto zauważyć, że każdy z parametrów okazał się niższy, niż w przypadku modelu zbudowanego na podstawie cieczy jonowych optymalizowanych osobno. Bezpośrednie porównanie jest jednak na tym etapie niemożliwe ze względu na zastosowanie innej metody podziału zbioru danych za zbiór uczący i testowy (podział według zmiennej X zamiast według zmiennej Y) oraz nieskorzystanie z algorytmu genetycznego oraz brak weryfikacji interpretowalności deskryptorów i testowania wymian skorelowanych ze sobą parametrów. Na Rysunku 4.8 przedstawiono zależność pomiędzy przewidywanymi a eksperymentalnymi (obliczonymi z wykorzystaniem dynamiki molekularnej) wartościami  $H_{CO_2}$  w modelu standardowym. Dane znajdują się Tabeli A6 w Materiałach uzupełniających.

### Wartości eksperymentalne vs przewidywane ( $\text{HCO}_2$ ) - Model standardowy

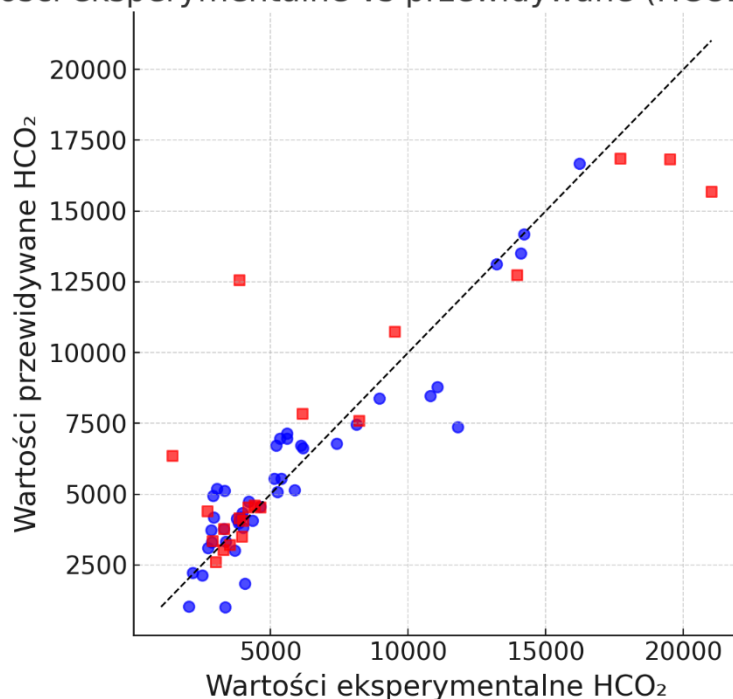


**Rysunek 4. 8.** Wykresy zależności wartości przewidywanej  $\text{HCO}_2$  od wartości eksperymentalnej dla modelu standardowego. Kolor niebieski – zbiór uczący, kolor czerwony – zbiór testowy

#### 4.2.2.3. Model FFV

FFV zostało wykorzystane podczas modelowania jako równorzędny z pozostałymi deskryptorami parametr, a proces został przeprowadzony w identyczny sposób, jak w przypadku modelu standardowego. Od samego początku FFV była identyfikowana jako jedna z istotniejszych statystycznie zmiennych, a na żadnym etapie nie dokonano ingerencji i nie wymuszono włączenia go do puli aktywnych deskryptorów. Ostatecznie wraz z pięcioma innymi zmiennymi trafił do końcowego modelu. Model ten, uzupełniony o obliczony z pomocą dynamiki molekularnej parametr FFV, wykazuje znaczące różnice w porównaniu z modelem standardowym. Największa rozbieżność zauważalna jest w zestawie deskryptorów. Cztery spośród pięciu powiązane są z anionami, a tylko jeden,  $\text{R7i}^+$ , z kationem. Najistotniejszym statystycznie deskryptorem jest  $\text{F01}[\text{C-O}]$ , a na drugim miejscu znajduje wspomniany już  $\text{R7i}^+$ , który pojawia się również w modelu standardowym jako trzeci pod względem wpływu (Tabela 4.5). Parametry statystyczne modelu uwzględniającego frakcyjną wolną objętość są wyższe w porównaniu do modelu zbudowanego bez FFV i wynoszą odpowiednio 0,874, 0,817 oraz 0,791 dla  $\text{R}^2$ ,  $\text{Q}^2$  oraz  $\text{R}^2_{\text{pred}}$ . Na Rysunku 4.9 przedstawiono zależność pomiędzy przewidywanymi a eksperymentalnymi (obliczonymi z wykorzystaniem dynamiki molekularnej) wartościami  $\text{HCO}_2$  w modelu FFV. Dane znajdują się Tabeli A7 w Materiałach uzupełniających.

Wartości eksperymentalne vs przewidywane ( $\text{HCO}_2$ ) - Model FFV



**Rysunek 4. 9.** Wykresy zależności wartości przewidywanej  $\text{HCO}_2$  od wartości eksperymentalnej dla modelu FFV. Kolor niebieski – zbiór uczący, kolor czerwony – zbiór testowy

#### 4.2.2.4. Omówienie i porównanie modeli

Parametry statystyczne modelu służącego do przewidywania frakcyjnej wolnej objętości zostały uznane za niewystarczające, aby móc je zastosować do ilościowego przewidywania wartości FFV. Oznacza to, że wolna objętość cieczy jonowych jest właściwością zależną od wielu czynników, a model oparty na deskryptorach molekularnych nie jest w stanie wyjaśnić wszystkich interakcji zachodzących pomiędzy poszczególnymi cząsteczkami tworzących strukturę cieczy. W tym konkretnym przypadku deskryptory, które opisują cechy mikroskopowe, dotyczące pojedynczych cząstek, nie są w stanie zastąpić symulacji dynamiki molekularnej uwzględniających interakcje na większą skalę, i które, w porównaniu do deskryptorów można uznać za makroskopowe. Dynamika molekularna, mimo czasochłonności, większego skomplikowania, a także większego kosztu obliczeń, pozostaje jedyną rzetelną metodą obliczania wolnej objętości cieczy jonowych. W tym miejscu należy jednak wspomnieć, że istnieje prawdopodobieństwo, że wartość parametru FFV udałoby się przewidzieć z większą dokładnością wykorzystując nieliniowe metody modelowania QSPR. Zbudowanie modelu pomogło jednak udowodnić, że wolna objętość zależy głównie od rodzaju kationu cieczy jonowej, a wpływ anionu jest w większości przypadków pomijalny.

Bezpośrednie porównanie obu modeli służących do przewidywania wartości  $\text{HCO}_2$  pozwala wyciągnąć kilka ważnych wniosków. Dane wejściowe dla obu modeli były wspólne, a etapy

modelowania przeprowadzone identycznie w celu wyeliminowania dodatkowych zmiennych mogących mieć wpływ na ich na parametry oraz uniemożliwiających ich rzetelne porównanie.

Model FFV charakteryzował się wyższymi wartościami wszystkich parametrów statystycznych służących do oceny modeli. Ich porównanie przedstawiono poniżej w Tabeli 4.7.

| Model                    | $R^2$ | $Q^2$ | $R^2_{pred}$ |
|--------------------------|-------|-------|--------------|
| FFV jako zmienna zależna | 0,683 | 0,501 | 0,560        |
| Model standardowy        | 0,795 | 0,697 | 0,747        |
| Model FFV                | 0,874 | 0,817 | 0,791        |

**Tabela 4. 7.** Porównanie parametrów statystycznych modeli

Model FFV uzyskał wartość współczynnika determinacji  $R^2$  na poziomie 0,874 w porównaniu do 0,795 w przypadku modelu standardowego. Oznacza to, że uwzględnienie wpływu frakcyjnej wolnej objętości na rozpuszczalność  $CO_2$  pozwoliło na dokładniejsze wyjaśnienie wariacji występującej w zbiorze danych. W konsekwencji interpretacja deskryptorów oraz płynące z niej wnioski mogą być traktowane z większą pewnością. Parametry walidacji wewnętrznej i zewnętrznej również wskazują na wyższość modelu FFV. Największa zauważalna różnica dotyczy  $Q^2$ , gdzie omawiany model uzyskał wynik 0,817 w porównaniu do 0,697 dla modelu standardowego. Uwzględnienie wpływu frakcyjnej wolnej objętości sprawiło, że otrzymany model był mniej podatny na zmiany w zbiorze uczącym i może zostać z powodzeniem zastosowany do szerokiej gamy cieczy jonowych. W konsekwencji, wartości  $H_{CO_2}$  obliczone dla cieczy nieuwzględnionych podczas budowy modelu, a więc znajdujących się w zbiorze testowym, były bardziej zbliżone do wartości zmierzonych eksperymentalnie w przypadku modelu FFV. Oznacza to, że model uwzględniający wpływ wolnej objętości jest w stanie oszacować rozpuszczalność  $CO_2$  w cieczach jonowych z większą dokładnością.

Porównując deskryptory, które trafiły do obu modeli, nie sposób nie zwrócić uwagi na fakt, że uwzględnienie FFV w modelowaniu przyczyniło się do odwrócenia stosunku liczby deskryptorów kationowych do anionowych w modelach. W modelu standardowym pięć na sześć deskryptorów dotyczyło kationów, natomiast w modelu FFV był to już tylko jeden z pięciu. Co ciekawe, w obu modelach największym wpływem wykazały się deskryptory anionowe. Można założyć, że parametr FFV, który jest zależny w przeważającej części przypadków głównie od kationu, uwzględnił ten wpływ również podczas modelowania  $H_{CO_2}$ , co w konsekwencji umożliwiło uwidocznienie w większym stopniu wpływu cech strukturalnych anionów na rozpuszczalność ditlenku węgla w cieczach jonowych.

#### 4.2.3. Interpretacja deskryptorów

Ze względu na niesatysfakcjonujące wyniki walidacji modelu służącego do przewidywania FFV zdecydowano się nie analizować szczegółowego wpływu deskryptorów budujących model. O ile

bezsprzecznym pozostaje fakt przeważającego wpływu kationów na wartość FFV, o tyle szczegółowa interpretacja deskryptorów oparta o ich istotność statystyczną w modelu mogłaby zostać poddana w wątpliwość. Z tego powodu poniższy podrozdział zawiera analizę obu modeli przewidujących wartość  $H_{CO_2}$  – modelu standardowego oraz modelu FFV.

#### 4.2.3.1. Model standardowy

Najbardziej istotnym, ale i jedynym związanym z anionami deskryptorem w modelu standardowym był B06[F-F], który podobnie jak opisywany już B02[F-F] jest deskryptorem jakościowym 2D i opisuje brak lub obecność dwóch atomów fluoru w odległości 6 atomów od siebie po przedstawieniu cząsteczki w postaci grafu.[84] Wartość  $H_{CO_2}$  maleje, gdy szukana para atomów fluoru zostanie odnaleziona, a B06[F-F] przyjmuje wartość równą 1. Obecność atomów fluoru została po raz kolejny zidentyfikowana jako ważny parametr zwiększający rozpuszczalność  $CO_2$  w cieczach jonowych.

Trzy z sześciu deskryptorów w modelu należą do tej samej kategorii 3D R-GETAWAY i są deskryptorami związanymi z kationami. Deskryptory te są obliczane na podstawie tworzonej macierzy wpływu/odległości i zależą od wartości wybranej właściwości atomu (wpływu) oraz odległości topologicznej pomiędzy atomami.[84] W przypadku omawianego modelu analizowaną właściwością jest energia jonizacji atomu, a dystans obliczany jest dla atomów w odległości 3 i 7 węzłów od siebie. Składowe deskryptorów (indeksów autokorelacji)  $R_{7i}$  oraz  $R_{7i+}$  obliczane są w bardzo podobny sposób, różnica polega jednak na ich końcowej wartości wykorzystanej w obliczanie wartości deskryptorów. W pierwszym przypadku wykorzystuje się sumę obliczonych indeksów autokorelacji, natomiast w drugim uwzględnia jedynie wybraną spośród nich najwyższą wartość.[246] Deskryptory te są bardziej wrażliwe na zmiany w układzie przestrzennym kationów, niż trzeci z omawianej grupy, czyli  $R_{3i}$ . Wynika to z większej odległości pomiędzy analizowanymi atomami, a więc i większych różnic w ułożeniu łańcucha alkilowego oraz wszystkich podstawników. Z tego powodu optymalizacja związków przed modelowaniem QSPR powinna uwzględniać wszystkie interakcje pomiędzy cząsteczkami, o ile jest to możliwe. Wzrost wartości deskryptorów  $R_{3i}$  oraz  $R_{7i+}$  negatywnie wpływa na rozpuszczalność  $CO_2$ , podczas gdy przeciwną zależność można zaobserwować dla deskryptora  $R_{7i}$ . Zachowanie to można wytłumaczyć powiązaniem wartości  $R_{7i}$  ze wzrostem rozmiarów kationu - ponieważ stanowi on sumę obliczonych indeksów, wzrasta z każdym wystąpieniem pary atomów w odległości siedmiu atomów od siebie. Jego wartość jest więc większa dla cząsteczek o dużej objętości i posiadających długie łańcuchy alkilowe, a największa dla kationu [P66614] zbudowanego z najdłuższych łańcuchów w analizowanym zbiorze ( $R_{7i} = 1,42$ ). Nie można zapomnieć jednak o zależności wartości omawianych deskryptorów od energii jonizacji. Właściwość ta okazuje się być ważnym parametrem i wpływa na trzy z sześciu zmiennych w modelu. Obserwacja ta znajduje potwierdzenie w literaturze. W swojej pracy Liu i inni

analizując rozpuszczalność ditlenku węgla w cieczach jonowych wykazali, że potencjał elektryczny oraz energia jonizacji atomów mają wzajemnie uzupełniający się wpływ na proces absorpcji CO<sub>2</sub> przez ciecz.[247,248]

Kolejnym omawianym deskrytorem kationowym jest CATS3D\_06\_AL, który jest analogiczny do opisywanego już wcześniej deskryptora CATS3D\_04\_AA. Stanowi on ilościowe przedstawienie występowania w strukturze cząsteczki obszarów lipofilowych oraz akceptorów wiązań wodorowych w odległości 6-7 Å od siebie. Wartość H<sub>CO2</sub> wzrasta wraz ze wzrostem wartości tego deskryptora. Najwyższe wartości stwierdzono dla cząsteczek posiadających długie niepodstawione łańcuchy alkilowe, [HMIM] oraz [OMIM] wynoszące w obu przypadkach 2.

Ostatnim omawianym deskrytorem jest C-042, jedyny kationowy i ilościowy deskryptor z grupy 2D. Reprezentuje on liczbę fragmentów X-CH...X w cząsteczce, gdzie X oznacza heteroatom, „-” oznacza wiązanie aromatyczne lub zdelokalizowane, a „...” jedynie wiązanie aromatyczne. Układ ten w analizowanym zbiorze danych występuje głównie w pierścieniach imidazolowych, a jego obecność zwiększa rozpuszczalność CO<sub>2</sub>. W tym momencie należy jednak nadmienić, że związki oparte na pierścieniach imidazolowych stanowią znaczną część zbioru danych i istnieje możliwość, że deskryptor ten został zakwalifikowany jako statystycznie istotny jedynie ze względu na ich nadreprezentację.

#### 4.2.3.2. Model FFV

W modelu FFV pojawiły się dwa deskryptory omówione już wcześniej. Pierwszym z nich jest B06[F-F], który podobnie jak w modelu standardowym jest odpowiedzialny za zmniejszenie wartości H<sub>CO2</sub> w przypadku stwierdzenia obecności atomów fluoru w anionie. Drugi z nich, to jedyny deskryptor powiązany z kationami w omawianym modelu. R7i+ jest drugim co do ważności deskrytorem w modelu, a stała Henry’ego cieczy jonowej rośnie proporcjonalnie wraz z nim.

Najistotniejszym statystycznie deskrytorem jest F01[C-O], który ilościowo określa liczbę wiązań węgiel-tlen w cząsteczce. Wiązania te występują w anionach organicznych, takich jak [OAc] oraz [Lac], a ich obecność negatywnie wpływa na zdolność cieczy do rozpuszczania CO<sub>2</sub>. Wyższe wartości deskryptora są powiązane z występowaniem grup -OH w cząsteczce anionu, które już wcześniej zostały zidentyfikowane jako niepożądane.[84]

Kolejny deskryptor, F02[C-C] opisuje występowanie dwóch atomów węgla w odległości dwóch węzłów od siebie. Stała Henry’ego jest odwrotnie proporcjonalna do jego wartości. Najwyższe wartości obliczono dla anionów [bFAP] (9) oraz [pFAP] (6), co wskazuje na powiązanie tego deskryptora z obecnością perfluorowanych łańcuchów alkilowych, o których wiadomo, że zwiększają rozpuszczalność

CO<sub>2</sub> w cieczach jonowych. Z drugiej strony, występowanie wielu fragmentów -CF<sub>2</sub>-CF<sub>2</sub>- w strukturze wspomnianych anionów mogło zaważyć na wskazaniu tego deskryptora jako statystycznie istotnego.

Ostatnim z omawianych deskryptorów jest DLS\_04, który został zdefiniowany aby służyć do oceny lekopodobności cząsteczek przez Chena i innych[249]. DLS\_04 (*ang. Drug Like Score*) zależy od siedmiu warunków, z których każdy analizowany jest osobno, a ostateczna wartość stanowi sumę cząstkowych punktów zdobytych przez związek w każdym z siedmiu testów. Do warunków, jakie musi spełnić związek, aby uzyskać wysoki wynik testu na lekopodobność należą: obecność pięciu lub mniej donorów wiązania wodorowego w cząsteczce, od 2 do 10 akceptorów wiązania wodorowego, masa cząsteczkowa zawierająca się w zakresie 78 do 500 u oraz logP Moriguchiego w zakresie od -0,5 do 5. Pozostałe trzy warunki wymagają obliczenia dodatkowych deskryptorów. Pierwszy jest zależny od stosunku atomów węgla o hybrydyzacji sp<sup>3</sup> do sumy wszystkich atomów niebędących halogenkami lub wodorem, drugi opisuje stosunek liczby atomów wodoru do wszystkich pozostałych atomów w cząsteczce, a trzeci jest miarą nienasycenia związku i zależy od stosunku sumy liczby pierścieni 5-, 6- lub 7-członowych oraz wiązań podwójnych, potrójnych i aromatycznych do sumy atomów niezwiązanych żadnym wiązaniem z halogenkami lub wodorem. Im wyższa jest ostateczna wartość DLS\_04, tym bardziej lekopodobny jest związek chemiczny, a w kontekście analizy wpływu na wartość H<sub>CO2</sub>, ze wzrostem wartości deskryptora jej wartość maleje, co przekłada się na lepszą rozpuszczalność ditlenku węgla.

Na koniec warto zauważyć, że wszystkie deskryptory dotyczące anionów należały do kategorii 2D, które można obliczyć wykorzystując zapisy w formacie SMILES bez konieczności przeprowadzania optymalizacji trójwymiarowych struktur. Z drugiej strony, deskryptory kationowe należały w większości do kategorii 3D. Możliwość zastosowania wzorów SMILES jest z całą pewnością ułatwieniem oraz oszczędnością czasu. Z drugiej strony jedynym modelem, gdzie deskryptory należały do kategorii 2D, był model, w którym uwzględniono wartości FFV obliczone za pomocą dynamiki molekularnej.

#### 4.3. Analiza metod optymalizacji

Badania nad wpływem sposobu optymalizacji cząsteczek składały się z dwóch części. W pierwszej zbudowano 32 modele, które po wykluczeniu niespełniających wymagań statystycznych poddano analizie dotyczącej wzajemnych powiązań pomiędzy metodą optymalizacji struktur chemicznych cieczy jonowych a sposobem budowania modeli regresyjnych. W drugiej części przeanalizowano wpływ sposobu optymalizacji na wartości obliczanych deskryptorów molekularnych na przykładzie najczęściej występującego w całym zbiorze danych jonu – bis(trifluorometanosulfonylo)imidku, czyli anionu [NTf<sub>2</sub>].

#### 4.3.1. Zestawienie parametrów statystycznych otrzymanych modeli

|                        |               | Próżnia |     |    |     | Etanol |     |    |     |
|------------------------|---------------|---------|-----|----|-----|--------|-----|----|-----|
| <b>HCO<sub>2</sub></b> | <b>Razem</b>  | MLR     | PLS | LR | SVM | MLR    | PLS | LR | SVM |
|                        | <b>Osobno</b> | MLR     | PLS | LR | SVM | MLR    | PLS | LR | SVM |
| <b>FFV</b>             | <b>Razem</b>  | MLR     | PLS | LR | SVM | MLR    | PLS | LR | SVM |
|                        | <b>Osobno</b> | MLR     | PLS | LR | SVM | MLR    | PLS | LR | SVM |

**Rysunek 4. 10.** Wizualne porównanie parametrów statystycznych otrzymanych modeli.

Kolor zielony:  $R^2$ ,  $Q^2$  oraz  $R^2_{pred}$  mają wartości większe od 0,5

Kolor żółty: jeden z parametrów jest mniejszy od 0,5

Kolor czerwony: więcej niż jeden parametr jest mniejszy od 0,5

Na Rysunku 4.10 przedstawiono wizualne porównanie wartości parametrów statystycznych 32 zbudowanych modeli. Kolorem zielonym oznaczono modele, dla których każdy z parametrów  $R^2$ ,  $Q^2$  oraz  $R^2_{pred}$  miał wartość większą od 0,5. Kolorem żółtym oznaczono te, gdzie jeden z wymienionych parametrów miał wartość mniejszą od 0,5, a kolorem czerwonym te, gdzie więcej niż jeden parametr nie spełnił zakładanych kryteriów. W Tabeli 4.8 znajdują się szczegółowe wartości wszystkich analizowanych parametrów statystycznych dla omawianych modeli.

Jak nietrudno zauważyć, modele dla których wszystkie parametry statystyczne spełniały warunki opisane w rozdziale 3.5 (wartość większa od 0,5) udało się uzyskać jedynie dla zbioru danych służącego do przewidywania wartości  $H_{CO_2}$ . Dla zbioru FFV, co najmniej jeden, z parametrów zawsze okazywał się niższy od zakładanej wartości, co potwierdza obserwacje uzyskane podczas analizy wolnej objętości, opisaną w rozdziale 4.2. Nie udało się zbudować żadnego modelu LR dla zestawu  $H_{CO_2}$ , a także żadnego modelu SVM dla zestawu FFV. Z analizy wykluczono również wszystkie modele FFV dla zestawu danych zoptymalizowanych osobno w środowisku etanolu ze względu na ujemny współczynnik determinacji w walidacji zewnętrznej. W rezultacie otrzymano 21 modeli, które poddano analizie.

| <b>H<sub>CO2</sub></b> |                                   |                                 |                  |                      |                                    |                      |                                     |                         |             |
|------------------------|-----------------------------------|---------------------------------|------------------|----------------------|------------------------------------|----------------------|-------------------------------------|-------------------------|-------------|
| <b>Model</b>           | <b>0 - próżnia<br/>1 - etanol</b> | <b>0 - osobno<br/>1 - razem</b> | <b>Parametry</b> | <b>R<sup>2</sup></b> | <b>R<sup>2</sup><sub>adj</sub></b> | <b>Q<sup>2</sup></b> | <b>R<sup>2</sup><sub>pred</sub></b> | <b>Y<sub>rand</sub></b> | <b>RMSE</b> |
| MLR                    | 0                                 | 0                               | 4x 2D            | 0.9423               | 0.9345                             | 0.8975               | 0.7125                              | 0.0918                  | 1114.9200   |
| PLS                    | 0                                 | 0                               | 1x 3D            | 0.9087               | 0.8970                             | 0.8660               | 0.7370                              | 0.1069                  | 1201.6100   |
| SVM                    | 0                                 | 0                               |                  | 0.8668               | 0.8497                             | 0.8579               | 0.7873                              | 0.0968                  | 1501.7000   |
| MLR                    | 1                                 | 0                               | 2x 2D            | 0.9126               | 0.9011                             | 0.8716               | 0.7311                              | 0.0976                  | 1221.1700   |
| PLS                    | 1                                 | 0                               | 3x 3D            | 0.9126               | 0.9011                             | 0.8627               | 0.7827                              | 0.0905                  | 1247.2500   |
| SVM                    | 1                                 | 0                               |                  | 0.7343               | 0.7016                             | 0.8716               | 0.7911                              | 0.1054                  | 2120.8500   |
| MLR                    | 0                                 | 1                               | FFV              | 0.9472               | 0.9399                             | 0.9117               | 0.9018                              | 0.1511                  | 867.9400    |
| PLS                    | 0                                 | 1                               | 3x 2D            | 0.9446               | 0.9369                             | 0.8724               | 0.8529                              | 0.1280                  | 1516.2600   |
| SVM                    | 0                                 | 1                               | 1x 3D            | 0.5418               | 0.4782                             | 0.9117               | 0.4482                              | 0.1437                  | 2556.7100   |
| MLR                    | 1                                 | 1                               | FFV              | 0.9147               | 0.9021                             | 0.8778               | 0.8655                              | 0.1360                  | 1085.8900   |
| PLS                    | 1                                 | 1                               | 0x 2D            | 0.8608               | 0.8425                             | 0.8053               | 0.8839                              | 0.1406                  | 1217.7900   |
| SVM                    | 1                                 | 1                               | 4x 3D            | 0.6863               | 0.6450                             | 0.8092               | 0.7844                              | 0.1329                  | 1844.6300   |
| <b>FFV</b>             |                                   |                                 |                  |                      |                                    |                      |                                     |                         |             |
| <b>Model</b>           | <b>0 - próżnia<br/>1 - etanol</b> | <b>0 - osobno<br/>1 - razem</b> | <b>Parametry</b> | <b>R<sup>2</sup></b> | <b>R<sup>2</sup><sub>adj</sub></b> | <b>Q<sup>2</sup></b> | <b>R<sup>2</sup><sub>pred</sub></b> | <b>Y<sub>rand</sub></b> | <b>RMSE</b> |
| MLR                    | 0                                 | 0                               | 5x 2D            | 0.7907               | 0.7568                             | 0.7156               | 0.3672                              | 0.1452                  | 0.0110      |
| PLS                    | 0                                 | 0                               | 1x 3D            | 0.7867               | 0.7521                             | 0.7106               | 0.3714                              | 0.1508                  | 0.0117      |
| LR                     | 0                                 | 0                               |                  | 0.7655               | 0.7274                             | 0.6827               | 0.3821                              | 0.1544                  | 0.0113      |
| MLR                    | 1                                 | 0                               | 3x 2D            | 0.8853               | 0.8651                             | 0.8277               | -0.9788                             | 0.1327                  | 0.0084      |
| PLS                    | 1                                 | 0                               | 3x 3D            | 0.8810               | 0.8599                             | 0.7657               | -0.3591                             | 0.1452                  | 0.0083      |
| LR                     | 1                                 | 0                               |                  | 0.8880               | 0.8682                             | 0.8214               | -0.5350                             | 0.1478                  | 0.0112      |
| MLR                    | 0                                 | 1                               | 6x 2D            | 0.8795               | 0.8594                             | 0.8385               | 0.2093                              | 0.1361                  | 0.0077      |
| PLS                    | 0                                 | 1                               | 0x 3D            | 0.7984               | 0.7665                             | 0.7268               | 0.3856                              | 0.1289                  | 0.0090      |
| LR                     | 0                                 | 1                               |                  | 0.7839               | 0.7498                             | 0.7133               | 0.2812                              | 0.1396                  | 0.0084      |
| MLR                    | 1                                 | 1                               | 0x 2D            | 0.7838               | 0.7406                             | 0.6969               | 0.3570                              | 0.1679                  | 0.0150      |
| PLS                    | 1                                 | 1                               | 6x 3D            | 0.7823               | 0.7388                             | 0.6905               | 0.4234                              | 0.1912                  | 0.0156      |
| LR                     | 1                                 | 1                               |                  | 0.6366               | 0.5774                             | 0.5407               | 0.3382                              | 0.1805                  | 0.0147      |

**Tabela 4. 8.** Parametry statystyczne omawianych modeli. W kolumnie Parametry podano, do jakich kategorii należały deskryptory, które znalazły się w modelach – 2D, 3D oraz czy wykorzystano wartości FFV jako zmienne niezależne.

Na rysunkach 4.11-4.14 przedstawiono wykresy zależności wartości przewidywanych od wartości eksperymentalnych dla każdego z modeli.

Rysunki 4.11a, 4.11b oraz 4.11c przedstawiają wykresy zależności H<sub>CO2</sub> dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych osobno w próżni. W modelach tych cztery deskryptory należą do kategorii 2D, a jeden do 3D. Uwagę zwraca charakterystyczna pozioma linia dla wartości przewidywanej ok. 4400, którą tworzą związku zarówno z zestawu uczącego, jak i testowego. Związane jest to z charakterem deskryptorów, z których zbudowano modele. Są one w większości deskryptorami o niewielkiej rozpiętości, B04[C-C]\_cat przyjmuje jedynie wartości 0 lub 1, F01[C-O]\_an wyrażany jest w liczbach całkowitych od 0 do 3, a G(N...O)\_cat przyjmuje wartość inną od 0 jedynie dla 15% związków. Mimo bardzo niewielkiej różnicy pomiędzy wartościami przewidzianymi dla znacznej części zbioru danych, wspomniane 15% związków odpowiada za bardzo wysoką wartość R<sup>2</sup> każdego z

modeli – wartości te mają wartość różną od 0 dla związków o najwyższych eksperymentalnych wartościach  $H_{CO_2}$ , co przekłada się na dobre przewidywanie wartości skrajnych.

Rysunki 4.11d, 4.11e oraz 4.11f przedstawiają wykresy zależności  $H_{CO_2}$  dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych osobno w środowisku etanolu. W modelach tych dwa deskryptory należą do kategorii 2D, a trzy do 3D. Zaobserwować można tutaj wyraźne zaniżanie wartości przewidywanych przez model SVM – trend, który prawdziwy jest również dla modeli SVM dla cieczy zoptymalizowanych wspólnie.

Rysunki 4.12a, 4.12b oraz 4.12c przedstawiają wykresy zależności  $H_{CO_2}$  dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych razem w próżni. W modelach tych trzy deskryptory należą do kategorii 2D, jeden do 3D, a dodatkowo jako jedna ze zmiennych występuje również parametr FFV. Model MLR uzyskał najwyższe wartości  $R^2$ ,  $Q^2$  oraz  $R^2_{pred}$  ze wszystkich modeli w badaniu. Ponownie uwagę zwraca zaniżanie wartości  $H_{CO_2}$  dla najwyższych wartości eksperymentalnych przez model SVM.

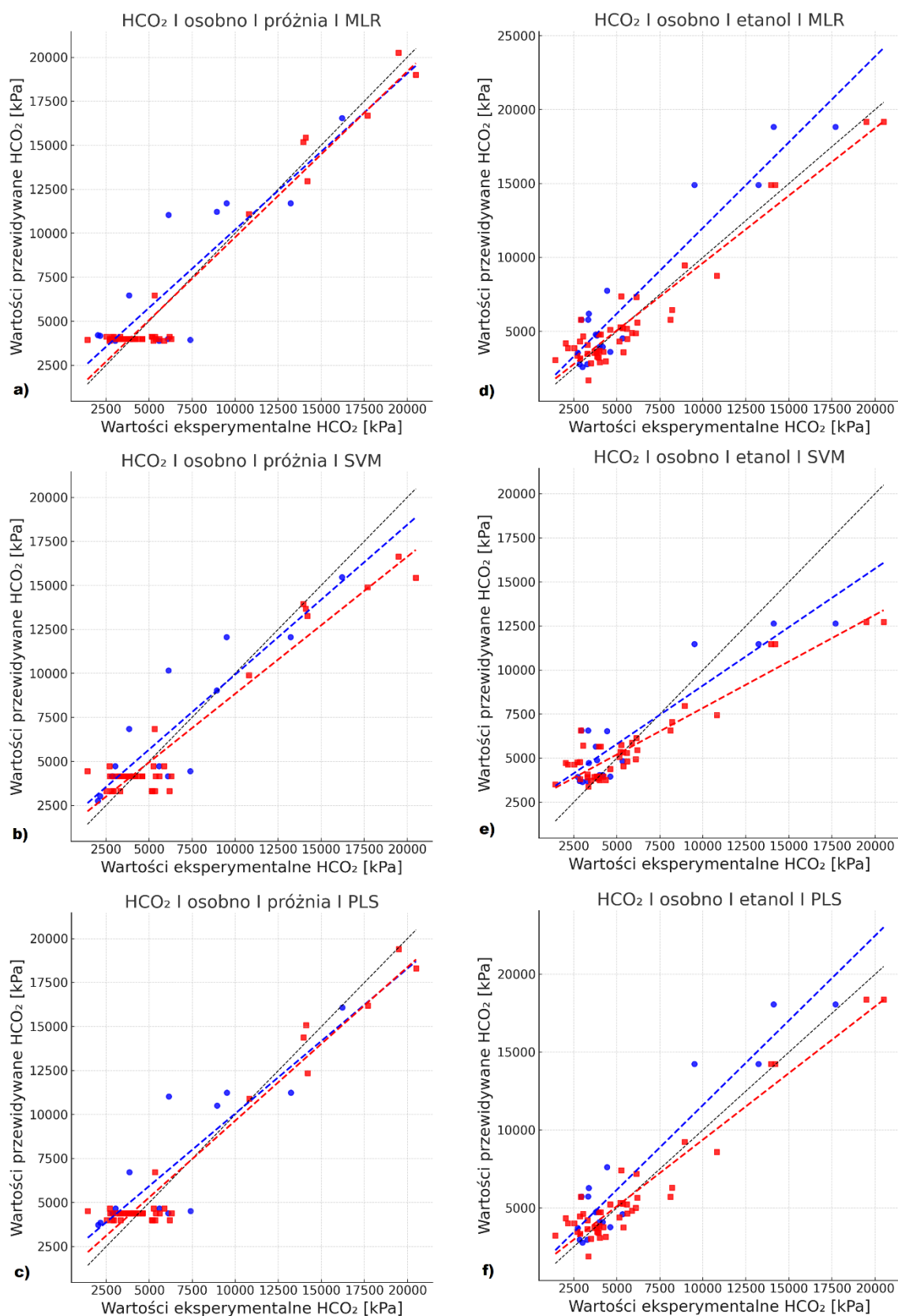
Rysunki 4.12d, 4.12e oraz 4.12f przedstawiają wykresy zależności  $H_{CO_2}$  dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych razem w środowisku etanolu. W modelach tych cztery deskryptory należą do kategorii 3D, pojawia się parametr FFV, a żaden z deskryptorów nie należy do kategorii 2D. Z tego powodu modele te są najtrudniejsze do zinterpretowania.

Rysunki 4.13a, 4.13b oraz 4.13c przedstawiają wykresy zależności FFV dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych osobno w próżni. W modelach tych pięć deskryptorów należy do kategorii 2D, a tylko jeden do 3D. Co ciekawe pomimo gorszych parametrów statystycznych w porównaniu do modeli  $H_{CO_2}$  widać, że zbiór danych charakteryzuje się bardziej równomiernym rozkładem wartości analizowanej właściwości. Wszystkie modele charakteryzują się zbliżonymi parametrami statystycznymi.

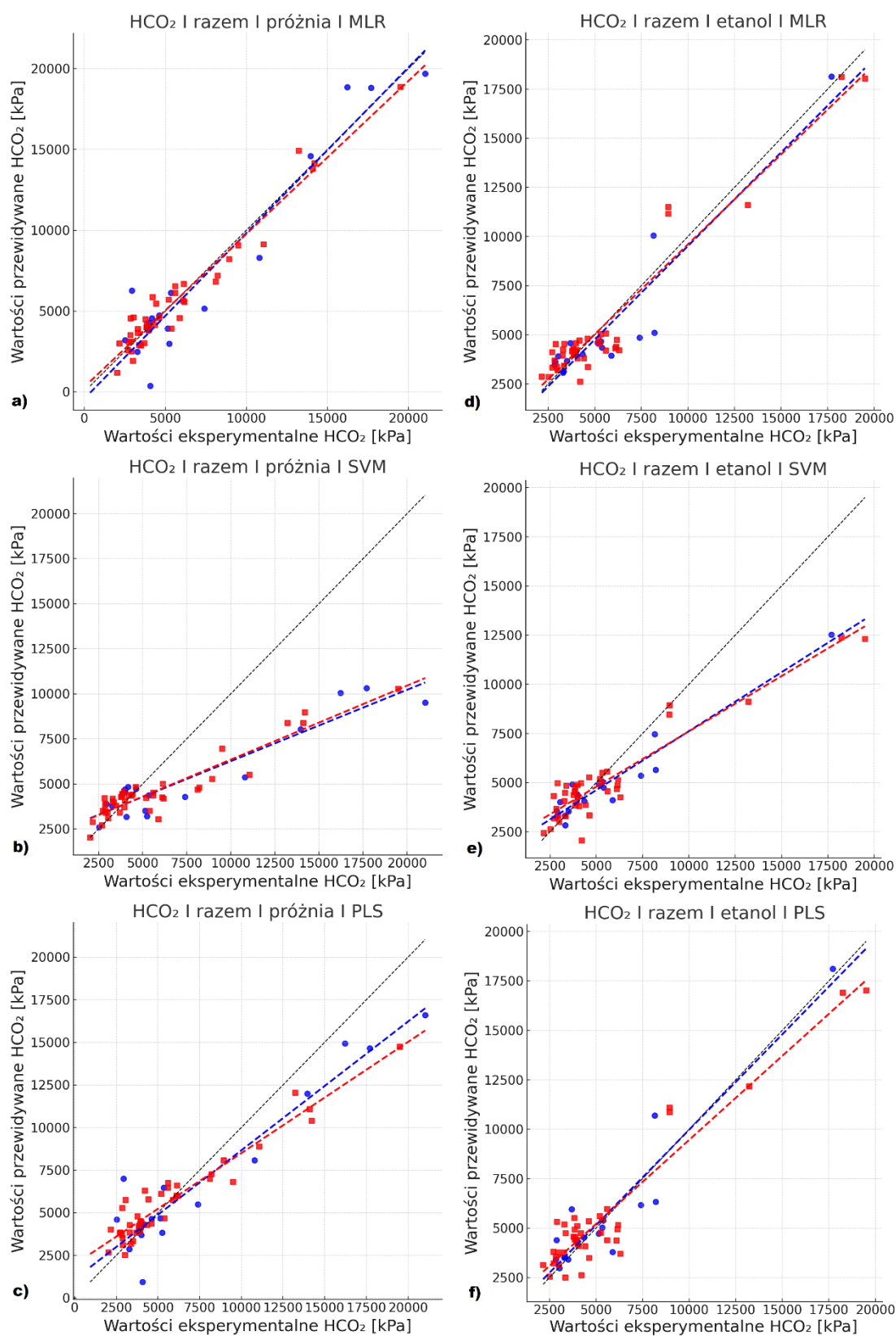
Rysunki 4.14a, 4.14b oraz 4.14c przedstawiają wykresy zależności FFV dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych razem w próżni. W modelach tych wszystkie sześć deskryptorów należy do kategorii 2D. Uwagę zwraca systematycznie największa w badaniu różnica pomiędzy dopasowaniem do zbioru uczącego a dopasowaniem do zbioru testowego, którą szczególnie widać w modelu MLR.

Rysunki 4.14d, 4.14e oraz 4.14f przedstawiają wykresy zależności FFV dla modeli zbudowanych na podstawie cieczy jonowych zoptymalizowanych razem w środowisku etanolu. W modelach tych wszystkie sześć deskryptorów należy do kategorii 3D. Zaobserwować można na nich charakterystyczną cechę nieobecną w żadnej innej grupie – niemal wszystkie wartości przewidywane dla zbioru uczącego są zaniżone.

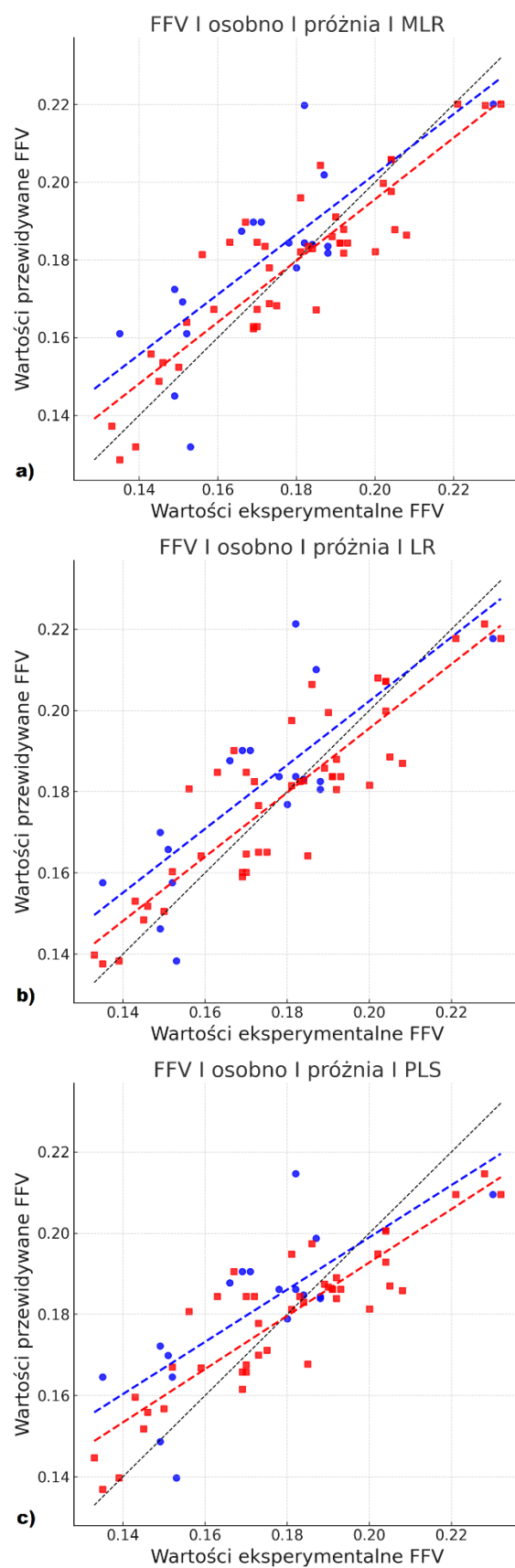
Porównując deskryptory we wszystkich modelach należy zauważyć, że w przypadku optymalizacji w środowisku etanolu przeważająca większość deskryptorów należy do kategorii 3D. Celem optymalizacji w środowisku o nadanej stałej dielektrycznej jest uwzględnienie interakcji pomiędzy cząsteczkami w trójwymiarowym układzie odniesienia, a konsekwencją tego działania jest podczas modelowania wzrost istotności deskryptorów 3D, które również zależne są od interakcji w trójwymiarowym układzie odniesienia wewnątrz samej cząsteczki. Negatywną konsekwencją takiego rozwiązania jest niestety wzrost trudności interpretacji otrzymanych modeli. Co dosyć zaskakujące, deskryptory obliczone dla cząsteczek zoptymalizowanych w ten sposób mają, średnio, gorsze parametry statystyczne jak wynika z zestawienia na Rysunku 4.10.



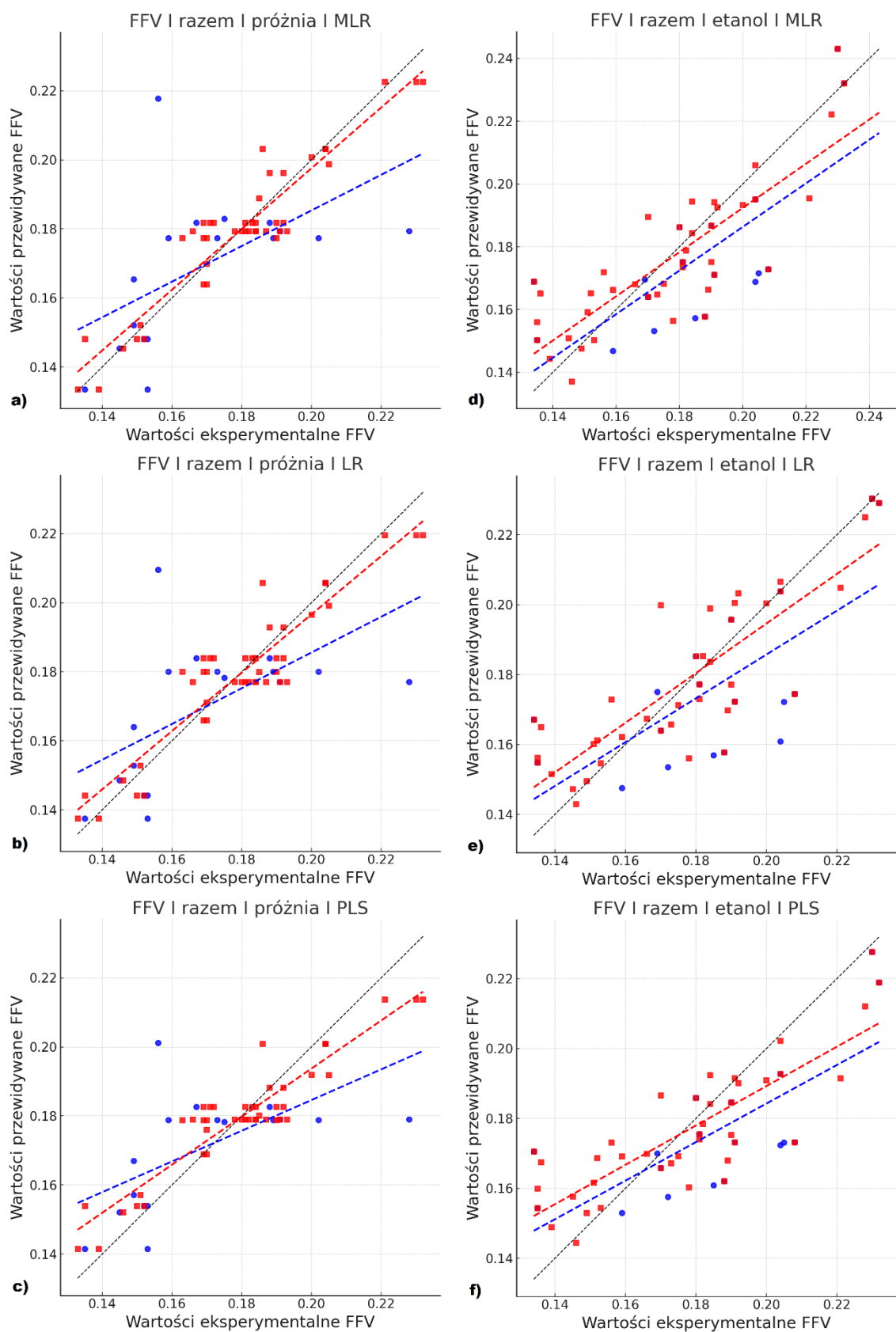
**Rysunek 4. 11.** Wykresy zależności wartości przewidywanej  $\text{HCO}_2$  od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych: a) osobno w próżni metodą MLR, b) osobno w próżni metodą SVM, c) osobno w próżni metodą PLS, d) osobno w etanolu metodą MLR, e) osobno w etanolu metodą SVM, f) osobno w etanolu metodą PLS. Kolorem niebieskim oznaczono zbiór uczący, kolorem czerwonym zbiór testowy.



**Rysunek 4. 12.** Wykresy zależności wartości przewidywanej  $\text{HCO}_2$  od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych: a) razem w próżni metodą MLR, b) razem w próżni metodą SVM, c) razem w próżni metodą PLS, d) razem w etanolu metodą MLR, e) razem w etanolu metodą SVM, f) razem w etanolu metodą PLS. Kolorem niebieskim oznaczono zbiór uczący, kolorem czerwonym zbiór testowy.



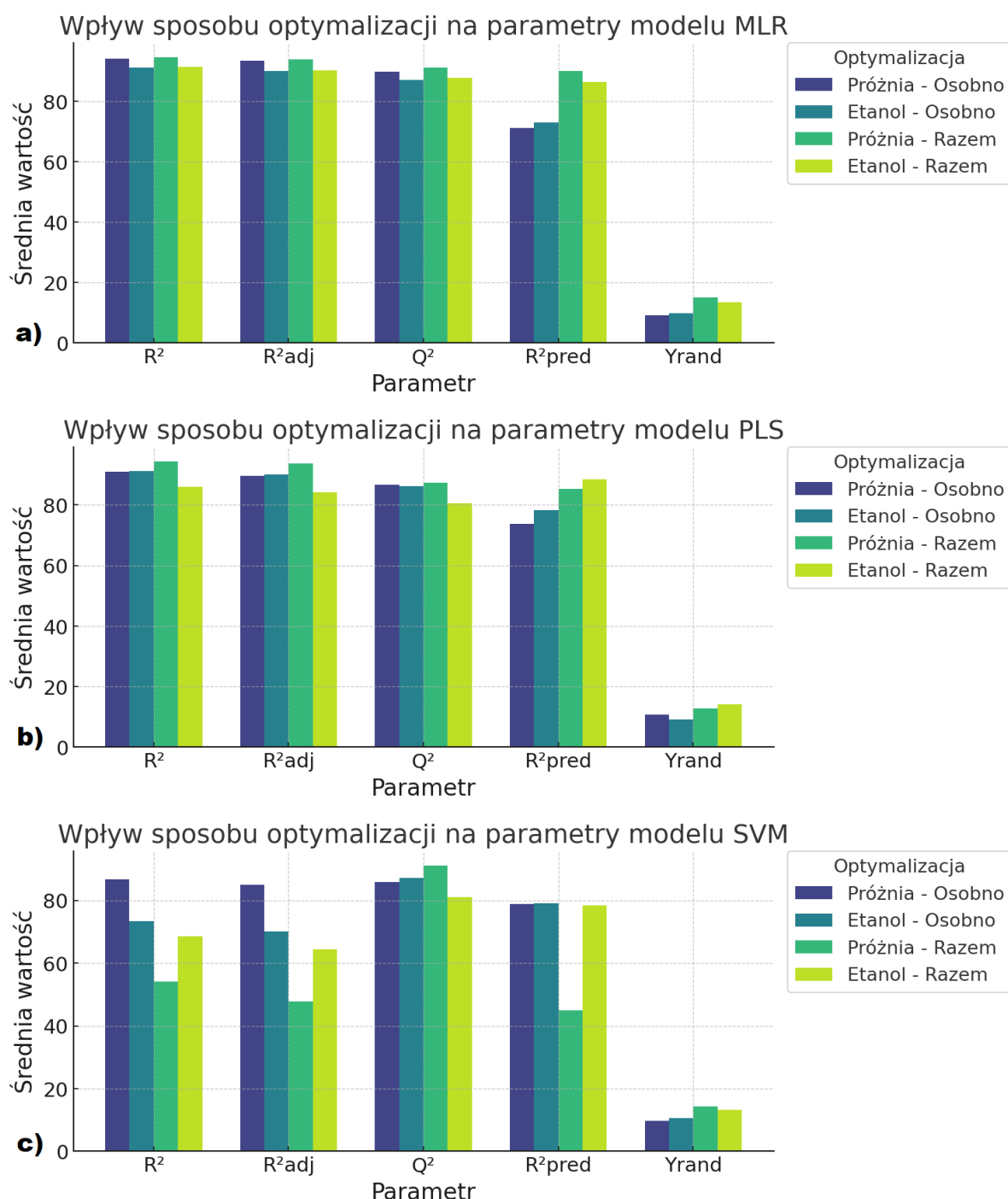
**Rysunek 4. 13.** Wykresy zależności wartości przewidywanej FFV od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych: a) osobno w próżni metodą MLR, b) osobno w próżni metodą LR, c) osobno w próżni metodą PLS. Kolorem niebieskim oznaczono zbiór uczący, kolorem czerwonym zbiór testowy.



**Rysunek 4. 14.** Wykresy zależności wartości przewidywanej FFV od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych: a) razem w próżni metodą MLR, b) razem w próżni metodą LR, c) razem w próżni metodą PLS, d) razem w etanolu metodą MLR, e) razem w etanolu metodą LR, f) razem w etanolu metodą PLS. Kolorem niebieskim oznaczono zbiór uczący, kolorem czerwonym zbiór testowy.

#### 4.3.2. Porównanie metod modelowania

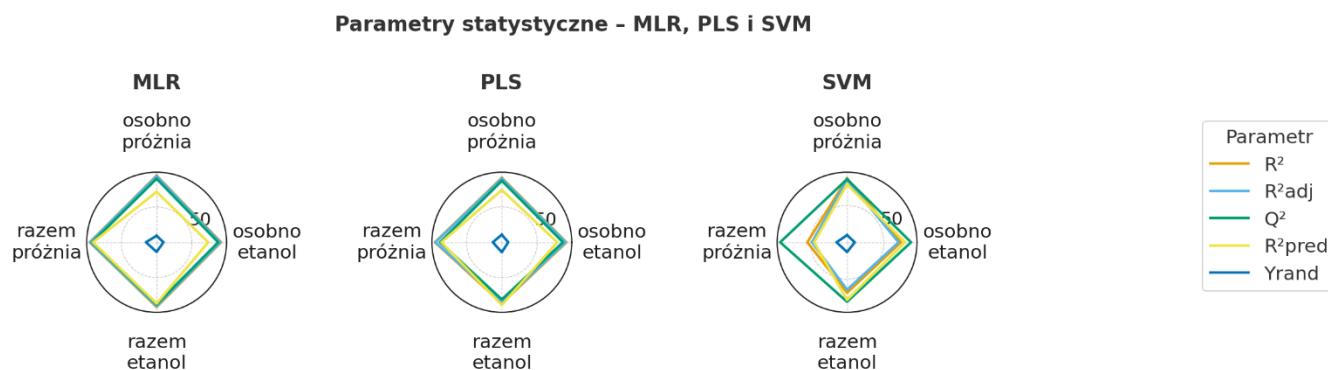
a)  $H_{CO_2}$



**Rysunek 4. 15.** Porównanie parametrów statystycznych modeli  $H_{CO_2}$ : a) MLR, b) PLS oraz c) SVM w zależności od sposobu i środowiska optymalizacji cieczy jonowych. Parametry wyrażono w wartościach procentowych.

Na rysunkach 4.15a, 4.15b oraz 4.15c przedstawiono wartości parametrów statystycznych ( $R^2$ ,  $R^2_{adj}$ ,  $Q^2$ ,  $R^2_{pred}$ ,  $Y_{rand}$ ) dla trzech rodzajów modeli regresyjnych  $H_{CO_2}$  w zależności od sposobu optymalizacji cieczy jonowych. Zaobserwować można, że modele MLR oraz PLS niezależnie od sposobu

optymalizacji charakteryzują się stabilnymi i zbliżonymi do siebie wartościami parametrów statystycznych, przy czym modele MLR-osobno negatywnie odbiegają od pozostałych jeśli chodzi o wynik walidacji zewnętrznej. Modele SVM radzą sobie zauważalnie gorzej, przez co nie są zalecane do modelowania wartości  $H_{CO_2}$ .

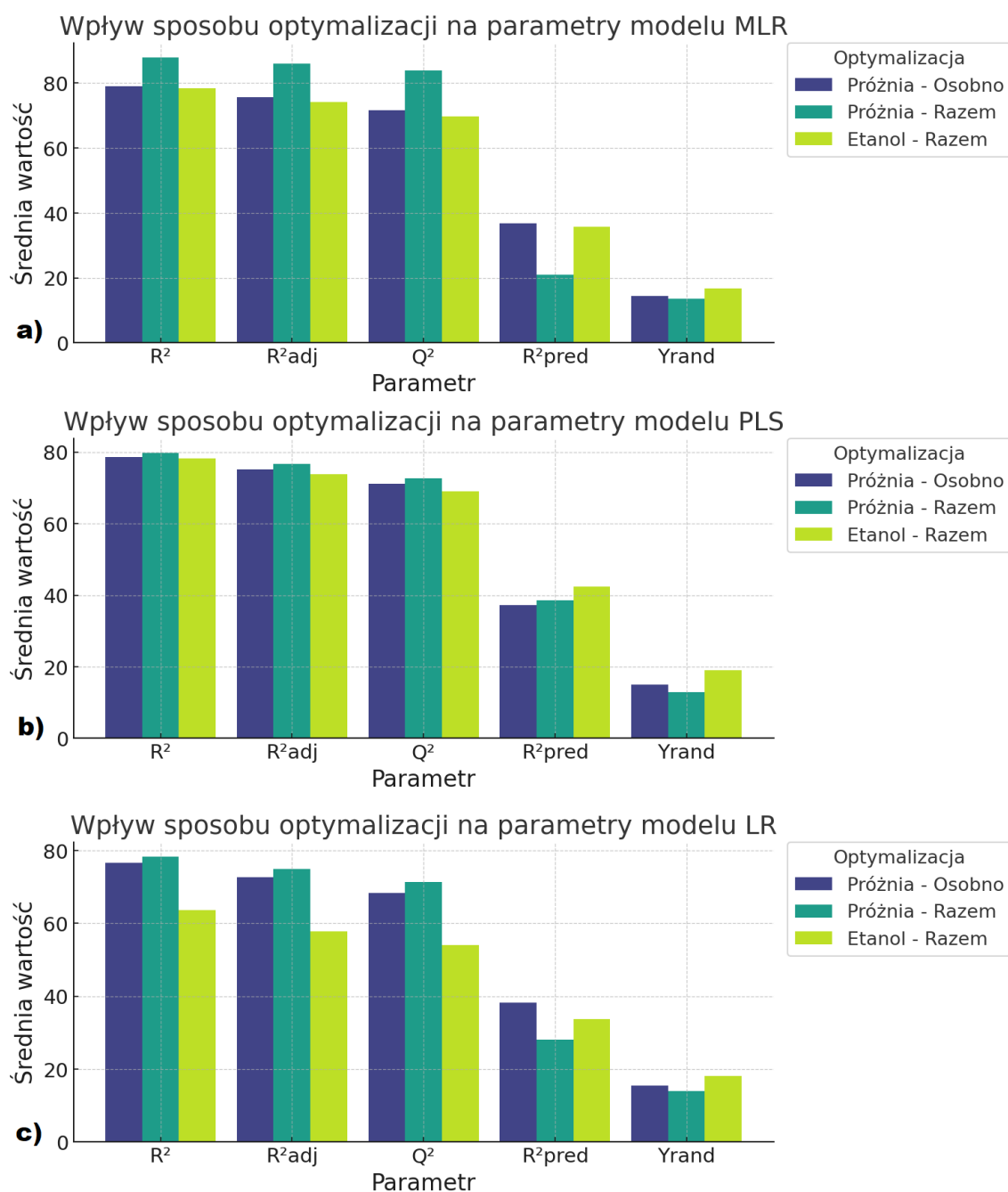


**Rysunek 4. 16.** Wykresy radarowe porównujące parametry statystyczne omawianych modeli  $H_{CO_2}$

Rysunek 4.16 przedstawia wykresy radarowe porównujące omawiane parametry statystyczne w zależności od wybranego modelu i warunków optymalizacji. Po raz kolejny zobaczyć można, że pomiędzy modelami MLR a PLS nie ma znaczących różnic, natomiast modele SVM nie radzą sobie w przypadku pracy z cieczami zoptymalizowanymi osobno w próżni oraz generalnie wypadają gorzej w przypadku optymalizacji w etanolu.

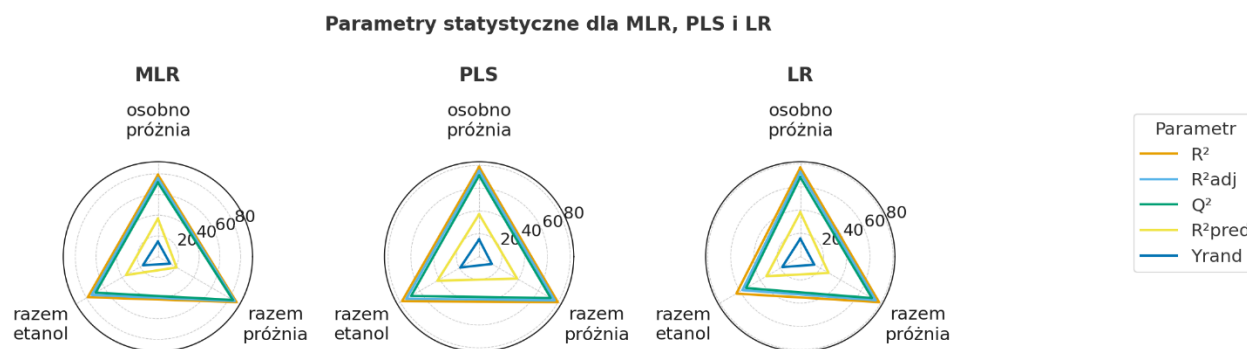
#### b) FFV

Na rysunkach 4.17a, 4.17b oraz 4.17c przedstawiono wartości parametrów statystycznych ( $R^2$ ,  $R^2_{adj}$ ,  $Q^2$ ,  $R^2_{pred}$ ,  $Y_{rand}$ ) dla trzech rodzajów modeli regresyjnych FFV w zależności od sposobu optymalizacji cieczy jonowych. Zauważyć można, że modele LR są stabilne najbardziej wrażliwe na metodę optymalizacji, co przekłada się na większe rozbieżności w wartościach poszczególnych statystyk. Modele MLR oraz PLS podobnie jak w przypadku modelowania  $H_{CO_2}$  charakteryzują się bardzo zbliżonymi parametrami statystycznymi. Najlepsze dopasowanie do zestawu uczącego w każdym z modeli uzyskuje się dla optymalizacji w próżni razem, jednak ten niewielki wzrost  $R^2$  nie rekompensuje znacznego pogorszenia wyników walidacji zewnętrznej.



**Rysunek 4. 17.** Porównanie parametrów statystycznych modeli FFV: a) MLR, b) PLS oraz c) LR w zależności od sposobu i środowiska optymalizacji cieczy jonowych. Parametry wyrażono w wartościach procentowych.

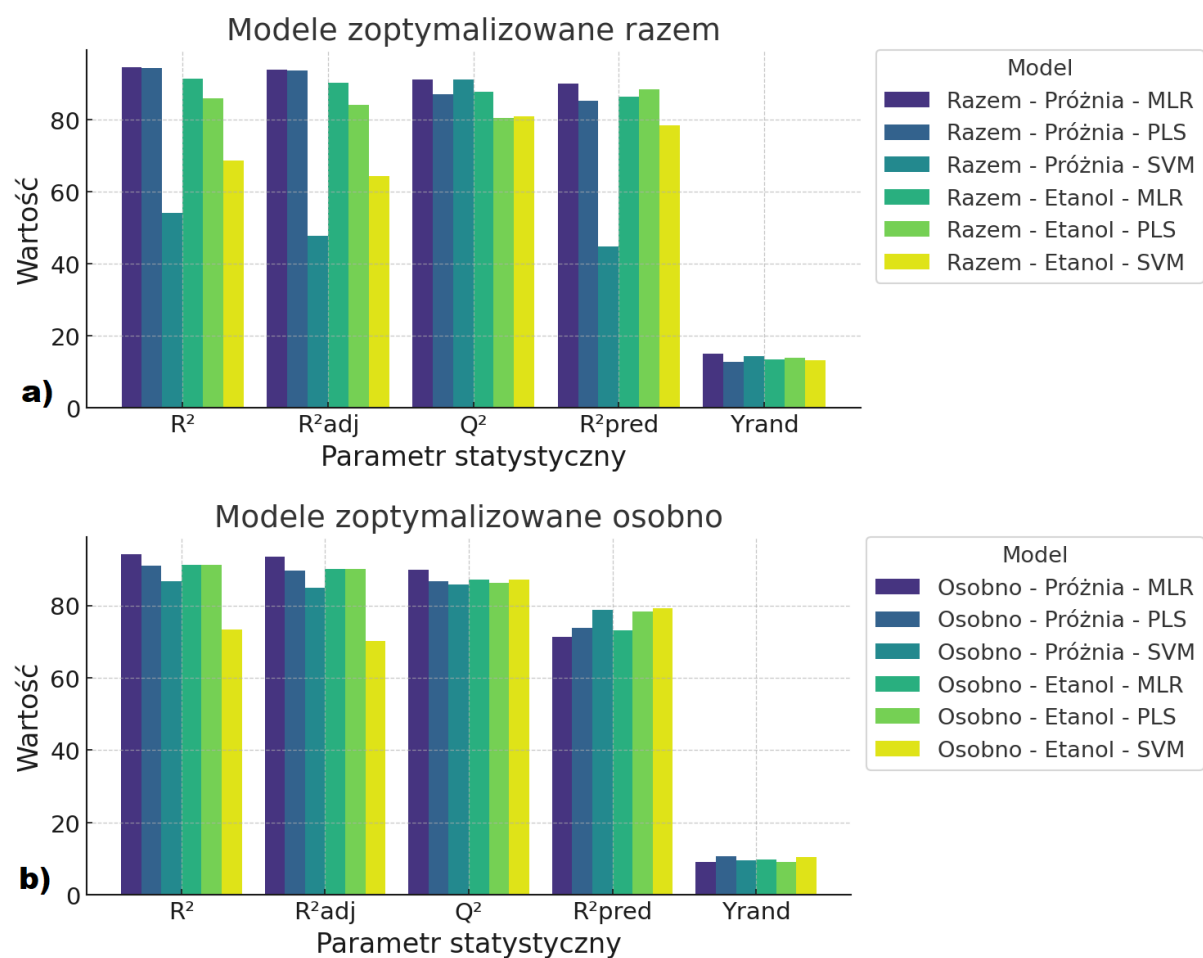
Rysunek 4.18 przedstawia wykres radarowy porównujący trzy rodzaje modeli regresyjnych. Wyraźnie widać, że żaden z modeli nie radzi sobie dobrze z walidacją zewnętrzną, czyli dopasowaniem do zbioru testowego. Parametr  $R^2_{pred}$  jest znacząco niższy od pozostałych, co pokazuje, że FFV jest zbyt specyficznym parametrem, aby dało się go z odpowiednią dokładnością wykorzystać do predykcji z zastosowaniem QSPR – wpływ na niego ma zbyt wiele czynników, których nie można zawrzeć w ramach deskryptorów molekularnych, nawet uwzględniając interakcje cząsteczek cieczy pomiędzy sobą.



**Rysunek 4. 18.** Wykresy radarowe porównujące parametry statystyczne omawianych modeli FFV

#### 4.3.3. Porównanie optymalizacji jonów osobno i razem

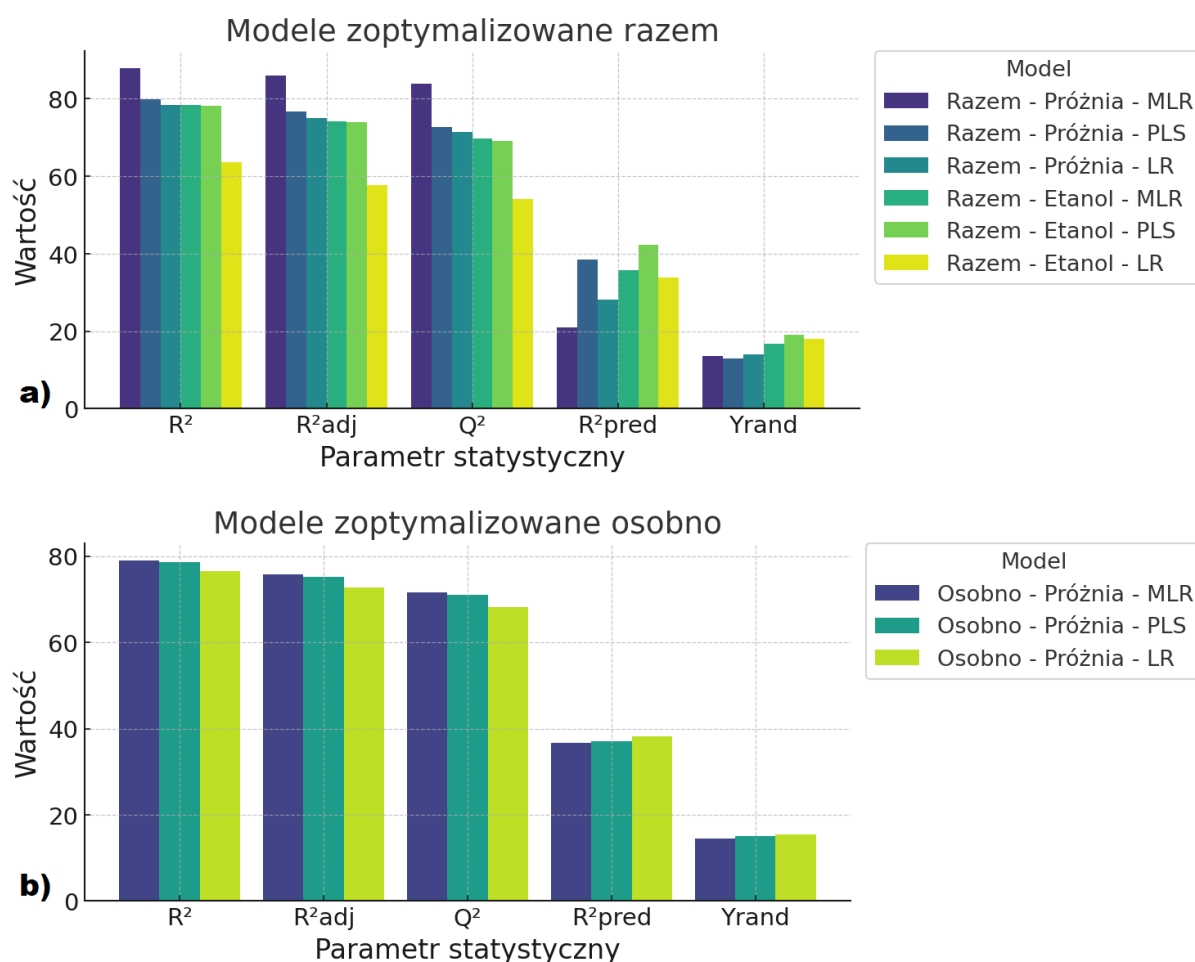
a)  $H_{CO_2}$



**Rysunek 4. 19.** Porównanie parametrów statystycznych modeli  $H_{CO_2}$ : a) razem, b) osobno w zależności od środowiska optymalizacji cieczy jonowych oraz metody modelowania. Parametry wyrażono w wartościach procentowych

Na rysunkach 4.19a i 4.19b przedstawiono wartości parametrów statystycznych ( $R^2$ ,  $R^2_{adj}$ ,  $Q^2$ ,  $R^2_{pred}$ ,  $Y_{rand}$ ) dla modeli  $H_{CO2}$  w zależności od optymalizacji cieczy osobno lub razem. Zauważyć można, że modele zoptymalizowane osobno są bardziej stabilne – wartości parametrów statystycznych nie różnią się aż tak bardzo w zależności od środowiska optymalizacji lub metody modelowania. Modele zoptymalizowane razem cechują się większymi wahaniami, jednak z drugiej strony lepiej radzą sobie z walidacją zewnętrzną, gdyż tylko w tych przypadkach wartości  $R^2_{pred}$  przekraczają 0,8. Wartość parametru  $Y_{rand}$  jest średnio wyższa (a więc gorsza) dla modeli zoptymalizowanych razem, zauważyć należy jednak że jest to parametr, który jest średnią z wielu wartości generowanych losowo, a żaden z nich nie przekracza wartości 0,2.

b) FFV



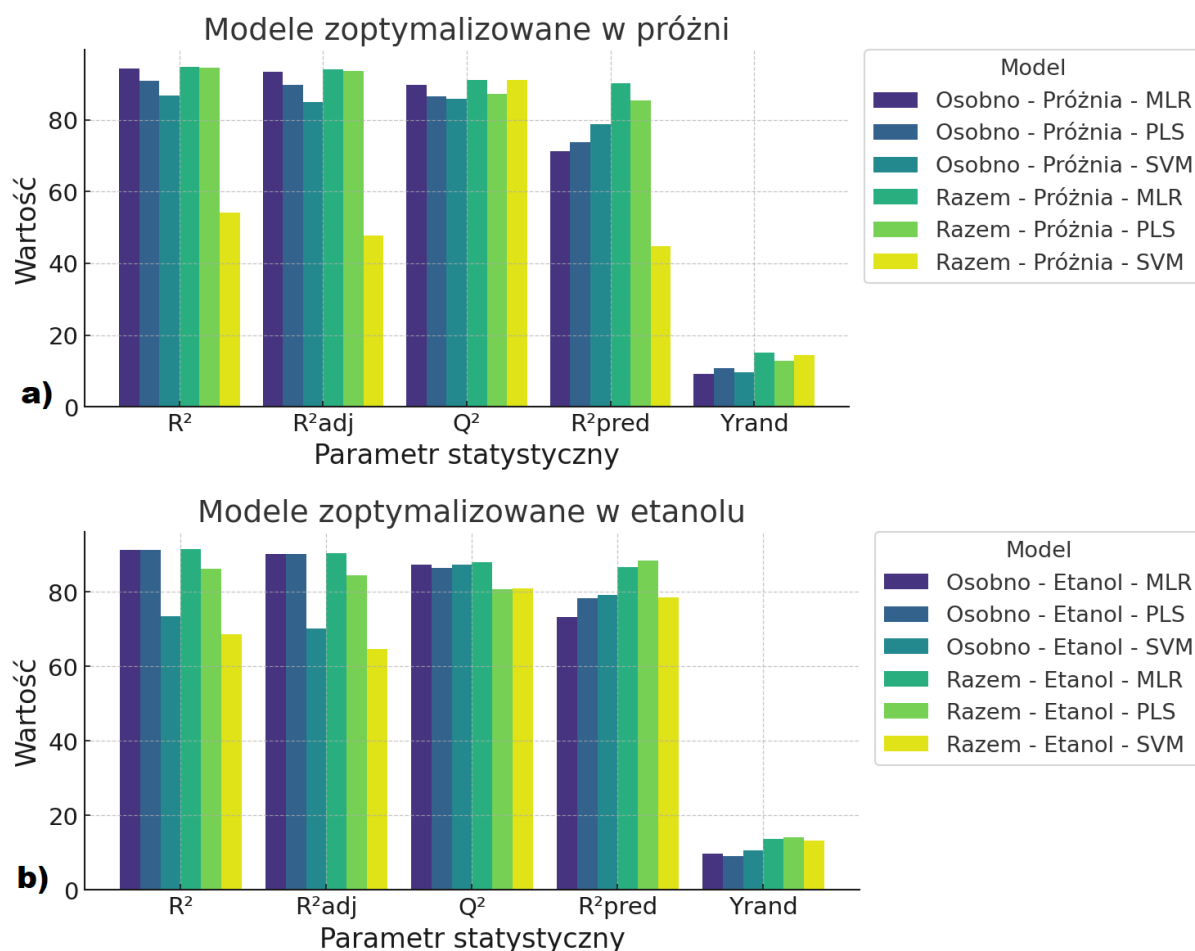
**Rysunek 4. 20.** Porównanie parametrów statystycznych modeli FFV: a) razem, b) osobno w zależności od środowiska optymalizacji cieczy jonowych oraz metody modelowania. Parametry wyrażono w wartościach procentowych

Na rysunkach 4.20a i 4.20b przedstawiono wartości parametrów statystycznych ( $R^2$ ,  $R^2_{adj}$ ,  $Q^2$ ,  $R^2_{pred}$ ,  $Y_{rand}$ ) dla modeli FFV w zależności od optymalizacji cieczy osobno lub razem. Ze względu na wykluczenie modeli z optymalizacją osobno w etanolu z analizy, rysunek X przedstawia jedynie trzy

modele zamiast sześciu, jak w przypadku większości wykresów. W przypadku parametru FFV nie zaobserwowano znaczących różnic pomiędzy optymalizacją osobno i razem.

#### 4.3.4. Porównanie optymalizacji jonów w próżni i w etanolu

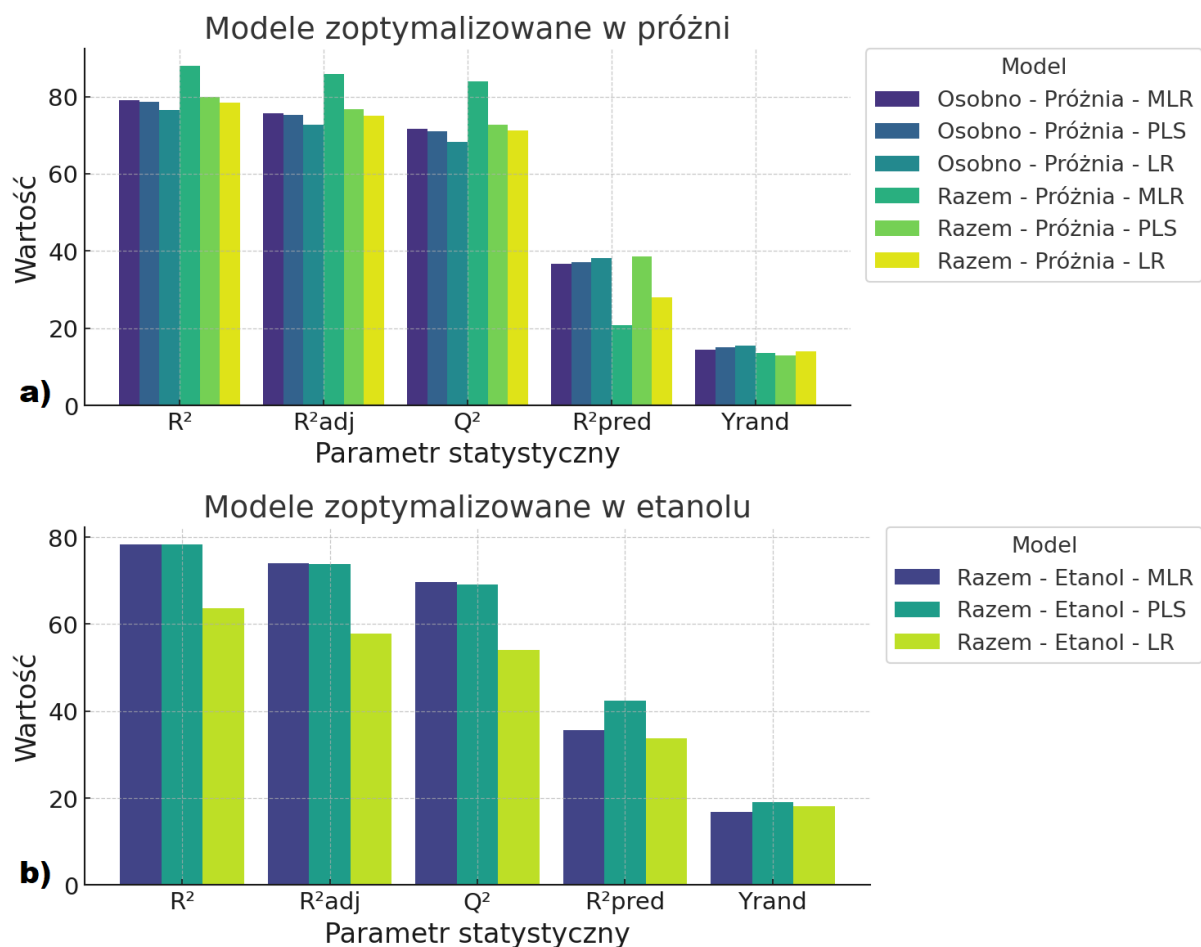
##### a) $H_{CO_2}$



**Rysunek 4. 21.** Porównanie parametrów statystycznych modeli  $H_{CO_2}$ : a) w próżni, b) w etanolu w zależności od sposobu optymalizacji cieczy jonowych oraz metody modelowania. Parametry wyrażono w wartościach procentowych

Na rysunkach 4.21a i 4.21b przedstawiono wartości parametrów statystycznych ( $R^2$ ,  $R^2_{adj}$ ,  $Q^2$ ,  $R^2_{pred}$ ,  $Y_{rand}$ ) dla modeli  $H_{CO_2}$  w zależności od optymalizacji cieczy w próżni lub w etanolu. Pomiedzy parametrami  $R^2$  oraz  $R^2_{adj}$  nie ma zauważalnej różnicy dla modeli MLR oraz PLS, natomiast model SVM poradził sobie lepiej w próżni w przypadku optymalizacji osobno oraz lepiej w etanolu przy optymalizacji razem. Porównując parametry walidacyjne zauważyć można, że optymalizacja w próżni przekłada się na nieznacznie lepszy wynik walidacji wewnętrznej, a optymalizacja w etanolu ma większy wpływ na wynik walidacji zewnętrznej. Jeśli model służyć ma tylko do analizy zależności budowy od właściwości, a nie do przewidywania nowych wartości, lepiej zdecydować się na optymalizację w próżni.

b) FFV

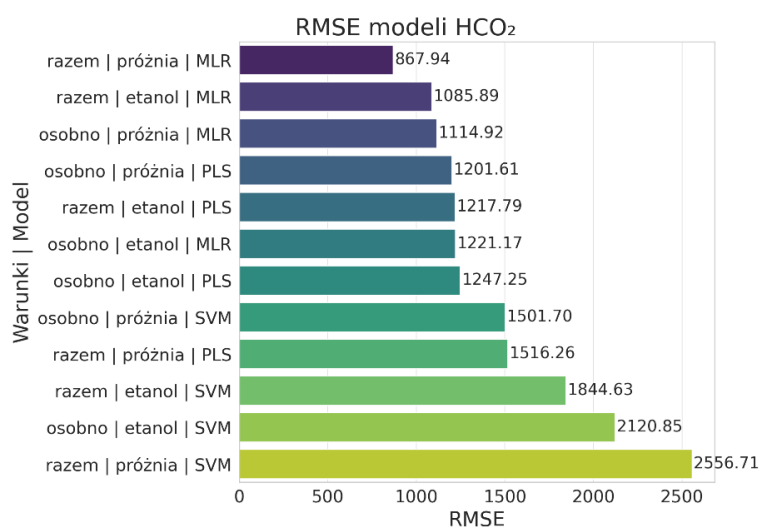


**Rysunek 4. 22.** Porównanie parametrów statystycznych modeli FFV: a) razem, b) osobno w zależności od sposobu optymalizacji cieczy jonowych oraz metody modelowania. Parametry wyrażono w wartościach procentowych

Na rysunkach 4.22a i 4.22b przedstawiono wartości parametrów statystycznych ( $R^2$ ,  $R^2_{adj}$ ,  $Q^2$ ,  $R^2_{pred}$ ,  $Y_{rand}$ ) dla modeli FFV w zależności od optymalizacji cieczy w próżni lub w etanolu. Podobnie jak w przypadku porównania optymalizacji osobno lub razem, dysponujemy ograniczonym zestawem danych ze względu na wykluczenie modelu FFV osobno w etanolu. Modele charakteryzują się bardzo zbliżonymi parametrami statystycznymi. Podobnie jak w przypadku modeli  $H_{CO_2}$ , walidacja zewnętrzna jest nieznacznie lepsza w przypadku optymalizacji w etanolu.

#### 4.3.5. RMSE

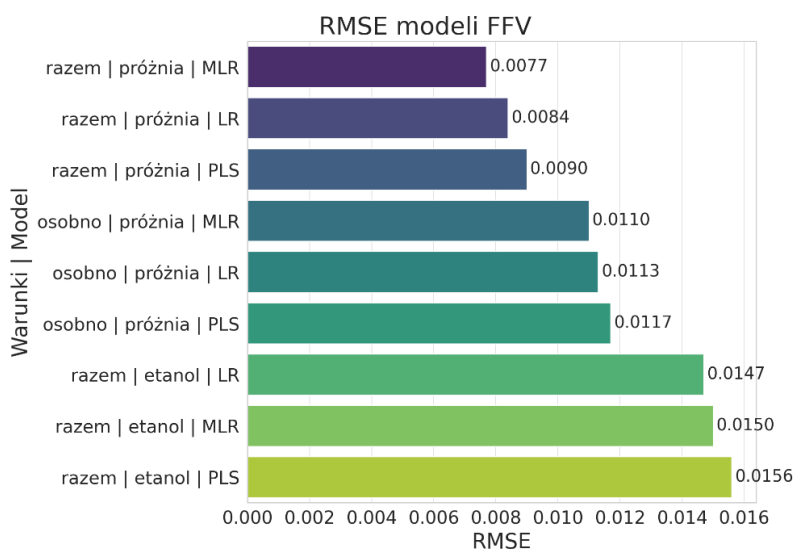
##### a) $H_{CO_2}$



**Rysunek 4. 23.** RMSE obliczone dla modeli  $H_{CO_2}$  uszeregowane rosnąco

Na rysunku 4.23 przedstawiono wartości RMSE obliczone dla modeli  $H_{CO_2}$  uszeregowane rosnąco. Zauważyć należy, że w przypadku analizowanego zbioru danych największy wpływ na wartość RMSE ma wybrana metoda modelowania. MLR charakteryzuje się najmniejszymi wartościami, PLS niewiele większymi, a najgorsze wyniki uzyskuje się dla modeli SVM. Optymalizacja razem oraz w próżni jest preferowana w przypadku metody MLR, natomiast nie ma większego znaczenia w modelowaniu metodami PLS. W przypadku metody SVM wyraźnie widać, że jest ona bardzo wrażliwa na sposób optymalizacji cząsteczek co przekłada się na skokowo zmieniające się wartości RMSE.

##### b) FFV



**Rysunek 4. 24.** RMSE obliczone dla modeli FFV uszeregowane rosnąco

Na rysunku 4.24 przedstawiono wartości RMSE obliczone dla modeli FFV uszeregowane rosnąco. Metoda modelowania ma bardzo niewielki wpływ na wartość RMSE, co jest znaczącą różnicą w porównaniu do modeli  $H_{CO_2}$ . Największy pozytywny wpływ ma optymalizacja w próżni, a następnie optymalizacja cząsteczek cieczy razem.

#### 4.3.6. Wpływ sposobu optymalizacji na wartości deskryptorów

W celu określenia wpływu sposobów optymalizacji na wartości deskryptorów obliczono wartości deskryptorów molekularnych dla bis(trifluorometanosulfonylo)imidku  $[NTf_2]$  w trzech konfiguracjach:

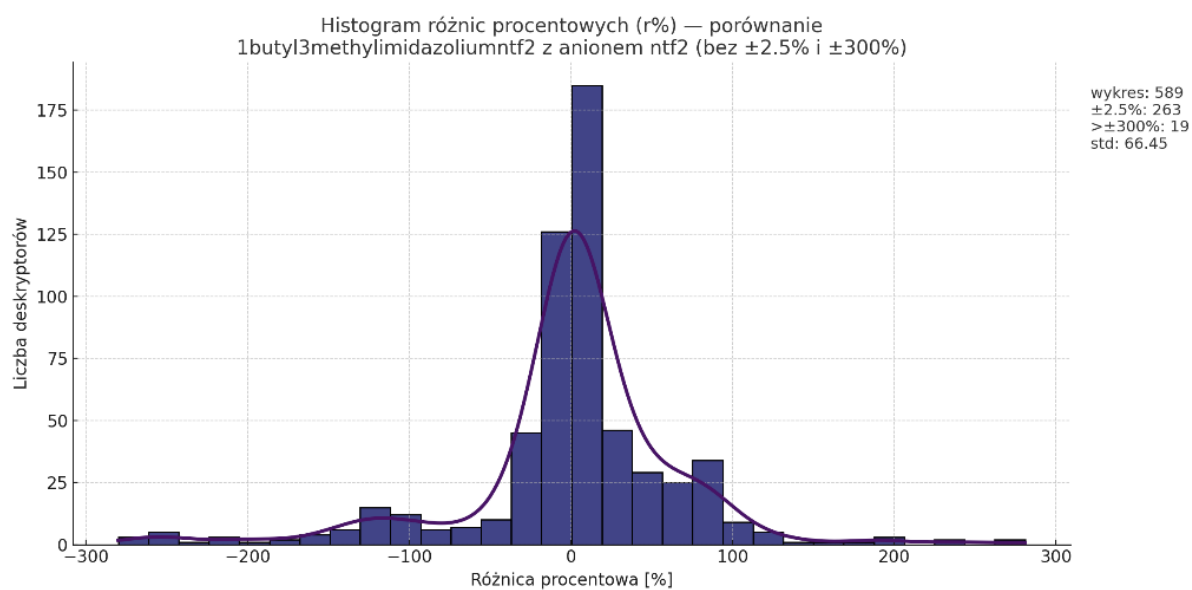
- I. 5270 deskryptorów dla jonu  $[NTf_2]$  zoptymalizowanego osobno w próżni
- II. 5270 deskryptorów dla jonu  $[NTf_2]$  dla 26 cieczy jonowych zoptymalizowanych razem w próżni
- III. 5270 deskryptorów dla jonu  $[NTf_2]$  dla 26 cieczy jonowych zoptymalizowanych razem w etanolu

Dla zestawów II oraz III obliczono następnie współczynnik zmienności, dzieląc odchylenie standardowe 26 wartości każdego deskryptora przez ich średnią arytmetyczną.

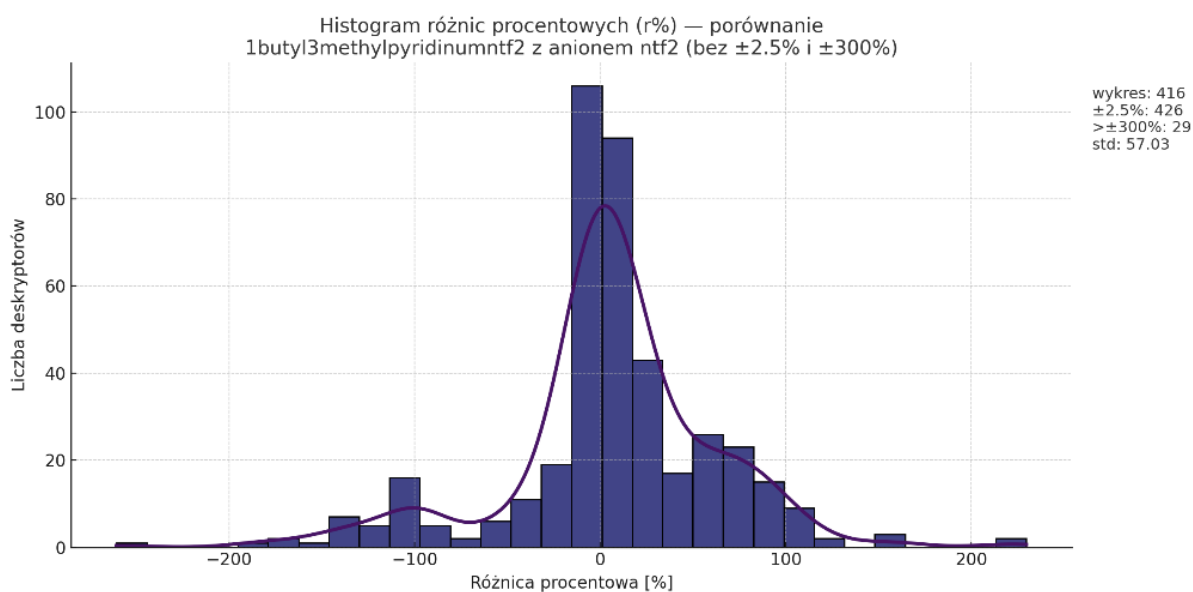
Na samym początku warto zaznaczyć, że wszystkie deskryptory 2D niezależnie od sposobu optymalizacji w każdym przypadku posiadały tę samą wartość. Spośród 279310 (5270 deskryptorów w 53 zestawach danych – jon  $[NTf_2]$  osobno w próżni oraz po 26 cieczy razem w próżni i razem w etanolu) analizowanych deskryptorów ani jeden deskryptor dwuwymiarowy nie wykazał odchylenia standardowego innego niż 0. Obserwacja ta była spodziewana, gdyż różnice w wartościach powinny wynikać jedynie z przestrzennego ułożenia atomów i wiązań w cząsteczce.

##### 4.3.6.1. Optymalizacja osobno i razem

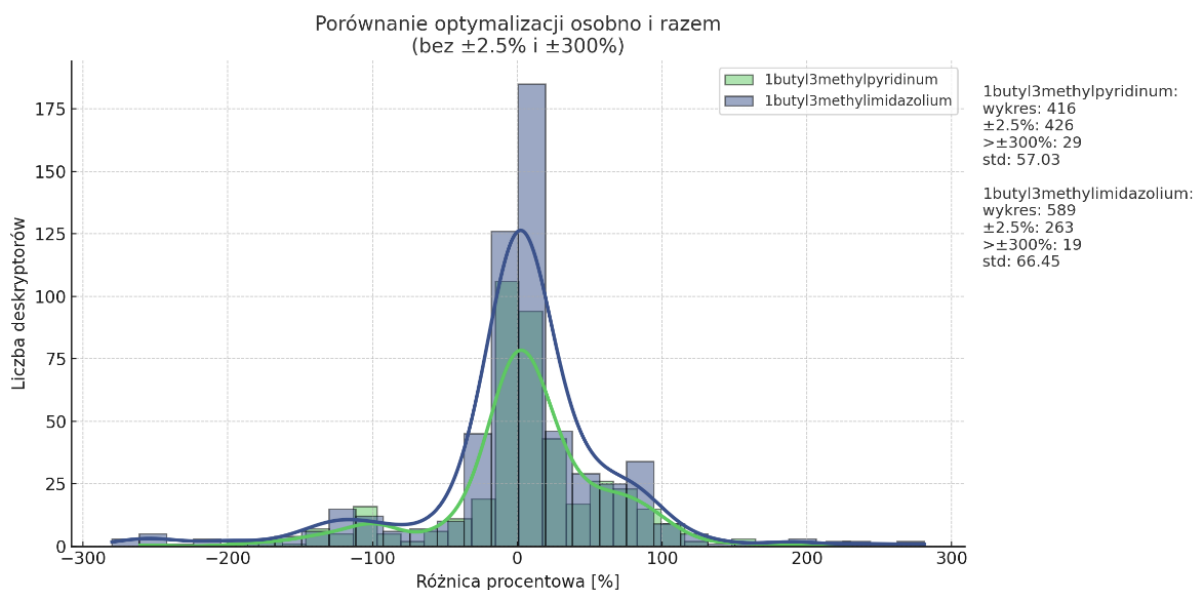
Pierwszym krokiem było porównanie wartości deskryptorów dla jonu zoptymalizowanego osobno oraz razem. W tym celu spośród zestawu cieczy zoptymalizowanych razem w próżni wybrano przykładowe dwie zawierające anion  $[NTf_2]$  oraz kationy: 1-butylo-3-metyloimidazoliowy oraz 1-butylo-3-metylopirydynowy. Na rysunkach 4.25 oraz 4.26 przedstawiono histogramy różnic procentowych wartości deskryptorów pomiędzy anionem  $[NTf_2]$  zoptymalizowanym osobno w próżni oraz anionem  $[NTf_2]$  zoptymalizowanym w próżni w obecności dwóch wspomnianych kationów. Na histogramach pokazanych na wszystkich rysunkach w rozprawie uwzględniono jedynie te spośród 5270 deskryptorów, które nie posiadały identycznych wartości, a także dla poprawy czytelności wykresu usunięto z nich deskryptory, których wartość procentowa różniła się o mniej niż 2,5% oraz więcej niż 300%. Informacja o liczbie pominiętych w ten sposób deskryptorów znajduje się po prawej stronie wykresu, wraz z wartością odchylenia standardowego dla prezentowanego zestawu danych.



**Rysunek 4. 25.** Histogram różnic procentowych wartości deskryptorów pomiędzy anionem [NTf<sub>2</sub>] zoptymalizowanym osobno w próżni oraz anionem [NTf<sub>2</sub>] zoptymalizowanym w próżni w obecności jonu [BMIM]



**Rysunek 4. 26.** Histogram różnic procentowych wartości deskryptorów pomiędzy anionem [NTf<sub>2</sub>] zoptymalizowanym osobno w próżni oraz anionem [NTf<sub>2</sub>] zoptymalizowanym w próżni w obecności jonu [B3MPy]

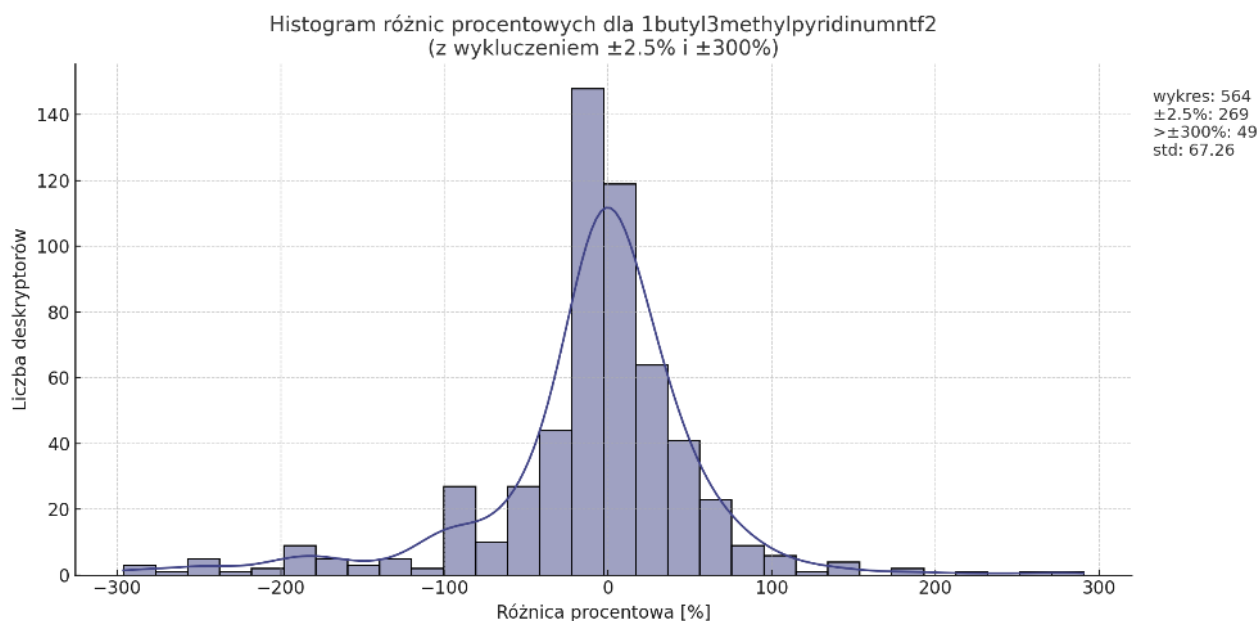


**Rysunek 4. 27.** Wspólny histogram różnic procentowych wartości deskryptorów pomiędzy anionem [NTf<sub>2</sub>] zoptymalizowanym osobno w próżni oraz anionem [NTf<sub>2</sub>] zoptymalizowanym w próżni w obecności kationów [BMIM] oraz [B3MPy]

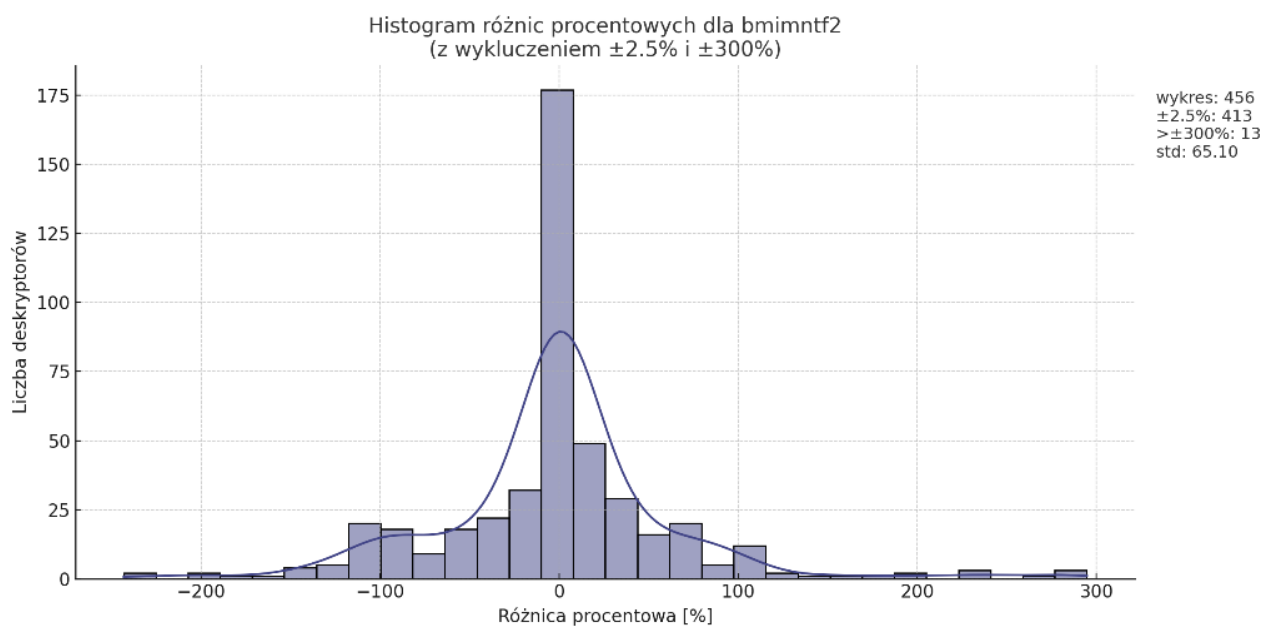
Na rysunku 4.27 przedstawiono zestawienie obu cieczy na jednym histogramie. Optymalizacja z kationem [B3MPy] cechowała się większą liczbą zarówno niewielkich odchyień, jak i wartości skrajnych, z kolei optymalizacja z kationem [BMIM] prowadziła do zdecydowanie większej liczby deskryptorów różniących się w zakresie  $\pm 2,5$ -100%. W obu przypadkach 871 deskryptorów wykazywało inne wartości w optymalizacji razem oraz osobno. Stanowi to jedynie 16% wszystkich obliczonych deskryptorów, ale aż 59% spośród wszystkich deskryptorów 3D, których obliczyć można 1470. Zmienność danych wejściowych na poziomie niemal 60% wskazuje na możliwe znaczące różnice w modelach QSPR, które można zbudować na ich podstawie.

#### 4.3.6.2. Optymalizacja w próżni i w etanolu

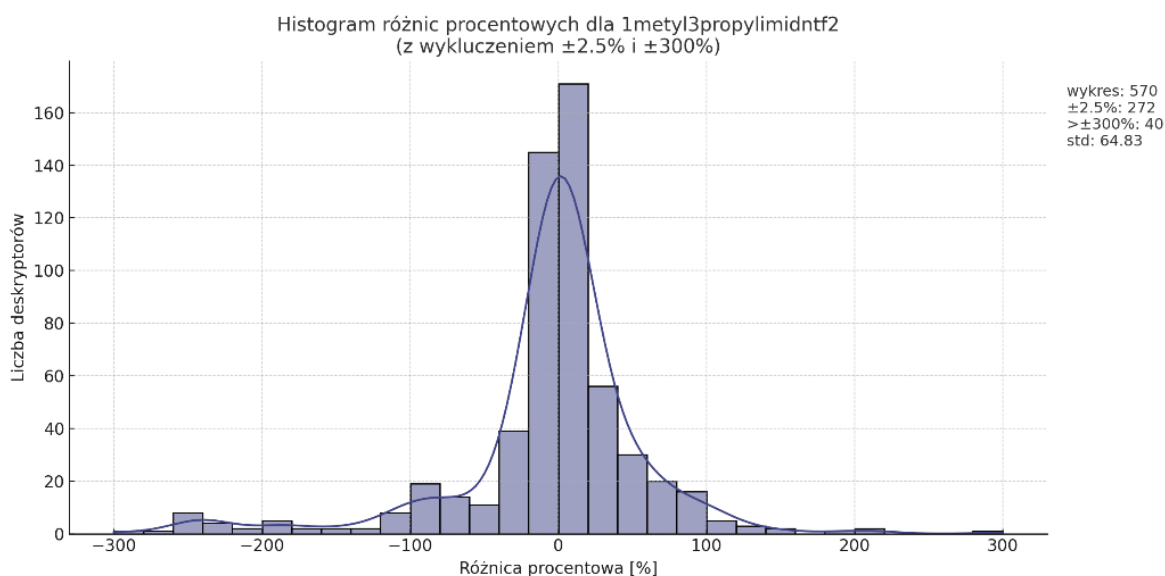
W kolejnym kroku analizie poddano różnice w wartościach deskryptorów obliczonych dla cieczy zoptymalizowanych razem w próżni oraz razem w środowisku etanolu. W tym celu przygotowano w identyczny sposób histogramy dla 26 cieczy jonowych. Na rysunkach 4.28-4.30 przedstawiono przykładowe wykresy dla cieczy [BMIM][NTf<sub>2</sub>], [B3MPy][NTf<sub>2</sub>] oraz [PMIM][NTf<sub>2</sub>]. Pozostałe histogramy znajdują się w Materiałach uzupełniających na rysunkach A1-A3.



**Rysunek 4. 28.** Histogram różnic procentowych wartości deskryptorów pomiędzy jonem  $[\text{NTf}_2]$  cieczy  $[\text{B3MPy}][\text{NTf}_2]$  zoptymalizowanej razem w próżni oraz razem w środowisku etanolu



**Rysunek 4. 29.** Histogram różnic procentowych wartości deskryptorów pomiędzy jonem  $[\text{NTf}_2]$  cieczy  $[\text{BMIM}][\text{NTf}_2]$  zoptymalizowanej razem w próżni oraz razem w środowisku etanolu

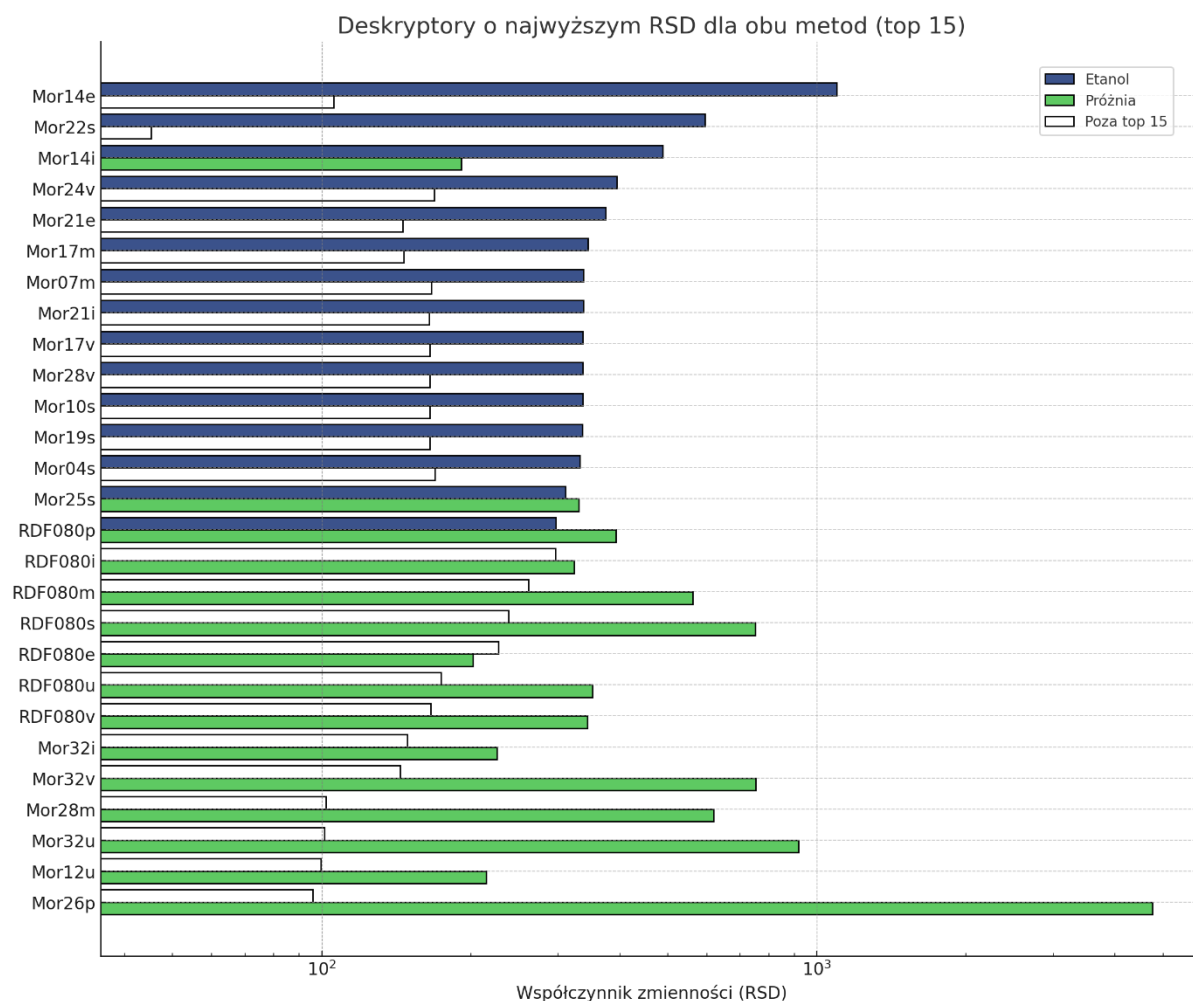


**Rysunek 4. 30.** Histogram różnic procentowych wartości deskryptorów pomiędzy jonem [NTf<sub>2</sub>] cieczy [PMIM][NTf<sub>2</sub>] zoptymalizowanej razem w próżni oraz razem w środowisku etanolu

Rozkład różnic procentowych przypomina w przeważającej większości przypadków rozkład normalny. Anomalnie wysokie słupki w okolicy -100%, które zaobserwować można na wszystkich trzech diagramach, związane są z przyjęciem wartości -100% dla wszystkich deskryptorów, które po zmianie sposobu optymalizacji z próżni na etanol przyjęły wartość 0, mimo posiadania wcześniej innej wartości. Liczba skrajnych różnic waha się pomiędzy 9 dla [BzMPyrr][NTf<sub>2</sub>] a 49 dla [B3MPy][NTf<sub>2</sub>]. Ciecze te różnią się jedynie podstawnikiem przy pierścieniu w kationie, anion dla którego obliczono deskryptory pozostaje bez zmian. Choć można byłoby założyć, że większy objętościowo kation z podstawnikiem benzylowym będzie cechował się większą liczbą skrajnie dużych różnic w wartościach deskryptorów, obserwacje pokazują, że jest odwrotnie. Najmniej różnic w zakresie  $\pm 2,5\%$  zaobserwowano dla cieczy [HMPyrr][NTf<sub>2</sub>], dla której obliczono ich jedynie 256. Posiada ona również najwięcej deskryptorów z zakresu  $\pm 2,5\text{-}100\%$  (613). Co ciekawe, w każdym przypadku zaobserwowano 882 deskryptory, na których wartość wpływ miała zmiana środowiska optymalizacji, co stanowi dokładnie 60% wszystkich deskryptorów 3D.

#### 4.3.6.3. Deskryptory o największej zmienności

W następnym etapie analizy skorzystano z obliczonego wcześniej współczynnika zmienności, aby sprawdzić, jakie deskryptory są najbardziej podatne na zmiany.

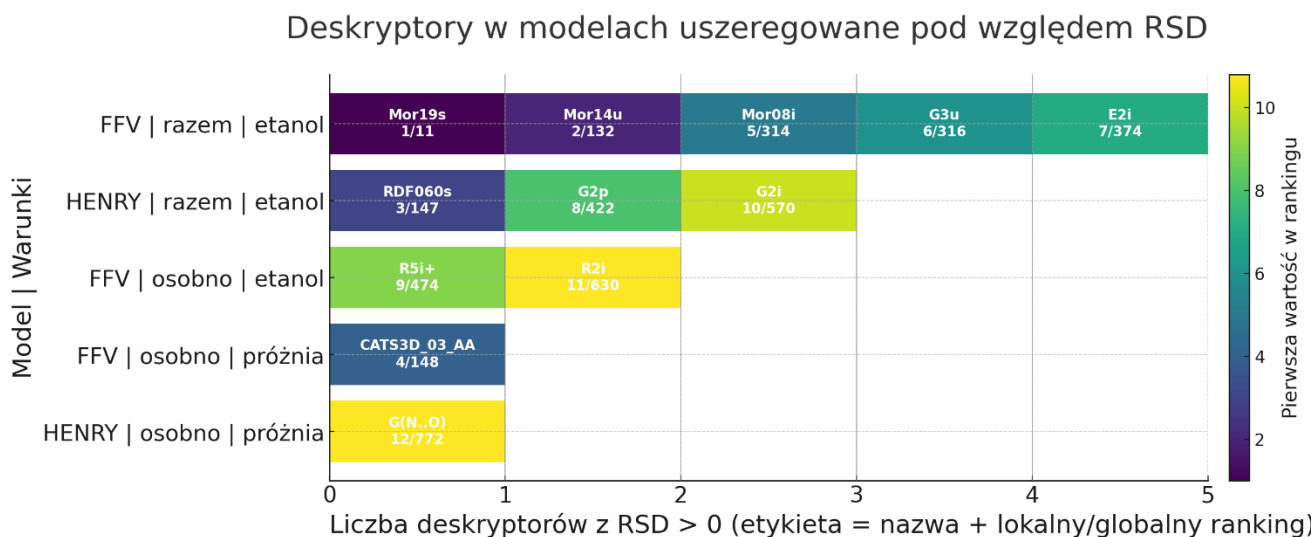


**Rysunek 4. 31.** 15 deskryptorów o największej zmienności dla każdej z metod. Kolorem niebieskim oznaczono optymalizację w etanolu, kolorem zielonym optymalizację w próżni. Na biało zaznaczone są wartości współczynnika zmienności danego deskryptora dla przeciwnej metody optymalizacji, jeśli znajdują się one poza zbiorem 15 najwyższych wartości.

Na rysunku 4.31 przedstawiono 15 deskryptorów o największej zmienności dla każdej z metod – optymalizacji razem w próżni oraz razem w etanolu. Jedynie trzy deskryptory (Mor14i, Mor25s oraz RDF080p) są wspólne dla obu metod, a pozostałe dwadzieścia cztery są unikalne, po dwanaście dla każdej metody. Największą zmienność ze wszystkich wykazuje deskryptor Mor26p. Analizując kategorie, najbardziej na zmiany wrażliwe są te z kategorii 3D-MorSE oraz RDF, do których należy odpowiednio 224 i 210 deskryptorów. Trend ten otrzymuje się po rozszerzeniu listy najbardziej zmiennych deskryptorów do 30. Pojawiają się na niej jedynie dwa z innej kategorii – VE2sign\_G/D oraz VE2sign\_G, które należą do 3D matrix-based descriptors. Warto zauważyć, że w tym przypadku już 19 deskryptorów na liście 30 najbardziej zmiennych występuje zarówno w przypadku optymalizacji w próżni, jak i w etanolu. Rysunek A4 znajduje się w Materiałach uzupełniających.

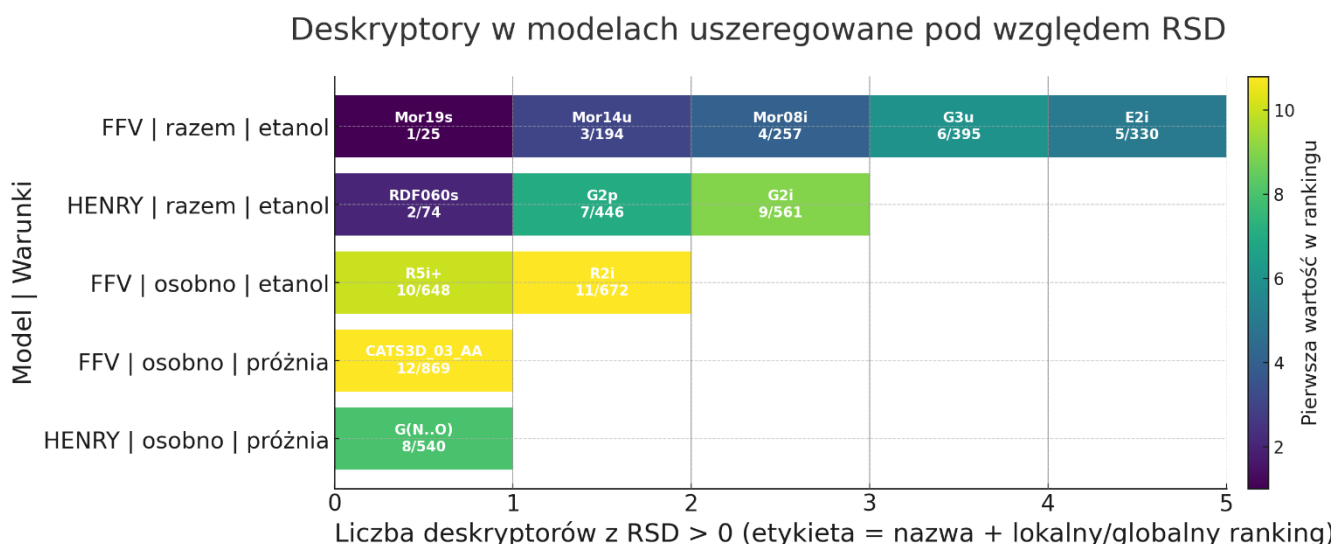
#### 4.3.7. Zmienność deskryptorów a otrzymane modele QSPR

Ostatnim krokiem było sprawdzenie, jak zmienność deskryptorów miała wpływ na modele QSPR opisane w rozdziale 4.3.1. Na rysunkach 4.32-4.33 przedstawiono deskryptory 3D, które wykazały zmienność i znalazły się w modelach. Uszeregowano je pod kątem wartości RSD obliczonych wcześniej dla dwóch metod – optymalizacji razem w próżni oraz razem w etanolu. Na każdym z rysunków przedstawiono je w dwóch rankingach. Ranking lokalny odnosi się jedynie do zbioru dwunastu deskryptorów przedstawionych na wykresach, a ranking globalny odnosi się do wszystkich obliczonych deskryptorów. Należy wspomnieć również, że analizę zmienności przeprowadzono dla deskryptorów obliczonych dla anionu, natomiast część deskryptorów 3D w analizowanych modelach powiązana jest z kationami. Należy przyjąć więc założenie, że podatność na zmianę metody optymalizacji jest podobna, niezależnie od rodzaju jonu, a poniższa analiza ma charakter jedynie orientacyjny. Do deskryptorów anionowych w badaniu należą: Mor08i, G3u, E2i, RDF060s oraz R5i+.



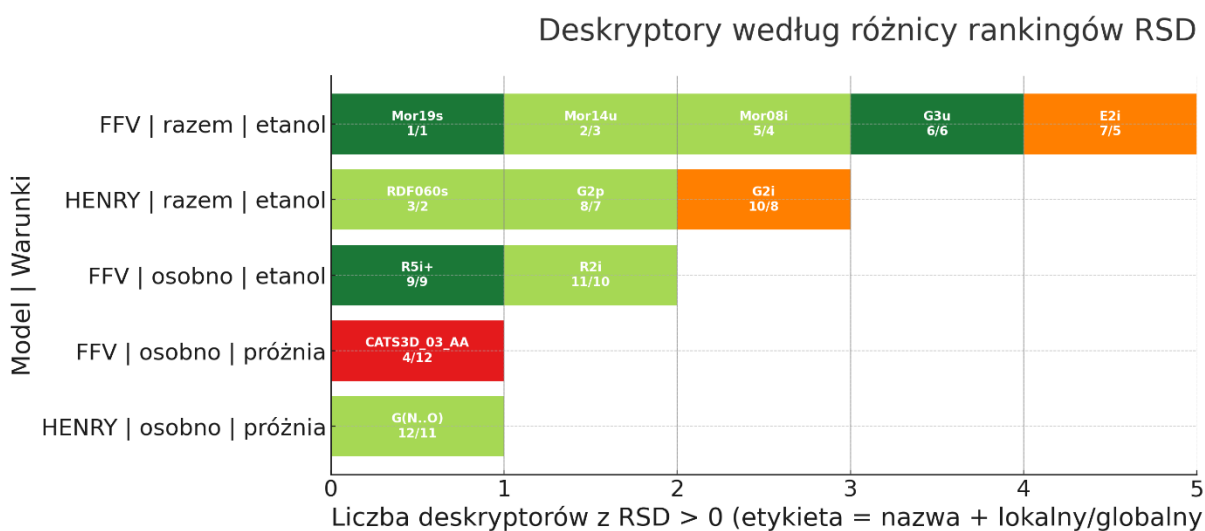
**Rysunek 4. 32.** Ranking zmienności deskryptorów dla RSD obliczonego przy optymalizacji razem w próżni

Rysunek 4.32 przedstawia ranking zmienności deskryptorów dla RSD obliczonego przy optymalizacji razem w próżni. Największą zmienność wykazuje Mor19s z modelu FFV przy optymalizacji razem w etanolu, a najmniejszą G(N...O) z modelu H<sub>CO2</sub> przy optymalizacji osobno w próżni. Warto wspomnieć, że żaden z omawianych deskryptorów 3D nie pojawił się w więcej, niż jednym modelu. Największą zmienność wykazują deskryptory z grupy 3D-MorSE, które należą do jednych z najtrudniejszych w interpretacji.



**Rysunek 4. 33.** Ranking zmienności deskryptorów dla RSD obliczonego przy optymalizacji razem w etanolu

Rysunek 4.33 przedstawia ranking zmienności deskryptorów przy optymalizacji razem w etanolu. Największą zmiennością charakteryzuje się ponownie Mor19s, natomiast najmniejszą CATS\_3D\_03\_AA, który w przypadku optymalizacji w próżni był na czwartej pozycji.



**Rysunek 4. 34.** Porównanie różnic w położeniu w rankingach lokalnych dla RSD obliczonych dla optymalizacji razem w próżni i razem w etanolu.

Rysunek 4.34 pokazuje porównanie rankingów lokalnych dla RSD obliczonych dla obu metod. Kolorem ciemnozielonym zaznaczono identyczną pozycję w obu rankingach, kolorem jasnozielonym różniącą się o 1, pomarańczowym różniącą się o 2, a na czerwono większą różnicę. Dziewięć z dwunastu deskryptorów zajmuje identyczne lub bardzo zbliżone pozycje w obu rankingach, co pokazuje, że wpływ środowiska optymalizacji na wartości deskryptorów 3D jest istotny, lecz w większości przypadków ogólny trend zostaje zachowany. Deskryptory z grup 3D-MoRSE, RDF oraz WHIM optymalizowane w próżni i w etanolu charakteryzują się nie identycznym, ale zbliżonym rozkładem zmienności. Największą

różnicę wykazuje wspomniany już CATS\_3D\_03\_AA, co może wskazywać na to, że grupa deskryptorów CATS3D, powiązanych z identyfikacją potencjalnych farmakoforów) jest znacznie bardziej wrażliwa na zmianę środowiska optymalizacji.

Powyższe zestawienia pokazują, porównując bezwzględne wartości deskryptorów, że są one bardzo podatne na zmiany sposobu i środowiska optymalizacji, co w konsekwencji może wywierać znaczny wpływ na modele QSPR. Efekt zmiany środowiska optymalizacji na etanol ma szczególne znaczenie w przypadku deskryptorów 3D. Ponadto, interesującą zależność można zauważyć analizując względne zmiany deskryptorów. Porównując zakres ich zmienności pomiędzy dwiema metodami optymalizacji okazuje się, że o ile ich wartości podlegają znaczącym wahaniom, to ogólny trend zmian pozostaje w większości przypadków niezaburzony. Z rysunku 3.34 wyraźnie wynika, że przeważająca część deskryptorów uszeregowana względnie, pod kątem współczynnika zmienności zajmuje identyczne lub bardzo podobne miejsca w przypadku obu zastosowanych metod optymalizacji. Po raz kolejny pokazuje to, że zmiana sposobu i środowiska optymalizacji struktury może nie mieć aż tak istotnego wpływu na model QSPR, jak wybór konkretnej metody modelowania.

## 5. WNIOSKI

### 5.1. Modelowanie stałej Henry'ego

W celu przeprowadzenia badań zebrano największy dostępny w literaturze zbiór danych dotyczących stałej Henry'ego dla ditlenku węgla w cieczach jonowych. Na podstawie zestawu zawierającego 73 ciecze jonowe o zróżnicowanej budowie stworzono trzy modele QSPR wykorzystując trzy metody modelowania – MLR, LR i PLS. Każdy z modeli pozytywnie przeszedł walidację statystyczną i porównując ich parametry statystyczne nie stwierdzono zauważalnych różnic w ich zdolnościach predykcyjnych. Analizie poddano deskryptory molekularne, na bazie których zbudowano modele.

W przeprowadzonym badaniu zidentyfikowano elementy strukturalne cząsteczek cieczy jonowych, które mają wpływ na wartość  $H_{CO_2}$ . W celu uzyskania najwyższej rozpuszczalności ditlenku węgla, kationy cieczy nie powinny posiadać w swojej strukturze grup nitrylowych, hydroksylowych ani atomów fluoru. Najdokładniej przebadane kationy oparte na pierścieniach imidazolowych charakteryzują się pośrednimi wartościami rozpuszczalności, przez co nie stanowią optymalnego wyboru zarówno w sytuacji, gdy pragnie się uzyskać wysoką, jak i niską rozpuszczalność  $CO_2$ . Podobnie jak w przypadku kationów, w celu otrzymania cieczy jonowej o najniższej wartości  $H_{CO_2}$  aniony nie powinny zawierać w swojej strukturze grup hydroksylowych. Preferowane są natomiast atomy fluoru oraz miejsca akceptorowe dla wiązań wodorowych. Najlepszymi właściwościami charakteryzowały się aniony [NTf<sub>2</sub>] oraz [C(CN)<sub>3</sub>]. Wśród przebadanych jonów nie ma jednak związku idealnego, a niektóre z nich, jak kationy amoniowe, poprzez dwa różne deskryptory wpływały na rozpuszczalność  $CO_2$  zarówno pozytywnie, jak i negatywnie.

Przeprowadzone badania potwierdzają doniesienia Muldoona i innych, w którym to artykule udowodniono, że kluczową rolę w zwiększaniu rozpuszczalności  $CO_2$  w cieczach jonowych ma obecność atomów fluoru w ich strukturze.[194] Tej ważnej cechy dotyczył drugi pod względem istotności statystycznej deskryptor w przedstawianym badaniu. Kolejną zidentyfikowaną właściwością cieczy jonowych, która ma wpływ na wartość  $H_{CO_2}$ , jest obecność akceptorów wiązań wodorowych, zarówno w cząsteczce kationu, jak i anionu. Zależność ta jest jednak odwrotna – w kationie obecność akceptorów wiązań H-B ma negatywny wpływ na rozpuszczalność, a w anionie pozytywny. Choć udowodnione zostało, że słabe oddziaływania C-H...O zwiększają stabilność kompleksów  $CO_2$  ze związkami organicznymi, w przypadku cieczy jonowych interakcje pomiędzy jonami, ich wzajemne ułożenie oraz stabilizacja przestrzenna poprzez tworzenie mostków wodorowych pomiędzy podstawnikami może okazać się ważniejsza. Kolejna ważna obserwacja dotyczy deskryptorów Chi\_G/D oraz GATS2s, które charakteryzują się najmniejszym, choć istotnym wpływem na wartość  $H_{CO_2}$ . Wartość Chi\_G/D zmniejsza się wraz ze wzrostem stopnia rozgałęzienia związku, a wartość GATS2s zależy od całkowitej liczby wiązań

w cząsteczce, co przekłada się na zajmowaną przez nią objętość. Obie te cechy pozytywnie wpływają na zdolność cieczy jonowych do rozpuszczania ditlenku węgla. Wolna objętość cząsteczkowa (*ang. Free Volume, FV*) jest wymieniana jako jeden z czynników posiadających prawdopodobny wpływ na rozpuszczalność CO<sub>2</sub>, a oba wspomniane deskryptory zdają się być z nią powiązane. Co więcej, udowodniony wpływ zdolności do tworzenia wiązań wodorowych oraz wzajemne interakcje jonów również przekładają się na teoretyczne występowanie w strukturach cieczy wolnych przestrzeni, które mogą zostać zajęte przez cząsteczki CO<sub>2</sub>.

## 5.2. Analiza wolnej objętości

Frakcyjne wolne objętości 73 cieczy jonowych zostały obliczone z wykorzystaniem dynamiki molekularnej, a proces ten wymagał czasu oraz znacznych zasobów obliczeniowych. Bazując na wynikach symulacji, zbudowano model QSPR, który służył do przewidywania wartości FFV bez konieczności przeprowadzania kolejnych obliczeń MD. Parametr ten okazał się jednak zbyt złożony, by przybliżyć go za pomocą deskryptorów, a otrzymany model nie spełnił wymagań statystycznych. W związku z tym dynamika molekularna pozostaje jedyną rzetelną metodą obliczania FFV *in silico*. Analizując wpływ budowy cieczy jonowej na wartości FFV potwierdzono, że kation wywiera na nią zdecydowanie większy wpływ od anionu.

Parametr FFV wykorzystano następnie do zbudowania modelu QSPR służącego do przewidywania HCO<sub>2</sub> i porównano go z identycznie przygotowanym modelem, w którym nie uwzględniono wpływu wolnej objętości. Oba modele charakteryzowały się zadowalającymi parametrami statystycznymi, lecz dodanie FFV jako zmiennej podczas modelowania poskutkowało ich znaczącą poprawą, w szczególności w przypadku walidacji wewnętrznej. Model standardowy składał się w większości z deskryptorów kationowych, natomiast proporcje w modelu FFV były odwrotne. Oznacza to, że uwzględnienie wolnej objętości, która zależy głównie od kationu, pozwoliło na dokładniejszą analizę subtelniejszego wpływu anionów na rozpuszczalność CO<sub>2</sub> w cieczach jonowych.

Do najważniejszych cech anionów, które obniżają wartość HCO<sub>2</sub> zaliczono brak grup -OH oraz wiązań C-O w strukturze cząsteczki, obecność atomów fluoru oraz dłuższych łańcuchów węglowych, niewielką liczbę donorów wiązania wodorowego oraz wysoki potencjał lekopodobności.

## 5.3. Analiza sposobów optymalizacji

W celu analizy wpływu sposobu optymalizacji cieczy jonowych na otrzymywane modele QSPR zbudowano 32 modele uwzględniające dwie zmienne (HCO<sub>2</sub> oraz FFV), cztery metody modelowania (MLR, PLS, LR oraz SVM) oraz cztery metody optymalizacji (jony razem i osobno oraz w próżni i w etanolu).

Modele MLR oraz PLS okazały się najbardziej stabilne oraz charakteryzowały się najlepszymi parametrami statystycznymi dla obu zmiennych. Metoda LR okazała się najbardziej wrażliwa na zmiany sposobu optymalizacji. Metoda SVM nie jest rekomendowana w celu modelowania wartości HCO<sub>2</sub>. RMSE jest najmniejsze dla metod MLR.

Porównując optymalizację jonów cieczy jonowej osobno i razem, zauważyć można, że modele zoptymalizowane razem są mniej stabilne i bardziej wrażliwe na zmiany pozostałych parametrów. Z drugiej strony gwarantują one lepsze wyniki walidacji zewnętrznej, a więc zapewniają lepszą zdolność predykcyjną modeli.

Optymalizacja cząsteczek w środowisku etanolu nie wpływa znacząco na parametry statystyczne otrzymanych modeli w porównaniu do tych optymalizowanych w próżni. Jedynym parametrem ulegającym nieznacznemu polepszeniu jest R<sup>2</sup>pred, czyli walidacja zewnętrzna. Zauważyć należy jednak, że przeprowadzenie optymalizacji w etanolu powoduje wzrost istotności deskryptorów 3D, przez co modele QSPR opierają się na nich w większej mierze, stając się jednocześnie trudniejszymi do interpretacji. Z drugiej strony pozwala to na zauważenie bardziej skomplikowanych lub subtelnych zależności. Optymalizacja w próżni zapewnia za to niższe wartości RMSE.

Wartość HCO<sub>2</sub> można przewidzieć metodą QSPR na wiele różnych sposobów i za każdym razem uzyskać satysfakcjonujące parametry statystyczne modeli. FFV wciąż pozostaje zbyt specyficznym parametrem, aby dało się przewidywać go bez przeprowadzania długich symulacji dynamiki molekularnej. Po raz kolejny udowodniono jednak, że uwzględnienie obliczonego parametru FFV w modelowaniu pozwala na dokładniejsze przewidywanie wartości HCO<sub>2</sub> ditlenku węgla w cieczach jonowych. Szersza użyteczność predykcyjna FFV jako źródła danych wymaga dalszych doświadczeń.

Aby ocenić wpływ sposobów optymalizacji cząsteczek na obliczone wartości deskryptorów molekularnych porównano wartości 5270 deskryptorów obliczonych dla jonu [NTf<sub>2</sub>] dla 26 cieczy w trzech wariantach: optymalizacji osobno, optymalizacji razem w próżni oraz razem w etanolu.

Sposób optymalizacji nie ma żadnego wpływu na deskryptory należące do kategorii 2D, natomiast ok. 60% deskryptorów 3D wykazuje różnice w wartościach w zależności od sposobu optymalizacji. Niezależnie od wybranego sposobu, rozkład różnic procentowych deskryptorów przypomina rozkład normalny.

Analizując bezwzględne wartości deskryptorów, do grup najbardziej wrażliwych na zmiany w optymalizacji należą deskryptory RDF oraz 3D-MoRSE, które charakteryzowały się zarówno najwyższymi wartościami współczynnika zmienności, jak i największą liczbą podatnych na zmiany deskryptorów. Analizując zmienność względną, okazuje się jednak, że w ramach różnych metod optymalizacji struktur

ogólny trend zmian wartości deskryptorów pozostaje w większości przypadków (wliczając grupy RDF oraz 3D-MoRSE) niemal niezaburzony, a spadki i wzrosty ich wartości są porównywalne. W tym przypadku do najbardziej wrażliwych należy kategoria CATS3D powiązana z występowaniem w strukturze cząsteczki potencjalnych farmakoforów.

Podsumowując, zmiana środowiska i sposobu optymalizacji struktur cząsteczek wywiera mniejszy wpływ na parametry otrzymywanych modeli QSPR, niż mogłoby się wydawać, a kluczowym okazuje się wybór odpowiedniej metody modelowania.

## 6. PODSUMOWANIE

W przedstawionej rozprawie omówiono istotne z praktycznego punktu widzenia właściwości cieczy jonowych oraz przedstawiono modelowanie QSPR jako metodą mogącą służyć do uzupełnienia luk w danych eksperymentalnych dotyczących ich właściwości fizykochemicznych.

W badaniach skupiono się na trzech aspektach:

1. Z powodzeniem wykorzystano modelowanie QSPR w celu przewidywania wartości stałej Henry'ego ditlenku węgla w cieczach jonowych. Ponadto, przeprowadzono analizę otrzymanych modeli matematycznych i na podstawie interpretacji deskryptorów molekularnych zidentyfikowano cechy strukturalne wpływające na rozpuszczalność CO<sub>2</sub> w cieczach jonowych.
2. Przeprowadzono analizę wpływu wolnej objętości w cieczach jonowych na ich zdolność do rozpuszczania CO<sub>2</sub>. Udowodniono, że wolna objętość w cieczach jonowych jest zależna głównie od struktury kationu. Uwzględnienie wolnej objętości w modelowaniu QSPR pozwoliło na odkrycie i analizę subtelного wpływu anionów na rozpuszczalność CO<sub>2</sub>. Zinterpretowano deskryptory molekularne i zidentyfikowane kolejne cechy strukturalne wpływające na rozpuszczalność CO<sub>2</sub> w cieczach jonowych. Udowodniono także, że do wyznaczenia wolnej objętości niezbędne są obliczenia dynamiki molekularnej, gdyż jest to zbyt złożony parametr, by udało się przybliżyć go modelowaniem QSPR.
3. Zbadano wpływ optymalizacji struktur cieczy jonowych na parametry otrzymywanych modeli QSPR w zależności od sposobu optymalizacji (jony osobno lub razem), środowiska optymalizacji (w próżni lub w etanolu) oraz od wybranej metody modelowania (MLR, PLS, LR, SVM). Udowodniono, że optymalizacja jonów razem nieznacznie zwiększa predykcyjność modeli, a optymalizacja w etanolu nie ma większego wpływu na ich parametry. Z drugiej strony wykazano, że optymalizacja w etanolu sprawia, że otrzymywane modele budowane są w większej mierze z deskryptorów 3D. W ostatnim etapie badań zbadano wrażliwość deskryptorów molekularnych na zmiany w sposobach optymalizacji struktur. Udowodniono, że deskryptory dwuwymiarowe są niepodatne na zmiany, a większa część deskryptorów trójwymiarowych charakteryzuje się sporą zmiennością w zależności od wybranego sposobu optymalizacji, lecz w ogólnym rozrachunku wartości deskryptorów zmieniają się z zachowaniem tego samego trendu dla każdej z metod.

W przyszłości autor planuje wykorzystać zgromadzoną wiedzę na temat modelowania QSPR oraz cieczy jonowych i skupić się badaniu i modelowaniu toksyczności cieczy jonowych wobec bakterii występujących w jamie ustnej oraz na analizie możliwości wykorzystania cieczy jonowych jako środki bakteriobójcze lub bakteriostatyczne w materiałach stomatologicznych.

## 7. BIBLIOGRAFIA

1. Seddon, K.R. A Taste of the Future. *Nature Materials* 2003 2:6 **2003**, 2, 363–365, doi:10.1038/nmat907.
2. Earle, M.J.; Seddon, K.R. Ionic Liquids. Green Solvents for the Future. In Proceedings of the Pure and Applied Chemistry; Walter de Gruyter GmbH, **2000**; Vol. 72, pp. 1391–1398.
3. Zhao, D.; Liao, Y.; Zhang, Z. Toxicity of Ionic Liquids. *Clean (Weinh)* **2007**, 35, 42–48, doi:10.1002/CLEN.200600015.
4. Flieger, J.; Flieger, M. Ionic Liquids Toxicity—Benefits and Threats. *International Journal of Molecular Sciences* 2020, Vol. 21, Page 6267 **2020**, 21, 6267, doi:10.3390/IJMS21176267.
5. Thuy Pham, T.P.; Cho, C.W.; Yun, Y.S. Environmental Fate and Toxicity of Ionic Liquids: A Review. *Water Res* **2010**, 44, 352–372, doi:10.1016/J.WATRES.2009.09.030.
6. Alexander H. Tullo For Ionic Liquids, the Time Is Now. *C&EN Global Enterprise* **2020**, 98, 24–27, doi:10.1021/CEN-09805-FEATURE2.
7. Walden, P. Ueber Die Molekulargrösse Und Elektrische Leitfähigkeit Einiger Geschmolzenen Salze. *Bulletin del'Academie Imperiale des Sciences de St.-Petersbourg*. **1914**, 6, 405–422.
8. Barrer, R.M. The Viscosity of Pure Liquids. II. Polymerised Ionic Melts. *Transactions of the Faraday Society* **1943**, 39, 59–67, doi:10.1039/TF9433900059.
9. Hurley, F.H.; Wler, T.P. The Electrodeposition of Aluminum from Nonaqueous Solutions at Room Temperature. *J Electrochem Soc* **1951**, 98, 207, doi:10.1149/1.2778133.
10. Li Chum, H.; Koch, V.R.; Miller, L.L.; Osteryoung, R.A. An Electrochemical Scrutiny of Organometallic Iron Complexes and Hexamethylbenzene in a Room Temperature Molten Salt. *J Am Chem Soc* **1975**, 97, 3264–3265, doi:10.1021/JA00844A081.
11. Wilkes, J.S.; Levisky, J.A.; Wilson, R.A.; Hussey, C.L. Dialkylimidazolium Chloroaluminate Melts: A New Class of Room-Temperature Ionic Liquids for Electrochemistry, Spectroscopy, and Synthesis. *Inorg Chem* **1982**, 21, 1263–1264, doi:10.1021/IC00133A078.
12. Robinson, J.; Osteryoung, R.A. An Electrochemical and Spectroscopic Study of Some Aromatic Hydrocarbons in the Room Temperature Molten Salt System Aluminum Chloride-n-Butylpyridinium Chloride. *J Am Chem Soc* **1979**, 101, 323–327, doi:10.1021/JA00496A008.
13. Wilkes, J.S.; Zaworotko, M.J. Air and Water Stable 1-Ethyl-3-Methylimidazolium Based Ionic Liquids. *J Chem Soc Chem Commun* **1992**, 965–967, doi:10.1039/C39920000965.
14. Gordon, C.M. New Developments in Catalysis Using Ionic Liquids. *Appl Catal A Gen* **2001**, 222, 101–117, doi:10.1016/S0926-860X(01)00834-1.
15. Welton, T. Ionic Liquids: A Brief History. *Biophysical Reviews* 2018 10:3 **2018**, 10, 691–706, doi:10.1007/S12551-018-0419-2.
16. Anastas, P.; Eghbali, N. Green Chemistry: Principles and Practice. *Chem Soc Rev* **2010**, 39, 301–312, doi:10.1039/b918763b.

17. Earle, M.J.; Esperança, J.M.S.S.; Gilea, M.A.; Canongia Lopes, J.N.; Rebelo, L.P.N.; Magee, J.W.; Seddon, K.R.; Widegren, J.A. The Distillation and Volatility of Ionic Liquids. *Nature* **2005** *439*:7078 **2006**, 439, 831–834, doi:10.1038/nature04451.
18. Nanda, B.; Sailaja, M.; Mohapatra, P.; Pradhan, R.K.; Nanda, B.B. Green Solvents: A Suitable Alternative for Sustainable Chemistry. *Mater Today Proc* **2021**, *47*, 1234–1240, doi:10.1016/J.MATPR.2021.06.458.
19. Aghaie, M.; Rezaei, N.; Zendehboudi, S. A Systematic Review on CO<sub>2</sub> Capture with Ionic Liquids: Current Status and Future Prospects. *Renewable and Sustainable Energy Reviews* **2018**, *96*, 502–525, doi:10.1016/J.RSER.2018.07.004.
20. Rajadurai, V.; Anguraj, B.L. Ionic Liquids to Remove Toxic Metal Pollution. *Environmental Chemistry Letters* **2020** *19*:2 **2020**, *19*, 1173–1203, doi:10.1007/S10311-020-01115-5.
21. Kumar, S.S.; Kumar, S.R. Ionic Liquids as Environmental Friendly Cutting Fluids - a Review. *Mater Today Proc* **2021**, *37*, 2121–2125, doi:10.1016/J.MATPR.2020.07.564.
22. Munir, N.; Awang, R.A.; Muhammad, N.; Ismail, N.H.; Khan, A.S.; Inayat, N. Potential Applications of Ionic Liquids in Dentistry: A Narrative Review. *ChemBioEng Reviews* **2023**, *10*, 1073–1082, doi:10.1002/CBEN.202300012.
23. Wasserscheid, P.; Welton, T. *Ionic Liquids in Synthesis: Second Edition*; John Wiley and Sons, **2008**; Vol. 1;.
24. Alreshidi, M.A.; Yadav, K.K.; Gunasekaran, S.; Gacem, A.; Sambandam, P.; Subbiah, G.; Bhutto, J.K.; Palanivel, S.; Fallatah, A.M.; Abo El-Khair, M.A.; et al. A Review on the Evolution of Ionic Liquids: Sustainable Synthesis, Applications, and Future Prospects. *Materials Today Sustainability* **2025**, *31*, 101160, doi:10.1016/J.MTSUST.2025.101160.
25. Kolbeck, C.; Niedermaier, I.; Deyko, A.; Lovelock, K.R.J.; Taccardi, N.; Wei, W.; Wasserscheid, P.; Maier, F.; Steinrück, H.P. Influence of Substituents and Functional Groups on the Surface Composition of Ionic Liquids. *Chemistry – A European Journal* **2014**, *20*, 3954–3965, doi:10.1002/CHEM.201304549.
26. Xue, H.; Shreeve, J.M. Ionic Liquids with Fluorine-Containing Cations. *Eur J Inorg Chem* **2005**, *2005*, 2573–2580, doi:10.1002/EJIC.200500129.
27. Schubert, T.J.S. Commercial Production of Ionic Liquids. In *Commercial Applications of Ionic Liquids*; Shiflett, M.B., Ed.; Springer International Publishing: Cham, **2020**; pp. 191–208 ISBN 978-3-030-35245-5.
28. Heiko Niedermeyer; P. Hallett, J.; J. Villar-Garcia, I.; A. Hunt, P.; Tom Welton Mixtures of Ionic Liquids. *Chem Soc Rev* **2012**, *41*, 7780–7802, doi:10.1039/C2CS35177C.
29. Al-Bodour, A.; Alomari, N.; Perez-Duran, G.; Rozas, S.; Iglesias-Silva, G.; Aparicio, S.; Gutierrez-Vega, A.; Atilhan, M. Ionic Liquids as Multidimensional Materials: A Review from Fundamentals to Applications. *Energy & Fuels* **2025**, *39*, 12791–12829, doi:10.1021/ACS.ENERGYFUELS.5C01280.
30. de Jesus, S.S.; Maciel Filho, R. Are Ionic Liquids Eco-Friendly? *Renewable and Sustainable Energy Reviews* **2022**, *157*, 112039, doi:10.1016/J.RSER.2021.112039.

31. Kosmulski, M.; Gustafsson, J.; Rosenholm, J.B. Thermal Stability of Low Temperature Ionic Liquids Revisited. *Thermochim Acta* **2004**, *412*, 47–53, doi:10.1016/J.TCA.2003.08.022.
32. Kaur, N.; Mithu, V.S.; Kumar, S. A Review on (Eco)Toxicity of Ionic Liquids and Their Interaction with Phospholipid Membranes. *J Mol Liq* **2024**, *397*, 124095, doi:10.1016/J.MOLLIQ.2024.124095.
33. Ewelina Wyrzykowska; Anna Rybińska-Fryca; Anita Sosnowska; Tomasz Puzyn Virtual Screening in the Design of Ionic Liquids as Environmentally Safe Bactericides. *Green Chemistry* **2019**, *21*, 1965–1973, doi:10.1039/C8GC03400A.
34. Anita Sosnowska; Maciej Barycki; Monika Zaborowska; Anna Rybinska; Tomasz Puzyn Towards Designing Environmentally Safe Ionic Liquids: The Influence of the Cation Structure. *Green Chemistry* **2014**, *16*, 4749–4757, doi:10.1039/C4GC00526K.
35. Rybinska, A.; Sosnowska, A.; Grzonkowska, M.; Barycki, M.; Puzyn, T. Filling Environmental Data Gaps with QSPR for Ionic Liquids: Modeling n-Octanol/Water Coefficient. *J Hazard Mater* **2016**, *303*, 137–144, doi:10.1016/J.JHAZMAT.2015.10.023.
36. V. Kapitanov, I.; Andrew Jordan; Yevgen Karpichev; Marcel Spulak; Lourdes Perez; Andrew Kellett; Klaus Kümmerer; Nicholas Gathergood Synthesis, Self-Assembly, Bacterial and Fungal Toxicity, and Preliminary Biodegradation Studies of a Series of L-Phenylalanine-Derived Surface-Active Ionic Liquids. *Green Chemistry* **2019**, *21*, 1777–1794, doi:10.1039/C9GC00030E.
37. Fan, M.; Ma, L.; Zhang, C.; Wang, Z.; Ruan, J.; Han, M.; Ren, Y.; Zhang, C.; Yang, D.; Zhou, F.; et al. Biobased Green Lubricants: Physicochemical, Tribological and Toxicological Properties of Fatty Acid Ionic Liquids. <https://doi.org/10.1080/10402004.2017.1290856> **2017**, *61*, 195–206, doi:10.1080/10402004.2017.1290856.
38. Greer, A.J.; Jacquemin, J.; Hardacre, C. Industrial Applications of Ionic Liquids. *Molecules* **2020**, *Vol. 25*, Page 5207 **2020**, *25*, 5207, doi:10.3390/MOLECULES25215207.
39. Freemantle, M. BASF's Smart Ionic Liquid. *Chemical and Engineering News* **2003**, *81*, 9, doi:10.1021/CEN-V081N013.P009.
40. Abbott, A.P.; Davies, D.L.; Capper, G.; Rasheed, R.K.; Tambyrajah, V. Ionic Liquids and Their Use **2002**, pat. WO2002026381A2.
41. Moganty, S.S.; Lee, J. Hybrid Ionic Liquid Electrolytes **2016**, pat. US20160164137A1.
42. Timken, H.K.C.; Elomari, S.; Trumbull, S.; Cleverdon, R. Integrated Alkylation Process Using Ionic Liquid Catalysts **2006**, WO2006073749A2.
43. Michael McCoy Chevron Embraces Ionic Liquids. *C&EN Global Enterprise* **2016**, *94*, 16–16, doi:10.1021/CEN-09439-NOTW11.
44. 3M™ Ionic Liquid Antistat FC-4400 | 3M United States Available online: [https://www.3m.com/3M/en\\_US/p/d/b10178970/](https://www.3m.com/3M/en_US/p/d/b10178970/) (dostęp 30 sierpnia 2021).
45. Home - Ioniqa Available online: <https://ioniqa.com/> (dostęp on 30 sierpnia 2021).
46. Kalb, R.S. Toward Industrialization of Ionic Liquids. **2020**, 261–282, doi:10.1007/978-3-030-35245-5\_11.

47. Dai, Z.; Wang, L.; Lu, X.; Ji, X. Melting Points of Ionic Liquids: Review and Evaluation. *Green Energy & Environment* **2024**, *9*, 1802–1811, doi:10.1016/J.GEE.2024.01.009.
48. López-Martin, I.; Burello, E.; Davey, P.N.; Seddon, K.R.; Rothenberg, G. Anion and Cation Effects on Imidazolium Salt Melting Points: A Descriptor Modelling Study. *ChemPhysChem* **2007**, *8*, 690–695, doi:10.1002/CPHC.200600637.
49. Lerch, S.; Strassner, T. Synthesis and Physical Properties of Tunable Aryl Alkyl Ionic Liquids (TAAILs). *Chemistry - A European Journal* **2021**, *27*, 15554–15557, doi:10.1002/CHEM.202102545.
50. Myrdek, T.; Popescu, C.; Kunz, W. Physical-Chemical Properties of Newly Synthesized Tetraalkylammonium Alkyl Ether Carboxylate Ionic Liquids. *J Mol Liq* **2021**, *322*, 114947, doi:10.1016/J.MOLLIQ.2020.114947.
51. Del Río, E.; Vidil, T.; Gati, W.; Grau, É.; Taton, D.; Cramail, H. Ester-Containing Imidazolium-Type Ionic Liquid Crystals Derived from Bio-Based Fatty Alcohols. *ACS Sustain Chem Eng* **2021**, *9*, 12687–12698, doi:10.1021/ACSSUSCHEMENG.1C05122/SUPPL\_FILE/SC1C05122\_SI\_004.CIF.
52. Niemczak, M.; Sobiech, Ł.; Grzanka, M. Iodosulfuron-Methyl-Based Herbicidal Ionic Liquids Comprising Alkyl Betainate Cation as Novel Active Ingredients with Reduced Environmental Impact and Excellent Efficacy. *J Agric Food Chem* **2020**, *68*, 13661–13671, doi:10.1021/ACS.JAFC.0C05850.
53. Freire, M.G.; Neves, C.M.S.S.; Carvalho, P.J.; Gardas, R.L.; Fernandes, A.M.; Marrucho, I.M.; Santos, L.M.N.B.F.; Coutinho, J.A.P. Mutual Solubilities of Water and Hydrophobic Ionic Liquids. *Journal of Physical Chemistry B* **2007**, *111*, 13082–13089, doi:10.1021/JP076271E.
54. Wang, L.S.; Wang, X.X.; Li, Y.; Jiang, K.; Shao, X.Z.; Du, C.J. Ionic Liquids: Solubility Parameters and Selectivities for Organic Solutes. *AIChE Journal* **2013**, *59*, 3034–3041, doi:10.1002/AIC.14047.
55. Hu, Y.; Xing, Y.; Yue, H.; Chen, T.; Diao, Y.; Wei, W.; Zhang, S. Ionic Liquids Revolutionizing Biomedicine: Recent Advances and Emerging Opportunities. *Chem Soc Rev* **2023**, *52*, 7262–7293, doi:10.1039/D3CS00510K.
56. McCrary, P.D.; Beasley, P.A.; Gurau, G.; Narita, A.; Barber, P.S.; Cojocaru, O.A.; Rogers, R.D. Drug Specific, Tuning of an Ionic Liquid's Hydrophilic–Lipophilic Balance to Improve Water Solubility of Poorly Soluble Active Pharmaceutical Ingredients. *New Journal of Chemistry* **2013**, *37*, 2196–2202, doi:10.1039/C3NJ00454F.
57. Canales, R.I.; Brennecke, J.F. Comparison of Ionic Liquids to Conventional Organic Solvents for Extraction of Aromatics from Aliphatics. *J Chem Eng Data* **2016**, *61*, 1685–1699, doi:10.1021/ACS.JCED.6B00077.
58. Sakhtemanian, L.; Duwadi, A.; Baldelli, S.; Ghatee, M.H. Simulating the Ionic Liquid Mixing with Organic-Solvent Clarifies the Mixture's SFG Spectral Behavior and the Specific Surface Region Originating SFG. *Sci Rep* **2024**, *14*, 1–14, doi:10.1038/S41598-024-74561-8;SUBJMETA.
59. Gaciño, F.M.; Regueira, T.; Lugo, L.; Comuñas, M.J.P.; Fernández, J. Influence of Molecular Structure on Densities and Viscosities of Several Ionic Liquids. *J Chem Eng Data* **2011**, *56*, 4984–4999, doi:10.1021/JE200883W.

60. Silva, W.; Zanatta, M.; Ferreira, A.S.; Corvo, M.C.; Cabrita, E.J. Revisiting Ionic Liquid Structure-Property Relationship: A Critical Analysis. *International Journal of Molecular Sciences* **2020**, Vol. 21, Page 7745 **2020**, 21, 7745, doi:10.3390/IJMS21207745.
61. Hayes, R.; Warr, G.G.; Atkin, R. Structure and Nanostructure in Ionic Liquids. *Chem Rev* **2015**, 115, 6357–6426, doi:10.1021/CR500411Q.
62. Kolbeck, C.; Lehmann, J.; Lovelock, K.R.J.; Cremer, T.; Paape, N.; Wasserscheid, P.; Fröba, A.P.; Maier, F.; Steinrück, H.P. Density and Surface Tension of Ionic Liquids. *Journal of Physical Chemistry B* **2010**, 114, 17025–17036, doi:10.1021/JP1068413.
63. Darabi, L.; Zare, M. Theoretical Study on the Structure and Electronic Properties of Alkylimidazolium Iodide Ionic Liquids: The Effect of Alkyl Chain Length. *New Journal of Chemistry* **2020**, 44, 4023–4032, doi:10.1039/C9NJ05687D.
64. Yu, G.; Wen, L.; Zhao, D.; Asumana, C.; Chen, X. QSPR Study on the Viscosity of Bis(Trifluoromethylsulfonyl)Imide-Based Ionic Liquids. *J Mol Liq* **2013**, 184, 51–59, doi:10.1016/J.MOLLIQ.2013.04.021.
65. Seki, S.; Kobayashi, T.; Kobayashi, Y.; Takei, K.; Miyashiro, H.; Hayamizu, K.; Tsuzuki, S.; Mitsugi, T.; Umebayashi, Y. Effects of Cation and Anion on Physical Properties of Room-Temperature Ionic Liquids. *J Mol Liq* **2010**, 152, 9–13, doi:10.1016/J.MOLLIQ.2009.10.008.
66. Laurenczy, G.; Dyson, P.J. Determination of the Viscosity of the Ionic Liquids [Bmim][PF<sub>6</sub>] and [Bmim][TF<sub>2</sub>N] under High CO<sub>2</sub> Gas Pressure Using Sapphire NMR Lubes. *Zeitschrift für Naturforschung - Section B Journal of Chemical Sciences* **2008**, 63, 681–684, doi:10.1515/ZNB-2008-0614/MACHINEREADABLECITATION/RIS.
67. Ghatee, M.H.; Bahrami, M.; Khanjari, N. Measurement and Study of Density, Surface Tension, and Viscosity of Quaternary Ammonium-Based Ionic Liquids ([N222(n)]Tf<sub>2</sub>N). *J Chem Thermodyn* **2013**, 65, 42–52, doi:10.1016/J.JCT.2013.05.031.
68. Almeida, H.F.D.; Freire, M.G.; Fernandes, A.M.; Lopes-Da-Silva, J.A.; Morgado, P.; Shimizu, K.; Filipe, E.J.M.; Canongia Lopes, J.N.; Santos, L.M.N.B.F.; Coutinho, J.A.P. Cation Alkyl Side Chain Length and Symmetry Effects on the Surface Tension of Ionic Liquids. *Langmuir* **2014**, 30, 6408–6418, doi:10.1021/LA501308Q.
69. Tariq, M.; Freire, M.G.; Saramago, B.; Coutinho, J.A.P.; Lopes, J.N.C.; Rebelo, L.P.N. Surface Tension of Ionic Liquids and Ionic Liquid Solutions. *Chem Soc Rev* **2012**, 41, 829–868, doi:10.1039/C1CS15146K.
70. Cui, G.; Wang, J.; Zhang, S. Active Chemisorption Sites in Functionalized Ionic Liquids for Carbon Capture. *Chem Soc Rev* **2016**, 45, 4307–4339, doi:10.1039/C5CS00462D.
71. Blath, J.; Deubler, N.; Hirth, T.; Schiestel, T. Chemisorption of Carbon Dioxide in Imidazolium Based Ionic Liquids with Carboxylic Anions. *Chemical Engineering Journal* **2012**, 181–182, 152–158, doi:10.1016/J.CEJ.2011.11.042.
72. Lei, Z.; Dai, C.; Chen, B. Gas Solubility in Ionic Liquids. *Chem Rev* **2013**, 114, 1289–1326, doi:10.1021/CR300497A.
73. Huang, J.; Riisager, A.; Berg, R.W.; Fehrmann, R. Tuning Ionic Liquids for High Gas Solubility and Reversible Gas Sorption. *J Mol Catal A Chem* **2008**, 279, 170–176, doi:10.1016/J.MOLCATA.2007.07.036.

74. Anthony, J.L.; Anderson, J.L.; Maginn, E.J.; Brennecke, J.F. Anion Effects on Gas Solubility in Ionic Liquids. *J Phys Chem B* **2005**, *109*, 6366–6374, doi:10.1021/jp046404l.
75. Hong, G.; Jacquemin, J.; Deetlefs, M.; Hardacre, C.; Husson, P.; Costa Gomes, M.F. Solubility of Carbon Dioxide and Ethane in Three Ionic Liquids Based on the Bis((Trifluoromethyl)Sulfonyl)imide Anion. *Fluid Phase Equilib* **2007**, *257*, 27–34, doi:10.1016/J.FLUID.2007.05.002.
76. Gomez-Herrero, E.; Tobajas, M.; Polo, A.; Rodriguez, J.J.; Mohedano, A.F. Toxicity and Inhibition Assessment of Ionic Liquids by Activated Sludge. *Ecotoxicol Environ Saf* **2020**, *187*, 109836, doi:10.1016/J.ECOENV.2019.109836.
77. Kuroda, K. A Simple Overview of Toxicity of Ionic Liquids and Designs of Biocompatible Ionic Liquids. *New Journal of Chemistry* **2022**, *46*, 20047–20052, doi:10.1039/D2NJ02634A.
78. Flieger, J.; Flieger, M. Ionic Liquids Toxicity—Benefits and Threats. *International Journal of Molecular Sciences* **2020**, Vol. 21, Page 6267 **2020**, *21*, 6267, doi:10.3390/IJMS21176267.
79. Gonçalves, A.R.P.; Paredes, X.; Cristino, A.F.; Santos, F.J.V.; Queirós, C.S.G.P. Ionic Liquids—A Review of Their Toxicity to Living Organisms. *International Journal of Molecular Sciences* **2021**, Vol. 22, Page 5612 **2021**, *22*, 5612, doi:10.3390/IJMS22115612.
80. Md Moshikur, R.; Chowdhury, M.R.; Moniruzzaman, M.; Goto, M. Biocompatible Ionic Liquids and Their Applications in Pharmaceuticals. *Green Chemistry* **2020**, *22*, 8116–8139, doi:10.1039/D0GC02387F.
81. Tzani, A.; Karadendrou, M.A.; Kalafateli, S.; Kakokefalou, V.; Detsi, A. Current Trends in Green Solvents: Biocompatible Ionic Liquids. *Crystals* **2022**, Vol. 12, Page 1776 **2022**, *12*, 1776, doi:10.3390/CRYST12121776.
82. Roy, K.; Kar, S.; Das, R.N. QSAR/QSPR Modeling: Introduction. In *A Primer on QSAR/QSPR Modeling: Fundamental Concepts*; Springer International Publishing: Cham, **2015**; pp. 1–36 ISBN 978-3-319-17281-1.
83. Gramatica, P. On the Development and Validation of QSAR Models. *Methods in Molecular Biology* **2013**, *930*, 499–526, doi:10.1007/978-1-62703-059-5\_21.
84. Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics*; Methods and Principles in Medicinal Chemistry; Wiley, **2009**; Vol. 41; ISBN 9783527318520.
85. Mauri, A. AlvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints BT - Ecotoxicological QSARs. **2020**, 801–820, doi:10.1007/978-1-0716-0150-1\_32.
86. Mauri, A.; Consonni, V.; Pavan, M.; Todeschini, R. Dragon Software: An Easy Approach to Molecular Descriptor Calculations. *MATCH-COMMUNICATIONS IN MATHEMATICAL AND IN COMPUTER CHEMISTRY* **2006**, *56*, 237–248.
87. Fujitsu Kyushu Systems Ltd. ADMEWORKS ModelBuilder **2010**.
88. Demšar, J.; Curk, T.; Erjavec, A.; Gorup, Č.; Hočevár, T.; Milutinovič, M.; Možina, M.; Polajnar, M.; Toplak, M.; Starič, A.; et al. Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research* **2013**, *14*, 2349–2353.
89. Gissi, A.; Tcheremenskaia, O.; Bossa, C.; Battistelli, C.L.; Browne, P. The OECD (Q)SAR Assessment Framework: A Tool for Increasing Regulatory Uptake of Computational

- Approaches. *Computational Toxicology* **2024**, *31*, 100326, doi:10.1016/J.COMTOX.2024.100326.
90. OECD OECD Principles for the Validation, for Regulatory Purposes, of (Q)SAR Models. In *OECD Series on Testing and Assessment*; Paris, **2004**; Vol. 49.
  91. Gramatica, P. Origin of the OECD Principles for QSAR Validation and Their Role in Changing the QSAR Paradigm Worldwide: An Historical Overview. *J Chemom* **2025**, *39*, e70014, doi:10.1002/CEM.70014.
  92. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. **2014**, doi:10.1787/9789264085442-EN.
  93. Boczar, D.; Michalska, K. A Review of Machine Learning and QSAR/QSPR Predictions for Complexes of Organic Molecules with Cyclodextrins. *Molecules* **2024**, Vol. 29, Page 3159 **2024**, *29*, 3159, doi:10.3390/MOLECULES29133159.
  94. Lewars, E.G. Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics* **2011**, 1–664, doi:10.1007/978-90-481-3862-3/COVER.
  95. Bursch, M.; Mewes, J.M.; Hansen, A.; Grimme, S. Best-Practice DFT Protocols for Basic Molecular Computational Chemistry. *Angewandte Chemie International Edition* **2022**, *61*, e202205735, doi:10.1002/ANIE.202205735.
  96. Mester, D.; Kállay, M. Basis Set Limit of CCSD(T) Energies: Explicit Correlation Versus Density-Based Basis-Set Correction. *J Chem Theory Comput* **2023**, *19*, 8210–8222, doi:10.1021/ACS.JCTC.3C00979.
  97. Bálint, D.; Jäntschi, L. Comparison of Molecular Geometry Optimization Methods Based on Molecular Descriptors. *Mathematics* **2021**, Vol. 9, Page 2855 **2021**, *9*, 2855, doi:10.3390/MATH9222855.
  98. Newhouse, P.F.; McGill, K.C. Schrödinger Equation Solutions That Lead to the Solution for the Hydrogen Atom. *J Chem Educ* **2004**, *81*, 424–426, doi:10.1021/ED081P424.
  99. Mitin, A. V. Polarization Functions for the Modified M6-31G Basis Sets for Atoms Ga through Kr. *J Comput Chem* **2013**, *34*, 2014–2019, doi:10.1002/JCC.23353.
  100. Dyall, K.G.; Tecmer, P.; Sunaga, A. Diffuse Basis Functions for Relativistic s and d Block Gaussian Basis Sets. *J Chem Theory Comput* **2022**, *19*, 198–210, doi:10.1021/ACS.JCTC.2C01050.
  101. Anthony, J.L.; Anderson, J.L.; Maginn, E.J.; Brennecke, J.F. Anion Effects on Gas Solubility in Ionic Liquids. *J Phys Chem B* **2005**, *109*, 6366–6374, doi:10.1021/JP046404L.
  102. Rybinska, A.; Sosnowska, A.; Barycki, M.; Puzyn, T. Geometry Optimization Method versus Predictive Ability in QSPR Modeling for Ionic Liquids. *Journal of Computer-Aided Molecular Design* **2016**, *30*, 165–176, doi:10.1007/S10822-016-9894-3.
  103. Stewart, J.J.P. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-Optimization of Parameters. *J Mol Model* **2013**, *19*, 1–32, doi:10.1007/S00894-012-1667-X/FIGURES/8.

104. Christensen, A.S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem Rev* **2016**, *116*, 5301, doi:10.1021/ACS.CHEMREV.5B00584.
105. Thiel, W. Semiempirical Quantum–Chemical Methods. *Wiley Interdiscip Rev Comput Mol Sci* **2014**, *4*, 145–157, doi:10.1002/WCMS.1161.
106. Cohen, A.J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem Rev* **2011**, *112*, 289–320, doi:10.1021/CR200107Z.
107. Puzyn, T.; Leszczynski, J.; Cronin, M.T. Recent Advances in QSAR Studies: Methods and Applications (Challenges and Advances in Computational Chemistry and Physics, 8), Springer Dordrecht, **2010**;
108. Keshavarz, M.H.; Pouretedal, H.R.; Saberi, E. A Simple Method for Prediction of Density of Ionic Liquids through Their Molecular Structure. *J Mol Liq* **2016**, *216*, 732–737, doi:10.1016/J.MOLLIQ.2016.01.095.
109. Qiao, Y.; Ma, Y.; Huo, Y.; Ma, P.; Xia, S. A Group Contribution Method to Estimate the Densities of Ionic Liquids. *J Chem Thermodyn* **2010**, *42*, 852–855, doi:10.1016/J.JCT.2010.02.007.
110. Paduszyński, K. Extensive Databases and Group Contribution QSPRs of Ionic Liquids Properties. 1. Density. *Ind Eng Chem Res* **2019**, *58*, 5322–5338, doi:10.1021/ACS.IECR.9B00130.
111. Zhang, S.; Jia, Q.; Yan, F.; Xia, S.; Wang, Q. Evaluating the Properties of Ionic Liquid at Variable Temperatures and Pressures by Quantitative Structure–Property Relationship (QSPR). *Chem Eng Sci* **2021**, *231*, 116326, doi:10.1016/J.CES.2020.116326.
112. Cerecedo-Cordoba, J.A.; Barbosa, J.J.G.; Solís, J.F.; Gallardo-Rivas, N.V. Melting Temperature Estimation of Imidazole Ionic Liquids with Clustering Methods. *J Chem Inf Model* **2019**, *59*, 3144–3153, doi:10.1021/ACS.JCIM.9B00203.
113. Katritzky, A.R.; Jain, R.; Lomaka, A.; Petrukhin, R.; Karelson, M.; Visser, A.E.; Rogers, R.D. Correlation of the Melting Points of Potential Ionic Liquids (Imidazolium Bromides and Benzimidazolium Bromides) Using the CODESSA Program. *J Chem Inf Comput Sci* **2002**, *42*, 225–231, doi:10.1021/CI0100494.
114. Varnek, A.; Kireeva, N.; Tetko, I. V.; Baskin, I.I.; Solov'ev, V.P. Exhaustive QSPR Studies of a Large Diverse Set of Ionic Liquids: How Accurately Can We Predict Melting Points? *J Chem Inf Model* **2007**, *47*, 1111–1122, doi:10.1021/CI600493X.
115. Farahani, N.; Gharagheizi, F.; Mirkhani, S.A.; Tumba, K. Ionic Liquids: Prediction of Melting Point by Molecular-Based Model. *Thermochim Acta* **2012**, *549*, 17–34, doi:10.1016/J.TCA.2012.09.011.
116. Lazzús, J.A. A Group Contribution Method to Predict the Melting Point of Ionic Liquids. *Fluid Phase Equilib* **2012**, *313*, 1–6, doi:10.1016/J.FLUID.2011.09.018.
117. Yan, F.; Xia, S.; Wang, Q.; Yang, Z.; Ma, P. Predicting the Melting Points of Ionic Liquids by the Quantitative Structure Property Relationship Method Using a Topological Index. *J Chem Thermodyn* **2013**, *62*, 196–200, doi:10.1016/J.JCT.2013.03.016.

118. Venkatraman, V.; Evjen, S.; Knuutila, H.K.; Fiksdahl, A.; Alsberg, B.K. Predicting Ionic Liquid Melting Points Using Machine Learning. *J Mol Liq* **2018**, *264*, 318–326, doi:10.1016/J.MOLLIQ.2018.03.090.
119. Paduszyński, K. Extensive Databases and Group Contribution QSPRs of Ionic Liquids Properties. 2. Viscosity. *Ind Eng Chem Res* **2019**, *58*, 17049–17066, doi:10.1021/ACS.IECR.9B03150.
120. Das, S.; Ojha, P.K.; Roy, K. Development of a Temperature Dependent 2D-QSPR Model for Viscosity of Diverse Functional Ionic Liquids. *J Mol Liq* **2017**, *240*, 454–467, doi:10.1016/J.MOLLIQ.2017.05.113.
121. Kang Koi, Z.; Nisa Yahya, W.Z.; Talip, R.A.A.; Adi Kurnia, K. Prediction of the Viscosity of Imidazolium-Based Ionic Liquids at Different Temperatures Using the Quantitative Structure Property Relationship Approach. *New Journal of Chemistry* **2019**, *43*, 16207–16217, doi:10.1039/C9NJ03436F.
122. Yan, F.; He, W.; Jia, Q.; Wang, Q.; Xia, S.; Ma, P. Prediction of Ionic Liquids Viscosity at Variable Temperatures and Pressures. *Chem Eng Sci* **2018**, *184*, 134–140, doi:10.1016/J.CES.2018.03.044.
123. Billard, I.; Marcou, G.; Ouadi, A.; Varnek, A. In Silico Design of New Ionic Liquids Based on Quantitative Structure–Property Relationship Models of Ionic Liquid Viscosity. *Journal of Physical Chemistry B* **2010**, *115*, 93–98, doi:10.1021/JP107868W.
124. Paduszyński, K.; Domańska, U. Viscosity of Ionic Liquids: An Extensive Database and a New Group Contribution Model Based on a Feed-Forward Artificial Neural Network. *J Chem Inf Model* **2014**, *54*, 1311–1324, doi:10.1021/CI500206U.
125. Jeon, H.-N.; Shin, H.K.; Hwang, S.; No, K.T. Development of an Infinite Dilution Activity Coefficient Prediction Model for Organic Solutes in Ionic Liquids with Modified Partial Equalization Orbital Electronegativity Method Derived Descriptors. *ACS Omega* **2021**, *6*, 15361–15373, doi:10.1021/ACSOMEGA.1C01690.
126. David M. Eike; Joan F. Brennecke, and; Maginn, E.J. Predicting Infinite-Dilution Activity Coefficients of Organic Solutes in Ionic Liquids. *Ind Eng Chem Res* **2004**, *43*, 1039–1048, doi:10.1021/IE034152P.
127. Xi, L.; Sun, H.; Li, J.; Liu, H.; Yao, X.; Gramatica, P. Prediction of Infinite-Dilution Activity Coefficients of Organic Solutes in Ionic Liquids Using Temperature-Dependent Quantitative Structure–Property Relationship Method. *Chemical Engineering Journal* **2010**, *163*, 195–201, doi:10.1016/J.CEJ.2010.07.023.
128. Paduszyński, K. In Silico Calculation of Infinite Dilution Activity Coefficients of Molecular Solutes in Ionic Liquids: Critical Review of Current Methods and New Models Based on Three Machine Learning Algorithms. *J Chem Inf Model* **2016**, *56*, 1420–1437, doi:10.1021/ACS.JCIM.6B00166.
129. Gonfa, G.; Bustam, M.A.; Shariff, A.M.; Muhammad, N.; Ullah, S. Quantitative Structure–Activity Relationships (QSARs) for Estimation of Activity Coefficient at Infinite Dilution of Water in Ionic Liquids for Natural Gas Dehydration. *J Taiwan Inst Chem Eng* **2016**, *66*, 222–229, doi:10.1016/J.JTICE.2016.06.027.

130. Mehraein, I.; Riahi, S. The QSPR Models to Predict the Solubility of CO<sub>2</sub> in Ionic Liquids Based on Least-Squares Support Vector Machines and Genetic Algorithm-Multi Linear Regression. *J Mol Liq* **2017**, *225*, 521–530, doi:10.1016/J.MOLLIQ.2016.10.133.
131. Safamirzaei, M.; Modarress, H. Correlating and Predicting Low Pressure Solubility of Gases in [Bmim][BF<sub>4</sub>] by Neural Network Molecular Modeling. *Thermochim Acta* **2012**, *545*, 125–130, doi:10.1016/J.TCA.2012.07.005.
132. Kang, X.; Qian, J.; Deng, J.; Latif, U.; Zhao, Y. Novel Molecular Descriptors for Prediction of H<sub>2</sub>S Solubility in Ionic Liquids. *J Mol Liq* **2018**, *265*, 756–764, doi:10.1016/J.MOLLIQ.2018.06.113.
133. Aghaie, M.; Zendehboudi, S. Estimation of CO<sub>2</sub> Solubility in Ionic Liquids Using Connectionist Tools Based on Thermodynamic and Structural Characteristics. *Fuel* **2020**, *279*, 117984, doi:10.1016/J.FUEL.2020.117984.
134. Jia, Q.; Shi, Q.; Yan, F.; Wang, Q. Norm Index-Based QSPR Model for Describing the n-Octanol/Water Partition Coefficients of Organics. *Environmental Science and Pollution Research* **2020**, *27*, 15454–15462, doi:10.1007/S11356-020-08020-Y.
135. Zhu, T.; Gu, L.; Chen, M.; Sun, F. Exploring QSPR Models for Predicting PUF-Air Partition Coefficients of Organic Compounds with Linear and Nonlinear Approaches. *Chemosphere* **2021**, *266*, 128962, doi:10.1016/J.CHEMOSPHERE.2020.128962.
136. Katritzky, A.R.; Kuanar, M.; Stoyanova-Slavova, I.B.; Slavov, S.H.; Dobchev, D.A.; Karelson, M.; William E. Acree, Jr. Quantitative Structure–Property Relationship Studies on Ostwald Solubility and Partition Coefficients of Organic Solutes in Ionic Liquids. *J Chem Eng Data* **2008**, *53*, 1085–1092, doi:10.1021/IE700607B.
137. Mutelet, F.; Ortega-Villa, V.; Moïse, J.-C.; Jaubert, J.-N.; William E. Acree, Jr. Prediction of Partition Coefficients of Organic Compounds in Ionic Liquids Using a Temperature-Dependent Linear Solvation Energy Relationship with Parameters Calculated through a Group Contribution Method. *J Chem Eng Data* **2011**, *56*, 3598–3606, doi:10.1021/IE200454D.
138. Khooshechin, S.; Dashtbozorgi, Z.; Golmohammadi, H.; Acree, W.E. QSPR Prediction of Gas-to-Ionic Liquid Partition Coefficient of Organic Solutes Dissolved in 1-(2-Hydroxyethyl)-1-Methylimidazolium Tris(Pentafluoroethyl)Trifluorophosphate Using the Replacement Method and Support Vector Regression. *J Mol Liq* **2014**, *196*, 43–51, doi:10.1016/J.MOLLIQ.2014.03.012.
139. Ghaslani, D.; Eshaghi Gorji, Z.; Ebrahimipoor Gorji, A.; Riahi, S. Descriptive and Predictive Models for Henry's Law Constant of CO<sub>2</sub> in Ionic Liquids: A QSPR Study. *Chemical Engineering Research and Design* **2017**, *120*, 15–25, doi:10.1016/j.cherd.2016.12.020.
140. Kang, X.; Liu, C.; Zeng, S.; Zhao, Z.; Qian, J.; Zhao, Y. Prediction of Henry's Law Constant of CO<sub>2</sub> in Ionic Liquids Based on SEP and  $\sigma$ -Profile Molecular Descriptors. *J Mol Liq* **2018**, *262*, 139–147, doi:10.1016/j.molliq.2018.04.026.
141. Safamirzaei, M.; Modarress, H. Application of Neural Network Molecular Modeling for Correlating and Predicting Henry's Law Constants of Gases in [Bmim][PF<sub>6</sub>] at Low Pressures. *Fluid Phase Equilib* **2012**, *332*, 165–172, doi:10.1016/J.FLUID.2012.07.009.
142. Roy, K.; Kar, S.; Das, R.N. Statistical Methods in QSAR/QSPR. In: Springer, Cham, **2015**; pp. 37–59.

143. Zhang, X.; Liu, Z.; Wang, W. Screening of Ionic Liquids to Capture CO<sub>2</sub> by COSMO-RS and Experiments. *AIChE Journal* **2008**, *54*, 2717–2728, doi:10.1002/aic.11573.
144. Gorji, A.E.; Sobati, M.A.; Alopaeus, V.; Uusi-Kyyny, P. Toward Solvent Screening in the Extractive Desulfurization Using Ionic Liquids: QSPR Modeling and Experimental Validations. *Fuel* **2021**, *302*, 121159, doi:10.1016/J.FUEL.2021.121159.
145. Perekhoda, L.A. Assessment of Quantitative Structure-Anticonvulsive Activity Relationships in a Series of Derivatives of 1,2,3-Triazole(1H), 1,2,4-Triazole(4H), 1,3,4-Oxadiazole(1H), and 1,3,4-Thiadiazole(1H). *Pharmaceutical Chemistry Journal* **2014**, *47*, 586–588, doi:10.1007/S11094-014-1012-Z.
146. Venkatraman, V.; Raj, J.J.; Evjen, S.; Lethesh, K.C.; Fiksdahl, A. In Silico Prediction and Experimental Verification of Ionic Liquid Refractive Indices. *J Mol Liq* **2018**, *264*, 563–570, doi:10.1016/J.MOLLIQ.2018.05.067.
147. Atkins Peter; de Paula Julio *Atkins' Physical Chemistry*; Oxford University Press, **2006**;
148. Buchowski H; Ufnalski W *Równowagi Chemiczne*; WNT: Warszawa, **1995**;
149. Buchowski H; Ufnalski W *Roztwory*; WNT: Warszawa, **1995**;
150. Stefanowski Bohdan; Staniszewski Bogumił *Termodynamika Techniczna*; Państwowe Wydawnictwo Naukowe: Warszawa, **1959**;
151. Gerrard, W. Solubility of Gases and Liquids. *Solubility of Gases and Liquids* **1976**, doi:10.1007/978-1-4899-2644-9.
152. Wilhelm, E. Solubility of Gases in Liquids: A Critical Review. *Pure and Applied Chemistry* **1985**, *57*, 303–322, doi:10.1351/PAC198557020303.
153. Pierotti, R.A. The Solubility of Gases in Liquids. *Journal of Physical Chemistry* **2002**, *67*, 1840–1845, doi:10.1021/J100803A024.
154. McMurry, John.; Fay, R.C.. Chemistry. **2001**, 1067.
155. Petrucci, R.H.. General Chemistry : Principles and Modern Applications. **2007**.
156. Simnick, J.J.; Sebastian, H.M.; Lin, H.M.; Chao, K.C. Solubility of Hydrogen in Toluene at Elevated Temperatures and Pressures. *J Chem Eng Data* **2002**, *23*, 339–340, doi:10.1021/JE60079A023.
157. Clarke, C.J.; Tu, W.C.; Levers, O.; Bröhl, A.; Hallett, J.P. Green and Sustainable Solvents in Chemical Processes. *Chem Rev* **2018**, *118*, 747–800, doi:10.1021/ACS.CHEMREV.7B00571.
158. Hong, S.U.; Wang, Y.; Soh, L.S.; Yong, W.F. Are Green Solvents Truly Green? Integrating Life Cycle Assessment and Techno-Economic Analysis for Sustainable Membrane Fabrication. *Green Chemistry* **2023**, *25*, 4501–4512, doi:10.1039/D3GC00776F.
159. Sheldon, R.A. Green Solvents for Sustainable Organic Synthesis: State of the Art. *Green Chemistry* **2005**, *7*, 267–278, doi:10.1039/B418069K.
160. Capello, C.; Fischer, U.; Hungerbühler, K. What Is a Green Solvent? A Comprehensive Framework for the Environmental Assessment of Solvents. *Green Chemistry* **2007**, *9*, 927–934, doi:10.1039/B617536H.

161. Winterton, N. The Green Solvent: A Critical Perspective. *Clean Technol Environ Policy* **2021**, *23*, 2499–2522, doi:10.1007/S10098-021-02188-8/SCHEMES/1.
162. Rao, J.; Tian, X.; Wang, H.; Liu, J.; Li, C. Ionic Liquids as Extractants in Extractive Distillation: A Review of Advances in Screening Methods and Their Mechanisms of Action. *Sep Purif Technol* **2025**, *364*, 132372, doi:10.1016/J.SEPPUR.2025.132372.
163. Liu, F.; Shen, Y.; Shen, L.; Zhang, Y.; Chen, W.; Wang, Q.; Li, S.; Zhang, S.; Li, W. Sustainable Ionic Liquid Organic Solution with Efficient Recyclability and Low Regeneration Energy Consumption for CO<sub>2</sub> Capture. *Sep Purif Technol* **2021**, *275*, 119123, doi:10.1016/J.SEPPUR.2021.119123.
164. Wappel, D.; Gronald, G.; Kalb, R.; Draxler, J. Ionic Liquids for Post-Combustion CO<sub>2</sub> Absorption. *International Journal of Greenhouse Gas Control* **2010**, *4*, 486–494, doi:10.1016/j.ijggc.2009.11.012.
165. Bates, E.D.; Mayton, R.D.; Ntai, I.; Davis, J.H. CO<sub>2</sub> Capture by a Task-Specific Ionic Liquid. *J Am Chem Soc* **2002**, *124*, 926–927, doi:10.1021/ja017593d.
166. Della Posta, S.; Frondaroli, M.C.; Fanali, C. Neoteric Solvents: Applications in Matrix Solid-Phase Dispersion Extraction. *Journal of Chromatography Open* **2024**, *6*, 100150, doi:10.1016/J.JCOA.2024.100150.
167. Azzouz, A.; Hayyan, M. Are Deep Eutectic Solvents Biodegradable? *Process Safety and Environmental Protection* **2023**, *176*, 1021–1025, doi:10.1016/J.PSEP.2023.06.079.
168. Halder, A.K.; Cordeiro, M.N.D.S. Probing the Environmental Toxicity of Deep Eutectic Solvents and Their Components: An In Silico Modeling Approach. *ACS Sustain Chem Eng* **2019**, *7*, 10649–10660, doi:10.1021/ACSSUSCHEMENG.9B01306.
169. Cvjetko Bubalo, M.; Vidović, S.; Radojčić Redovniković, I.; Jokić, S. Green Solvents for Green Technologies. *Journal of Chemical Technology & Biotechnology* **2015**, *90*, 1631–1639, doi:10.1002/JCTB.4668.
170. Płotka-Wasyłka, J.; Rutkowska, M.; Owczarek, K.; Tobiszewski, M.; Namieśnik, J. Extraction with Environmentally Friendly Solvents. *TrAC Trends in Analytical Chemistry* **2017**, *91*, 12–25, doi:10.1016/J.TRAC.2017.03.006.
171. Corsini, E.; Luebke, R.W.; Germolec, D.R.; DeWitt, J.C. PERFLUORINATED COMPOUNDS: EMERGING POPs WITH POTENTIAL IMMUNOTOXICITY. *Toxicol Lett* **2014**, *230*, 263, doi:10.1016/J.TOXLET.2014.01.038.
172. Tavener, S.J.; Clark, J.H. Can Fluorine Chemistry Be Green Chemistry? *J Fluor Chem* **2003**, *123*, 31–36, doi:10.1016/S0022-1139(03)00140-4.
173. Marset, X.; Montilla-Verdú, S.; Rico, E.; Guijarro, N. Beyond Conventional Organic Electrosynthesis: The Role of Fluorinated Solvents. *ACS Electrochemistry* **2024**, *1*, 3–19, doi:10.1021/ACSELECTROCHEM.4C00129.
174. Häckl, K.; Kunz, W. Some Aspects of Green Solvents. *Comptes Rendus Chimie* **2018**, *21*, 572–580, doi:10.1016/J.CRCI.2018.03.010.
175. Parmentier, M.; Gabriel, C.M.; Guo, P.; Isley, N.A.; Zhou, J.; Gallou, F. Switching from Organic Solvents to Water at an Industrial Scale. *Curr Opin Green Sustain Chem* **2017**, *7*, 13–17, doi:10.1016/J.COGSC.2017.06.004.

176. Tobiszewski, M.; Namieśnik, J.; Pena-Pereira, F. Environmental Risk-Based Ranking of Solvents Using the Combination of a Multimedia Model and Multi-Criteria Decision Analysis. *Green Chemistry* **2017**, *19*, 1034–1042, doi:10.1039/C6GC03424A.
177. Gills, B.; Morgan, J. Global Climate Emergency: After COP24, Climate Science, Urgency, and the Threat to Humanity. *Globalizations* **2020**, *17*, 885–902.
178. Figueroa, J.D.; Fout, T.; Plasynski, S.; McIlvried, H.; Srivastava, R.D. Advances in CO<sub>2</sub> Capture Technology-The U.S. Department of Energy's Carbon Sequestration Program. *International Journal of Greenhouse Gas Control* **2008**, *2*, 9–20.
179. Keith, D.W. Why Capture CO<sub>2</sub> from the Atmosphere? *Science* **2009**, *325*, 1654–1655.
180. de Coninck, H.; Davidson, O.; Metz, B.; Loos, M.; Meyer, L. *IPCC Special Report on Carbon Dioxide Capture and Storage*; **2005**;
181. WMO WMO Greenhouse Gas Bulletin - The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2014. In *World Meteorological Organization and Global Atmosphere Watch*; **2015**; pp. 1–4 ISBN 2078-0796.
182. Simon, N.M.; Zanatta, M.; dos Santos, F.P.; Corvo, M.C.; Cabrita, E.J.; Dupont, J. Carbon Dioxide Capture by Aqueous Ionic Liquid Solutions. *ChemSusChem* **2017**, *10*, 4927–4933, doi:10.1002/cssc.201701044.
183. Rochelle, G.T. Amine Scrubbing for CO<sub>2</sub> Capture. *Science* **2009**, *325*, 1652–1654.
184. Dutcher, B.; Fan, M.; Russell, A.G. Amine-Based CO<sub>2</sub> Capture Technology Development from the Beginning of 2013-A Review. *ACS Appl Mater Interfaces* **2015**, *7*, 2137–2148, doi:10.1021/am507465f.
185. Anthony, J.L.; Aki, S.N.V.K.; Maginn, E.J.; Brennecke, J.F. Feasibility of Using Ionic Liquids for Carbon Dioxide Capture. *International Journal of Environmental Technology and Management* **2004**, *4*, 105–115.
186. Sheridan, Q.R.; Schneider, W.F.; Maginn, E.J. Role of Molecular Modeling in the Development of CO<sub>2</sub>-Reactive Ionic Liquids. *Chem Rev* **2018**, *118*, 5242–5260, doi:10.1021/acs.chemrev.8b00017.
187. Shannon, M.S.; Tedstone, J.M.; Danielsen, S.P.O.; Hindman, M.S.; Irvin, A.C.; Bara, J.E. Free Volume as the Basis of Gas Solubility and Selectivity in Imidazolium-Based Ionic Liquids. *Ind Eng Chem Res* **2012**, *51*, 5565–5576, doi:10.1021/IE202916E/SUPPL\_FILE/IE202916E\_SI\_001.PDF.
188. Dai, C.; Lei, Z.; Chen, B. Gas Solubility in Long-Chain Imidazolium-Based Ionic Liquids. *AIChE Journal* **2017**, *63*, 1792–1798, doi:10.1002/aic.15711.
189. Wenny, M.B.; Molinari, N.; Slavney, A.H.; Thapa, S.; Lee, B.; Kozinsky, B.; Mason, J.A. Understanding Relationships between Free Volume and Oxygen Absorption in Ionic Liquids. *Journal of Physical Chemistry B* **2022**, *126*, 1268–1274, doi:10.1021/acs.jpcc.2c00202.
190. Horne, J.W.; Shannon, M.S.; Bara, J.E. Correlating Fractional Free Volume to CO<sub>2</sub> Selectivity in [Rmim][Tf<sub>2</sub>N] Ionic Liquids. *Journal of Chemical Thermodynamics* **2014**, *77*, 190–196, doi:10.1016/j.jct.2014.03.012.

191. Dong, Q.; Muzny, C.D.; Kazakov, A.; Diky, V.; Magee, J.W.; Widegren, J.A.; Chirico, R.D.; Marsh, K.N.; Frenkel, M. Reviews ILThermo: A Free-Access Web Database for Thermodynamic Properties of Ionic Liquids †. **2007**, *52*, 1151–1159, doi:10.1021/je700171f.
192. Kazakov, A.F.; Magee, J.W.; Chirico, R.D.; Diky, V.; Kroenlein, K.G.; Muzny, C.D.; Frenkel, M.D. Ionic Liquids Database - ILThermo (v2.0). (dostęp 2 kwietnia 2022)
193. Soriano, A.N.; Doma, B.T.; Li, M.H. Carbon Dioxide Solubility in Some Ionic Liquids at Moderate Pressures. *J Taiwan Inst Chem Eng* **2009**, *40*, 387–393, doi:10.1016/j.jtice.2008.12.002.
194. Muldoon, M.J.; Aki, S.N.V.K.; Anderson, J.L.; Dixon, J.K.; Brennecke, J.F. Improving Carbon Dioxide Solubility in Ionic Liquids. *Journal of Physical Chemistry B* **2007**, *111*, 9001–9009, doi:10.1021/jp071897q.
195. Ramdin, M.; Olasagasti, T.Z.; Vlucht, T.J.H.; De Loos, T.W. High Pressure Solubility of CO<sub>2</sub> in Non-Fluorinated Phosphonium-Based Ionic Liquids. *Journal of Supercritical Fluids* **2013**, *82*, 41–49, doi:10.1016/j.supflu.2013.06.004.
196. Carvalho, P.J.; Álvarez, V.H.; Machado, J.J.B.; Pauly, J.; Daridon, J.L.; Marrucho, I.M.; Aznar, M.; Coutinho, J.A.P. High Pressure Phase Behavior of Carbon Dioxide in 1-Alkyl-3-Methylimidazolium Bis(Trifluoromethylsulfonyl)Imide Ionic Liquids. *Journal of Supercritical Fluids* **2009**, *48*, 99–107, doi:10.1016/j.supflu.2008.10.012.
197. Camper, D.; Bara, J.; Koval, C.; Noble, R. Bulk-Fluid Solubility and Membrane Feasibility of Rmim-Based Room-Temperature Ionic Liquids. *Ind Eng Chem Res* **2006**, *45*, 6279–6283, doi:10.1021/ie060177n.
198. Baltus, R.E.; Culbertson, B.H.; Dai, S.; Luo, H.; DePaoli, D.W. Low-Pressure Solubility of Carbon Dioxide in Room-Temperature Ionic Liquids Measured with a Quartz Crystal Microbalance. *J Phys Chem B* **2003**, *108*, 721–727, doi:10.1021/jp036051a.
199. Camper, D.; Scovazzo, P.; Koval, C.; Noble, R. Gas Solubilities in Room-Temperature Ionic Liquids. *Ind Eng Chem Res* **2004**, *43*, 3049–3054, doi:10.1021/ie034097k.
200. Anderson, J.L.; Dixon, J.K.; Maginn, E.J.; Brennecke, J.F. Measurement of SO<sub>2</sub> Solubility in Ionic Liquids. *J Phys Chem B* **2006**, *110*, 15059–15062, doi:10.1021/jp063547u.
201. Hou, Y.; Baltus, R.E. Experimental Measurement of the Solubility and Diffusivity of CO<sub>2</sub> in Room-Temperature Ionic Liquids Using a Transient Thin-Liquid-Film Method. *Ind Eng Chem Res* **2007**, *46*, 8166–8175, doi:10.1021/ie070501u.
202. Althuluth, M.; T. Mota-Martinez, M.; C. Kroon, M.; J. Peters, C. Solubility of Carbon Dioxide in the Ionic Liquid 1-Ethyl-3-Methylimidazolium Tris(Pentafluoroethyl)Trifluorophosphate. *Journal of Chemical & Engineering Data* **2012**, *57*, 3422–3425, doi:10.1021/je300521y.
203. Ferguson, L.; Scovazzo, P. Solubility, Diffusivity, and Permeability of Gases in Phosphonium-Based Room Temperature Ionic Liquids: Data and Correlations. *Ind Eng Chem Res* **2007**, *46*, 1369–1374, doi:10.1021/ie0610905.
204. Mahurin, S.M.; Yeary, J.S.; Baker, S.N.; Jiang, D. en; Dai, S.; Baker, G.A. Ring-Opened Heterocycles: Promising Ionic Liquids for Gas Separation and Capture. *J Memb Sci* **2012**, *401–402*, 61–67, doi:10.1016/j.memsci.2012.01.042.

205. Moya, C.; Gonzalez-Miquel, M.; Rodriguez, F.; Soto, A.; Rodriguez, H.; Palomar, J. Non-ideal behavior of ionic liquid mixtures to enhance CO<sub>2</sub> capture. *Fluid Phase Equilib* **2017**, *450*, 175–183, doi:10.1016/j.fluid.2017.07.014.
206. Carvalho, P.J.; Álvarez, V.H.; Schröder, B.; Gil, A.M.; Marrucho, I.M.; Aznar, M.; Santos, L.M.N.B.F.; Coutinho, J.A.P. Specific Solvation Interactions of CO<sub>2</sub> on Acetate and Trifluoroacetate Imidazolium Based Ionic Liquids at High Pressures. *Journal of Physical Chemistry B* **2009**, *113*, 6803–6812, doi:10.1021/jp901275b.
207. Mahurin, S.M.; Dai, T.; Yeary, J.S.; Luo, H.; Dai, S. Benzyl-Functionalized Room Temperature Ionic Liquids for CO<sub>2</sub>/N<sub>2</sub> Separation. *Ind Eng Chem Res* **2011**, *50*, 14061–14069, doi:10.1021/ie201428k.
208. Zhang, J.; Zhang, Q.; Qiao, B.; Deng, Y. Solubilities of the Gaseous and Liquid Solutes and Their Thermodynamics of Solubilization in the Novel Room-Temperature Ionic Liquids at Infinite Dilution by Gas Chromatography. *J Chem Eng Data* **2007**, *52*, 2277–2283, doi:10.1021/je700297c.
209. Jalili, A.H.; Mehdizadeh, A.; Shokouhi, M.; Sakhaeinia, H.; Taghikhani, V. Solubility of CO<sub>2</sub> in 1-(2-Hydroxyethyl)-3-Methylimidazolium Ionic Liquids with Different Anions. *Journal of Chemical Thermodynamics* **2010**, *42*, 787–791, doi:10.1016/j.jct.2010.02.002.
210. Galán Sánchez, L.M.; Meindersma, G.W.; de Haan, A.B. Solvent Properties of Functionalized Ionic Liquids for CO<sub>2</sub> Absorption. *Chemical Engineering Research and Design* **2007**, *85*, 31–39, doi:10.1205/cherd06124.
211. Carvalho, P.J.; Álvarez, V.H.; Marrucho, I.M.; Aznar, M.; Coutinho, J.A.P. High Pressure Phase Behavior of Carbon Dioxide in 1-Butyl-3-Methylimidazolium Bis(Trifluoromethylsulfonyl)Imide and 1-Butyl-3-Methylimidazolium Dicyanamide Ionic Liquids. *Journal of Supercritical Fluids* **2009**, *50*, 105–111, doi:10.1016/j.supflu.2009.05.008.
212. Cadena, C.; Anthony, J.L.; Shah, J.K.; Morrow, T.I.; Brennecke, J.F.; Maginn, E.J. Why Is CO<sub>2</sub> So Soluble in Imidazolium-Based Ionic Liquids? *J Am Chem Soc* **2004**, *126*, 5300–5308, doi:10.1021/ja039615x.
213. Finotello, A.; Bara, J.E.; Camper, D.; Noble, R.D. Room-Temperature Ionic Liquids: Temperature Dependence of Gas Solubility Selectivity. *Ind Eng Chem Res* **2008**, *47*, 3453–3459, doi:10.1021/ie0704142.
214. Yuan, X.; Zhang, S.; Liu, J.; Lu, X. Solubilities of CO<sub>2</sub> in Hydroxyl Ammonium Ionic Liquids at Elevated Pressures. *Fluid Phase Equilib* **2007**, *257*, 195–200, doi:10.1016/j.fluid.2007.01.031.
215. Kurnia, K.A.; Harris, F.; Wilfred, C.D.; Abdul Mutalib, M.I.; Murugesan, T. Thermodynamic Properties of CO<sub>2</sub> Absorption in Hydroxyl Ammonium Ionic Liquids at Pressures of (100–1600) KPa. *Journal of Chemical Thermodynamics* **2009**, *41*, 1069–1073, doi:10.1016/j.jct.2009.04.003.
216. Mattedi, S.; Carvalho, P.J.; Coutinho, J.A.P.; Alvarez, V.H.; Iglesias, M. High Pressure CO<sub>2</sub> Solubility in N-Methyl-2-Hydroxyethylammonium Protic Ionic Liquids. In *Proceedings of the Journal of Supercritical Fluids; Elsevier*, **2011**; Vol. 56, pp. 224–230.

217. Comesana, A.E.; Huntington, T.T.; Scown, C.D.; Niemeyer, K.E.; Rapp, V.H. A Systematic Method for Selecting Molecular Descriptors as Features When Training Models for Predicting Physiochemical Properties. *Fuel* **2022**, *321*, 123836, doi:10.1016/J.FUEL.2022.123836.
218. Hanwell, M.D.; Curtis, D.E.; Lonie, D.C.; Vandermeersch, T.; Zurek, E.; Hutchison, G.R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J Cheminform* **2012**, *4*, doi:10.1186/1758-2946-4-17.
219. Halgren, T.A. MMFF VII. Characterization of MMFF94, MMFF94s, and Other Widely Available Force Fields for Conformational Energies and for Intermolecular-Interaction Energies and Geometries. *J Comput Chem* **1999**, *20*, 730–748, doi:10.1002/(SICI)1096-987X(199905)20:7<730::AID-JCC8>3.0.CO;2-T.
220. Frisch, M.J.; Trucks, G.W.; Schlegel, H.E.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; O., Farkas.; et al. Gaussian 16. *Gaussian, Inc., Wallingford CT*, 2016.
221. Hypercube Inc. HyperChem(TM) Professional.
222. Talete Dragon7, Molecular Descriptors for Computational Chemistry. Talete Srl **2015**.
223. Sukumar, N.; Prabhu, G.; Saha, P. Applications of Genetic Algorithms in QSAR/QSPR Modeling. In *Applications of Metaheuristics in Process Engineering*; Springer International Publishing, **2014**; Vol. 9783319065083, pp. 315–324 ISBN 9783319065083.
224. Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*; 1st ed.; Addison-Wesley Longman Publishing Co., Inc.: USA, **1989**; ISBN 0201157675.
225. Beichel, W.; Yu, Y.; Dlubek, G.; Krause-Rehberg, R.; Pionteck, J.; Pfeifferkorn, D.; Bulut, S.; Bejan, D.; Friedrich, C.; Krossing, I. Free Volume in Ionic Liquids: A Connection of Experimentally Accessible Observables from PALS and PVT Experiments with the Molecular Structure from XRD Data. *Physical Chemistry Chemical Physics* **2013**, *15*, 8821–8830, doi:10.1039/c3cp43306d.
226. Low, K.; Tan, S.Y.S.; Izgorodina, E.I. An Ab Initio Study of the Structure and Energetics of Hydrogen Bonding in Ionic Liquids. *Front Chem* **2019**, *7*, 433455, doi:10.3389/FCHEM.2019.00208/BIBTEX.
227. Chen, S.; Izgorodina, E.I. Prediction of <sup>1</sup>H NMR Chemical Shifts for Clusters of Imidazolium-Based Ionic Liquids. *Physical Chemistry Chemical Physics* **2017**, *19*, 17411–17425, doi:10.1039/C7CP02951A.
228. Roy, K. On Some Aspects of Validation of Predictive Quantitative Structure-Activity Relationship Models. *Expert Opin Drug Discov* **2007**, *2*, 1567–1577.
229. Martin, T.M.; Harten, P.; Young, D.M.; Muratov, E.N.; Golbraikh, A.; Zhu, H.; Tropsha, A. Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling? *J Chem Inf Model* **2012**, *52*, 2570–2578, doi:10.1021/ci300338w.
230. Leonard, J.T.; Roy, K. On Selection of Training and Test Sets for the Development of Predictive QSAR Models. *QSAR Comb Sci* **2006**, *25*, 235–251, doi:10.1002/qsar.200510161.
231. Kennard, R.W.; Stone, L.A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137–148, doi:10.1080/00401706.1969.10490666.

232. Roy Kunal [https://Teqip.Jdvu.Ac.in/QSAR\\_Tools/](https://Teqip.Jdvu.Ac.in/QSAR_Tools/). (dostęp 20 sierpnia 2020)
233. Yee, L.C.; Wei, Y.C. Current Modeling Methods Used in QSAR/QSPR. In *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*; John Wiley & Sons, Ltd, **2012**; pp. 1–31 ISBN 9783527645121.
234. Wold, S.; Sjöström, M.; Eriksson, L. PLS-Regression: A Basic Tool of Chemometrics. In *Proceedings of the Chemometrics and Intelligent Laboratory Systems*; Elsevier, **2001**; Vol. 58, pp. 109–130.
235. Alexander, D.L.J.; Tropsha, A.; Winkler, D.A. Beware of R<sup>2</sup>: Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *J Chem Inf Model* **2015**, *55*, 1316, doi:10.1021/ACS.JCIM.5B00206.
236. Hodson, T.O. Root-Mean-Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not. *Geosci Model Dev* **2022**, *15*, 5481–5487, doi:10.5194/GMD-15-5481-2022.
237. Spiess, A.N.; Neumeyer, N. An Evaluation of R<sup>2</sup> as an Inadequate Measure for Nonlinear Models in Pharmacological and Biochemical Research: A Monte Carlo Approach. *BMC Pharmacol* **2010**, *10*, 1–11, doi:10.1186/1471-2210-10-6/FIGURES/4.
238. Piepho, H.P. A Coefficient of Determination (R<sup>2</sup>) for Generalized Linear Mixed Models. *Biometrical Journal* **2019**, *61*, 860–872, doi:10.1002/BIMJ.201800270.
239. Rücker, C.; Rücker, G.; Meringer, M. Y-Randomization and Its Variants in QSPR/QSAR. *J Chem Inf Model* **2007**, *47*, 2345–2357, doi:10.1021/CI700157B.
240. Gajewicz, A. How to Judge Whether QSAR/Read-across Predictions Can Be Trusted: A Novel Approach for Establishing a Model's Applicability Domain. *Environ Sci Nano* **2018**, *5*, 408–421, doi:10.1039/C7EN00774D.
241. Kar, S.; Roy, K.; Leszczynski, J. Applicability Domain: A Step Toward Confident Predictions and Decidability for QSAR Modeling. *Methods in Molecular Biology* **2018**, *1800*, 141–169, doi:10.1007/978-1-4939-7899-1\_6.
242. Sahigara, F.; Mansouri, K.; Ballabio, D.; Mauri, A.; Consonni, V.; Todeschini, R. Comparison of Different Approaches to Define the Applicability Domain of QSAR Models. *Molecules* **2012**, *17*, 4791–4810, doi:10.3390/MOLECULES17054791.
243. Kier, L.B.; Hall, L.H. An Electrotopological-State Index for Atoms in Molecules. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists* **1990**, *7*, 801–807, doi:10.1023/A:1015952613760/METRICS.
244. Todeschini, R.; Consonni, V.; Ballabio, D.; Mauri, A.; Cassotti, M.; Lee, S.; West, A.; Cartledge, D. QSPR STUDY OF RHEOLOGICAL AND MECHANICAL PROPERTIES OF CHLOROPRENE RUBBER ACCELERATORS. *Rubber Chemistry and Technology* **2014**, *87*, 219–238, doi:10.5254/rct.13.87918.
245. Ling, Y. Characterization and Evaluation of Cyclodextrin Polymers as Alternative Adsorbents to Activated Carbon for Micropollutant Removal during Water and Wastewater Treatment, Cornell University, **2019**.

246. Consonni, V.; Todeschini, R.; Pavan, M. Structure/Response Correlations and Similarity/Diversity Analysis by GETAWAY Descriptors. 1. Theory of the Novel 3D Molecular Descriptors. *J Chem Inf Comput Sci* **2002**, *42*, 682–692, doi:10.1021/ci015504a.
247. Liu, X.; O’Harra, K.E.; Bara, J.E.; Turner, C.H. Molecular Insight into the Anion Effect and Free Volume Effect of CO<sub>2</sub> solubility in Multivalent Ionic Liquids. *Physical Chemistry Chemical Physics* **2020**, *22*, 20618–20633, doi:10.1039/d0cp03424j.
248. Politzer, P.; Murray, J.S.; Concha, M.C. The Complementary Roles of Molecular Surface Electrostatic Potentials and Average Local Ionization Energies with Respect to Electrophilic Processes. *Int J Quantum Chem* **2002**, *88*, 19–27, doi:10.1002/qua.10109.
249. Chen, G.; Zheng, S.; Luo, X.; Shen, J.; Zhu, W.; Liu, H.; Gui, C.; Zhang, J.; Zheng, M.; Chum, M.P.; et al. Focused Combinatorial Library Design Based on Structural Diversity, Druglikeness and Binding Affinity Score. *J Comb Chem* **2005**, *7*, 398–406, doi:10.1021/CC049866H.

## 8. DOROBEK NAUKOWY

1. Marcinkowski, Ł.; **Eichenlaub, J.**; Ghasemi, E.; Polkowska, Ż.; Kloskowski, A. Measurements of Activity Coefficients at Infinite Dilution for Organic Solutes in the Ionic Liquids *N*-Ethyl- and *N*-Octyl-*N*-methylmorpholinium Bis(trifluoromethanesulfonyl)imide. A Useful Tool for Solvent Selection. *Molecules* **2020**, 25, 634. <https://doi.org/10.3390/molecules25030634>

Punkty Ministerialne: 140

Impact Factor (2025): 4,6

2. **Eichenlaub, J.**; Rakowska, P.W.; Kloskowski, A. User-assisted methodology targeted for building structure interpretable QSPR models for boosting CO<sub>2</sub> capture with ionic liquids. *Journal of Molecular Liquids* **2022**, 350, 118511. <https://10.1016/j.molliq.2022.118511>

Punkty Ministerialne: 100

Impact Factor (2025): 5,2

3. **Eichenlaub, J.**; Baran, K.; Śmiechowski, M.; Kloskowski, A. Free volume in physical absorption of carbon dioxide in ionic liquids: Molecular dynamics supported modeling. *Separation and Purification Technology* **2023**, 313, 123464. <https://doi.org/10.1016/j.seppur.2023.123464>

Punkty Ministerialne: 140

Impact Factor (2025): 9,0

4. **Eichenlaub, J.**; Baran, K.; Urbański, K.; Robakowska, M.; Kalinowska, J.; Racka-Pilszak, B.; Kloskowski, A. In Search of the Perfect Composite Material—A Chemoinformatics Approach Towards the Easier Handling of Dental Materials. *International Journal of Molecular Sciences*. **2025**, 26, 8283. <https://doi.org/10.3390/ijms26178283>

Punkty Ministerialne: 140

Impact Factor (2025): 4,9

Publikacja przygotowana we współpracy z Gdańskim Uniwersytetem Medycznym

## SPIS ILUSTRACJI I TABEL

|                                                                                                                                                                                                       |           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Rysunek 1. 1.</b> Wybrane wydarzenia w historii badań nad cieczami jonowymi .....                                                                                                                  | <b>11</b> |
| <b>Rysunek 1.2.</b> Liczba publikacji dotyczących cieczy jonowych indeksowanych w bazie SCOPUS na przestrzeni lat .....                                                                               | <b>12</b> |
| <b>Rysunek 1.3.</b> Wybrane wzory kationów i anionów tworzących ciecze jonowe.....                                                                                                                    | <b>14</b> |
| <b>Rysunek 1.4.</b> Schemat procedury modelowania QSPR.....                                                                                                                                           | <b>22</b> |
| <b>Rysunek 3.1.</b> Schemat postępowania w przypadku wystąpienia skorelowanych ze sobą deskryptorów .....                                                                                             | <b>50</b> |
| <b>Rysunek 3.2.</b> Graficzna reprezentacja procesu przygotowania danych oraz procesu modelowania ....                                                                                                | <b>51</b> |
| <b>Rysunek 4.1.</b> Wykresy zależności wartości przewidywanej $H_{CO_2}$ od wartości eksperymentalnej dla trzech modeli. a) MLR, b) LR, c) PLS .....                                                  | <b>59</b> |
| <b>Rysunek 4.2.</b> Wykresy Williamsa dla modeli a) MLR, b) LR, c) PLS .....                                                                                                                          | <b>60</b> |
| <b>Rysunek 4.3.</b> Graficzna reprezentacja struktur jonów wraz z odpowiadającymi im wartościami analizowanych deskryptorów .....                                                                     | <b>62</b> |
| <b>Rysunek 4.4.</b> Luki w cieczach o niskim, średnim i wysokim FFV. a) [THE][OAc], b) [HMIM][Cl], c) [P66614][N(CN) <sub>2</sub> ].....                                                              | <b>67</b> |
| <b>Rysunek 4.5.</b> Porównanie liczby luk o różnych rozmiarach w cieczy [P66614][DCA]. a) 200 pm, b) 170 pm, c) 140 pm.....                                                                           | <b>67</b> |
| <b>Rysunek 4.6.</b> a) rozkład wartości FFV, b) rozkład wartości FFV w odniesieniu do $H_{CO_2}$ .....                                                                                                | <b>68</b> |
| <b>Rysunek 4.7.</b> Wykresy zależności wartości przewidywanej FFV od wartości eksperymentalnej dla modelu FFV jako zmienna zależna. ....                                                              | <b>72</b> |
| <b>Rysunek 4.8.</b> Wykresy zależności wartości przewidywanej $H_{CO_2}$ od wartości eksperymentalnej dla modelu standardowego. ....                                                                  | <b>73</b> |
| <b>Rysunek 4.9.</b> Wykresy zależności wartości przewidywanej $H_{CO_2}$ od wartości eksperymentalnej dla modelu FFV. ....                                                                            | <b>74</b> |
| <b>Rysunek 4.10.</b> Wizualne porównanie parametrów statystycznych otrzymanych modeli. ....                                                                                                           | <b>79</b> |
| <b>Rysunek 4.11.</b> Wykresy zależności wartości przewidywanej $H_{CO_2}$ od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych osobno w próżni lub w etanolu różnymi metodami..... | <b>83</b> |
| <b>Rysunek 4.12.</b> Wykresy zależności wartości przewidywanej $H_{CO_2}$ od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych razem w próżni lub w etanolu różnymi metodami.....  | <b>84</b> |
| <b>Rysunek 4.13.</b> Wykresy zależności wartości przewidywanej FFV od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych osobno w próżni różnymi metodami.....                      | <b>85</b> |
| <b>Rysunek 4.14.</b> Wykresy zależności wartości przewidywanej FFV od wartości eksperymentalnej dla modeli zoptymalizowanych i zbudowanych razem w próżni lub etanolu różnymi metodami .....          | <b>86</b> |
| <b>Rysunek 4.15.</b> Porównanie parametrów statystycznych modeli $H_{CO_2}$ : a) MLR, b) PLS oraz c) SVM w zależności od sposobu i środowiska optymalizacji cieczy jonowych.....                      | <b>87</b> |
| <b>Rysunek 4.16.</b> Wykresy radarowe porównujące parametry statystyczne omawianych modeli $H_{CO_2}$ ....                                                                                            | <b>88</b> |
| <b>Rysunek 4.17.</b> Porównanie parametrów statystycznych modeli FFV: a) MLR, b) PLS oraz c) LR w zależności od sposobu i środowiska optymalizacji cieczy jonowych.....                               | <b>89</b> |
| <b>Rysunek 4.18.</b> Wykresy radarowe porównujące parametry statystyczne omawianych modeli FFV .....                                                                                                  | <b>90</b> |
| <b>Rysunek 4.19.</b> Porównanie parametrów statystycznych modeli $H_{CO_2}$ : a) razem, b) osobno w zależności od środowiska optymalizacji cieczy jonowych oraz metody modelowania. ....              | <b>90</b> |
| <b>Rysunek 4.20.</b> Porównanie parametrów statystycznych modeli FFV: a) razem, b) osobno w zależności od środowiska optymalizacji cieczy jonowych oraz metody modelowania .....                      | <b>91</b> |
| <b>Rysunek 4.21.</b> Porównanie parametrów statystycznych modeli $H_{CO_2}$ : a) w próżni, b) w etanolu w zależności od sposobu optymalizacji cieczy jonowych oraz metody modelowania .....           | <b>92</b> |

|                                                                                                                                                                                                                                                                  |            |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Rysunek 4.22.</b> Porównanie parametrów statystycznych modeli FFV: a) razem, b) osobno w zależności od sposobu optymalizacji cieczy jonowych oraz metody modelowania .....                                                                                    | <b>93</b>  |
| <b>Rysunek 4.23.</b> RMSE obliczone dla modeli H <sub>CO2</sub> uszeregowane rosnąco .....                                                                                                                                                                       | <b>94</b>  |
| <b>Rysunek 4.24.</b> RMSE obliczone dla modeli FFV uszeregowane rosnąco .....                                                                                                                                                                                    | <b>94</b>  |
| <b>Rysunek 4.25.</b> Histogram różnic procentowych wartości deskryptorów pomiędzy anionem [NTf <sub>2</sub> ] zoptymalizowanym osobno w próżni oraz anionem [NTf <sub>2</sub> ] zoptymalizowanym w próżni w obecności jonu [BMIM] .....                          | <b>96</b>  |
| <b>Rysunek 4.26.</b> Histogram różnic procentowych wartości deskryptorów pomiędzy anionem [NTf <sub>2</sub> ] zoptymalizowanym osobno w próżni oraz anionem [NTf <sub>2</sub> ] zoptymalizowanym w próżni w obecności jonu [B3MPy] .....                         | <b>96</b>  |
| <b>Rysunek 4.27.</b> Wspólny histogram różnic procentowych wartości deskryptorów pomiędzy anionem [NTf <sub>2</sub> ] zoptymalizowanym osobno w próżni oraz anionem [NTf <sub>2</sub> ] zoptymalizowanym w próżni w obecności kationów [BMIM] oraz [B3MPy] ..... | <b>97</b>  |
| <b>Rysunek 4.28.</b> Histogram różnic procentowych wartości deskryptorów pomiędzy jonem [NTf <sub>2</sub> ] cieczy [B3MPy][NTf <sub>2</sub> ] zoptymalizowanej razem w próżni oraz razem w środowisku etanolu .....                                              | <b>98</b>  |
| <b>Rysunek 4.29.</b> Histogram różnic procentowych wartości deskryptorów pomiędzy jonem [NTf <sub>2</sub> ] cieczy [BMIM][NTf <sub>2</sub> ] zoptymalizowanej razem w próżni oraz razem w środowisku etanolu .....                                               | <b>98</b>  |
| <b>Rysunek 4.30.</b> Histogram różnic procentowych wartości deskryptorów pomiędzy jonem [NTf <sub>2</sub> ] cieczy [PMIM][NTf <sub>2</sub> ] zoptymalizowanej razem w próżni oraz razem w środowisku etanolu .....                                               | <b>99</b>  |
| <b>Rysunek 4.31.</b> 15 deskryptorów o największej zmienności dla każdej z metod .....                                                                                                                                                                           | <b>100</b> |
| <b>Rysunek 4.32.</b> Ranking zmienności deskryptorów dla RSD obliczonego przy optymalizacji razem w próżni .....                                                                                                                                                 | <b>101</b> |
| <b>Rysunek 4.33.</b> Ranking zmienności deskryptorów dla RSD obliczonego przy optymalizacji razem w etanolu .....                                                                                                                                                | <b>102</b> |
| <b>Rysunek 4.34.</b> Porównanie różnic w położeniu w rankingach lokalnych dla RSD obliczonych dla optymalizacji razem w próżni i razem w etanolu .....                                                                                                           | <b>102</b> |

|                                                                                                                                                                                                                                                              |           |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <b>Tabela 1.1.</b> Wybrane publikacje dotyczące modelowania QSPR w celu przewidywania właściwości fizykochemicznych cieczy jonowych .....                                                                                                                    | <b>28</b> |
| <b>Tabela 3.1.</b> Zestawienie cieczy jonowych wykorzystanych w badaniach, wraz z ich zmierzonymi eksperymentalnie wartościami stałej Henry'ego dla ditlenku węgla, źródłami literaturowymi oraz z obliczonymi wartościami frakcyjnej wolnej objętości ..... | <b>43</b> |
| <b>Tabela 4.1.</b> Zestawienie definicji analizowanych deskryptorów wraz z kategoriami, do których należą .....                                                                                                                                              | <b>56</b> |
| <b>Tabela 4.2.</b> Współczynniki regresji deskryptorów w analizowanych modelach .....                                                                                                                                                                        | <b>56</b> |
| <b>Tabela 4.3.</b> Parametry statystyczne modeli .....                                                                                                                                                                                                       | <b>57</b> |
| <b>Tabela 4.4.</b> Zestawienie parametrów istotności statystycznej deskryptorów .....                                                                                                                                                                        | <b>61</b> |
| <b>Tabela 4.5.</b> Deskryptory w poszczególnych modelach wraz z ich definicjami, kategorią, do której należą oraz wartością parametru t .....                                                                                                                | <b>70</b> |
| <b>Tabela 4.6.</b> Współczynniki regresji deskryptorów w modelach .....                                                                                                                                                                                      | <b>70</b> |
| <b>Tabela 4.7.</b> Porównanie parametrów statystycznych modeli .....                                                                                                                                                                                         | <b>75</b> |
| <b>Tabela 4.8.</b> Parametry statystyczne omawianych modeli .....                                                                                                                                                                                            | <b>80</b> |

## MATERIAŁY UZUPEŁNIAJĄCE

**Tabela A1.** Liczba deskryptorów z poszczególnych kategorii pozostałych po każdym z czterech etapów algorytmu genetycznego

| Kat. | Opis kategorii (ang/pl)                                                                                                                                                     | Kationy |     |     |    | Aniony |     |     |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|-----|-----|----|--------|-----|-----|----|
|      |                                                                                                                                                                             | I       | II  | III | IV | I      | II  | III | IV |
| 1    | Constitutional indices/Indeksy strukturalne                                                                                                                                 | 41      | 23  | 11  | 2  | 42     | 26  | 10  | 1  |
| 2    | Ring descriptors/Deskryptory pierścieniowe                                                                                                                                  | 15      | 6   | 5   | 1  | 0      | 0   | 0   | 0  |
| 3    | Topological indices/Indeksy topologiczne                                                                                                                                    | 76      | 25  | 13  | 2  | 73     | 20  | 10  | 2  |
| 4    | Walk and path counts/Ścieżki i drogi                                                                                                                                        | 46      | 12  | 2   | 2  | 43     | 8   | 6   | 2  |
| 5    | Connectivity indices/Indeksy wiązań                                                                                                                                         | 38      | 12  | 8   | 2  | 38     | 15  | 9   | 2  |
| 6    | Information indices/Indeksy informacyjne                                                                                                                                    | 51      | 24  | 10  | 3  | 51     | 26  | 8   | 2  |
| 7    | 2D matrix-based descriptors/Deskryptory na bazie matryc 2D                                                                                                                  | 596     | 82  | 20  | 3  | 589    | 54  | 10  | 3  |
| 8    | 2D autocorrelations/Autokorelacje 2D                                                                                                                                        | 214     | 105 | 24  | 2  | 214    | 64  | 10  | 2  |
| 9    | Burden eigenvalues/Własności własne Burdena                                                                                                                                 | 97      | 29  | 14  | 2  | 97     | 31  | 8   | 2  |
| 10   | P_VSA-like descriptors/Deskryptory P_VSA                                                                                                                                    | 51      | 26  | 12  | 2  | 49     | 36  | 9   | 2  |
| 11   | ETA indices/Rozszerzone atomowe indeksy topochemiczne                                                                                                                       | 24      | 16  | 8   | 2  | 24     | 15  | 8   | 3  |
| 12   | Edge adjacency indices/Indeksy macierzy sąsiedztwa                                                                                                                          | 323     | 69  | 23  | 2  | 324    | 46  | 10  | 4  |
| 13   | Geometrical descriptors/Deskryptory geometryczne                                                                                                                            | 39      | 17  | 14  | 4  | 39     | 16  | 7   | 2  |
| 14   | 3D matrix-based descriptors/Deskryptory na bazie matryc 3D                                                                                                                  | 100     | 22  | 12  | 11 | 100    | 24  | 8   | 4  |
| 15   | 3D autocorrelations/Autokorelacje 3D                                                                                                                                        | 81      | 40  | 15  | 3  | 81     | 23  | 10  | 2  |
| 16   | RDF descriptors/Deskryptory funkcji rozkładu radialnego                                                                                                                     | 204     | 25  | 11  | 3  | 210    | 76  | 7   | 2  |
| 17   | 3D-MoRSE descriptors/Deskryptory 3D-MoRSE oparte na dyfrakcji elektronów                                                                                                    | 255     | 101 | 19  | 5  | 225    | 103 | 8   | 3  |
| 18   | WHIM (Weighted Holistic Invariant Molecular) descriptors/Deskryptory WHIM (Ważone, Holistyczne, Niezmienne, Molekularne)                                                    | 115     | 32  | 18  | 3  | 115    | 39  | 9   | 4  |
| 19   | GETAWAY (GEometry, Topology, and Atom-Weights Assembly) descriptors/Deskryptory GETAWAY (Zespolone deskryptory geometryczne, topologiczne i ważne po właściwościach atomów) | 272     | 118 | 22  | 2  | 274    | 71  | 10  | 3  |
| 20   | Randic molecular profiles/Profile cząsteczkowe Randica                                                                                                                      | 42      | 3   | 3   | 3  | 42     | 5   | 4   | 4  |
| 21   | Functional group counts/Grupy funkcyjne                                                                                                                                     | 21      | 15  | 12  | 1  | 12     | 11  | 6   | 4  |
| 22   | Atom-centred fragments/Fragmenty strukturalne                                                                                                                               | 32      | 24  | 15  | 3  | 21     | 15  | 8   | 3  |
| 23   | Atom-type E-state indices/Zależne od atomu indeksy elektrotopologiczne                                                                                                      | 33      | 14  | 13  | 2  | 27     | 14  | 7   | 2  |
| 24   | CATS 2D descriptors/                                                                                                                                                        | 46      | 28  | 17  | 3  | 33     | 28  | 6   | 3  |
| 25   | 2D Atom Pairs/Pary atomów 2D                                                                                                                                                | 112     | 44  | 24  | 2  | 129    | 52  | 10  | 2  |
| 26   | 3D Atom Pairs/Pary atomów 3D                                                                                                                                                | 5       | 4   | 4   | 2  | 9      | 3   | 3   | 3  |
| 27   | Charge descriptors/Deskryptory ładunku                                                                                                                                      | 5       | 4   | 3   | 3  | 7      | 4   | 3   | 3  |
| 28   | Molecular properties/Właściwości cząsteczki                                                                                                                                 | 21      | 12  | 9   | 1  | 18     | 7   | 6   | 3  |
| 29   | Drug-like indices/Lekopodobność                                                                                                                                             | 17      | 14  | 10  | 0  | 11     | 8   | 5   | 0  |
| 30   | CATS 3D descriptors                                                                                                                                                         | 51      | 34  | 20  | 3  | 57     | 53  | 6   | 6  |

**Tabela A2.** Podział zbioru danych na zbiory uczące i testowe z wyszczególnieniem obserwacji odstających (outlierów)

| Kation                                              | Anion                 | HENRY   | FFV     |
|-----------------------------------------------------|-----------------------|---------|---------|
| [(N <sub>111</sub> ) <sub>2</sub> CH]               | [NTf2]                | uczący  | uczący  |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]                | uczący  | uczący  |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [C(CN) <sub>3</sub> ] | uczący  | uczący  |
| [1,2BzMPY]                                          | [NTf2]                | testowy | testowy |
| [1,3BzMPY]                                          | [NTf2]                | uczący  | uczący  |
| [1,4BzMPY]                                          | [NTf2]                | testowy | uczący  |
| [B3MPy]                                             | [NTf2]                | uczący  | testowy |
| [B4MPy]                                             | [NTf2]                | uczący  | uczący  |
| [BDMIM]                                             | [BF4]                 | uczący  | uczący  |
| [BDMIM]                                             | [PF6]                 | uczący  | uczący  |
| [BHEA]                                              | [OAc]                 | uczący  | uczący  |
| [BHEA]                                              | [Lac]                 | testowy | testowy |
| [BMIM]                                              | [NTf2]                | uczący  | uczący  |
| [BMIM]                                              | [TA]                  | uczący  | testowy |
| [BMIM]                                              | [PF6]                 | uczący  | uczący  |
| [BMIM]                                              | [BF4]                 | uczący  | uczący  |
| [BMIM]                                              | [N(CN) <sub>2</sub> ] | testowy | uczący  |
| [BMIM]                                              | [TfO]                 | outlier | testowy |
| [BMPyrr]                                            | [eFAP]                | uczący  | uczący  |
| [BMPyrr]                                            | [NTf2]                | uczący  | uczący  |
| [BMPyrr]                                            | [TA]                  | testowy | outlier |
| [BzMIM]                                             | [NTf2]                | testowy | uczący  |
| [BzMPyrr]                                           | [NTf2]                | uczący  | testowy |
| [BzPY]                                              | [NTf2]                | uczący  | uczący  |
| [C3CNDMIM]                                          | [N(CN) <sub>2</sub> ] | testowy | outlier |
| [C3CNDMIM]                                          | [NTf2]                | outlier | testowy |
| [C3CNMIM]                                           | [NTf2]                | uczący  | uczący  |
| [C3CNMIM]                                           | [N(CN) <sub>2</sub> ] | testowy | uczący  |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]                | uczący  | uczący  |
| [C <sub>8</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]                | testowy | uczący  |
| [DMIM]                                              | [BF4]                 | outlier | uczący  |
| [DMPIM]                                             | [NTf2]                | testowy | testowy |
| [EDMIM]                                             | [NTf2]                | uczący  | testowy |
| [EMIM]                                              | [NTf2]                | uczący  | uczący  |
| [EMIM]                                              | [BF4]                 | uczący  | uczący  |
| [EMIM]                                              | [N(CN) <sub>2</sub> ] | uczący  | testowy |
| [EMIM]                                              | [eFAP]                | testowy | uczący  |
| [EMIM]                                              | [TfO]                 | outlier | uczący  |
| [ETT]                                               | [eFAP]                | uczący  | uczący  |
| [HEA]                                               | [HCOO]                | uczący  | uczący  |
| [HEA]                                               | [OAc]                 | uczący  | testowy |
| [HEA]                                               | [Lac]                 | uczący  | uczący  |
| [HEEA]                                              | [Lac]                 | uczący  | uczący  |
| [HEEA]                                              | [HCOO]                | outlier | testowy |
| [HEEA]                                              | [OAc]                 | outlier | outlier |
| [HHEMEA]                                            | [Lac]                 | uczący  | testowy |
| [HHEMEA]                                            | [OAc]                 | outlier | outlier |
| [HMIM]                                              | [eFAP]                | uczący  | uczący  |
| [HMIM]                                              | [NTf2]                | uczący  | testowy |
| [HMIM]                                              | [Cl]                  | uczący  | uczący  |
| [HMIM]                                              | [pFAP]                | testowy | uczący  |
| [HMPy]                                              | [NTf2]                | testowy | testowy |
| [HOEtMIM]                                           | [BF4]                 | uczący  | uczący  |
| [HOEtMIM]                                           | [NTf2]                | testowy | testowy |
| [HOEtMIM]                                           | [PF6]                 | testowy | uczący  |
| [HOEtMIM]                                           | [TfO]                 | testowy | testowy |
| [MHEA]                                              | [HCOO]                | uczący  | testowy |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]                | uczący  | uczący  |
| [OMIM]                                              | [NTf2]                | uczący  | testowy |
| [OMIM]                                              | [BF4]                 | uczący  | uczący  |
| [OMIM]                                              | [PF6]                 | testowy | uczący  |
| [P66614]                                            | [Br]                  | uczący  | uczący  |

|           |          |         |         |
|-----------|----------|---------|---------|
| [P66614]  | [N(CN)2] | uczący  | testowy |
| [P66614]  | [Cl]     | outlier | uczący  |
| [P66614]  | [NTf2]   | outlier | uczący  |
| [PentMIM] | [bFAP]   | uczący  | uczący  |
| [PentMIM] | [NTf2]   | testowy | testowy |
| [PhBIM]   | [NTf2]   | outlier | outlier |
| [PhDBIM]  | [NTf2]   | testowy | outlier |
| [PMIM]    | [PF6]    | uczący  | uczący  |
| [PMIM]    | [NTf2]   | testowy | uczący  |
| [THEA]    | [OAc]    | uczący  | outlier |
| [THEA]    | [Lac]    | outlier | outlier |

---

Tabela A3. Porównanie eksperymentalnych i obliczonych wartości  $H_{CO_2}$  na etapie modelowania stałej Henry'ego

| Kation                                              | Anion    | Rzeczywista | LR       | MLR      | PLS      |
|-----------------------------------------------------|----------|-------------|----------|----------|----------|
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | [NTf2]   | 4070        | 3503,73  | 3325,51  | 3549,54  |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]   | 3370        | 4169,05  | 4611,96  | 4772,65  |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [C(CN)3] | 5250        | 4544,29  | 5741,01  | 5964,56  |
| [1,2BzMPY]                                          | [NTf2]   | 4170        | 4089,24  | 3978,14  | 4126,69  |
| [1,3BzMPY]                                          | [NTf2]   | 3980        | 3802,02  | 3618,58  | 3806,12  |
| [1,4BzMPY]                                          | [NTf2]   | 3930        | 3904,37  | 3757,56  | 3929,25  |
| [BzPY]                                              | [NTf2]   | 4610        | 3661,25  | 3246,44  | 3500,19  |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]   | 2840        | 3507,98  | 3543,20  | 3702,90  |
| [C <sub>8</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]   | 2730        | 3087,57  | 2654,63  | 2928,51  |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]   | 4350        | 3822,36  | 3944,87  | 4037,98  |
| [HOEtMIM]                                           | [NTf2]   | 10800       | 9603,80  | 9336,32  | 9134,25  |
| [HOEtMIM]                                           | [PF6]    | 4420        | 5776,77  | 6684,03  | 6725,01  |
| [HOEtMIM]                                           | [BF4]    | 6140        | 8195,76  | 8444,57  | 8379,94  |
| [HOEtMIM]                                           | [TfO]    | 8940        | 9603,80  | 9336,32  | 9134,25  |
| [C3CNMIM]                                           | [NTf2]   | 4210        | 7210,85  | 7878,25  | 7657,44  |
| [C3CNMIM]                                           | [N(CN)2] | 11060       | 12828,99 | 12000,34 | 11546,84 |
| [DMPIM]                                             | [NTf2]   | 3850        | 3648,11  | 3585,91  | 3765,69  |
| [BzMPyrr]                                           | [NTf2]   | 4000        | 3837,38  | 3783,61  | 3920,71  |
| [BzMIM]                                             | [NTf2]   | 3990        | 3516,60  | 3268,31  | 3500,01  |
| [BMPyrr]                                            | [TA]     | 2850        | 4075,83  | 4031,26  | 4291,44  |
| [BMPyrr]                                            | [eFAP]   | 3860        | 3519,33  | 3166,33  | 3394,09  |
| [BMPyrr]                                            | [NTf2]   | 5320        | 3923,95  | 3803,19  | 4098,51  |
| [BDMIM]                                             | [BF4]    | 6100        | 5908,41  | 6222,08  | 6161,36  |
| [BDMIM]                                             | [PF6]    | 6180        | 5908,41  | 6222,08  | 6161,36  |
| [BMIM]                                              | [NTf2]   | 3300        | 3255,32  | 2820,79  | 3101,29  |
| [BMIM]                                              | [TA]     | 3850        | 3595,71  | 3457,65  | 3805,71  |
| [BMIM]                                              | [N(CN)2] | 5340        | 5087,97  | 5473,07  | 5510,53  |
| [BMIM]                                              | [PF6]    | 5600        | 5087,97  | 5473,07  | 5510,53  |
| [BMIM]                                              | [BF4]    | 5590        | 5967,50  | 6942,88  | 6990,69  |
| [B3MPy]                                             | [NTf2]   | 3300        | 3869,06  | 3812,75  | 3963,74  |
| [B4MPy]                                             | [NTf2]   | 3830        | 3829,03  | 3758,30  | 3915,50  |
| [EDMIM]                                             | [NTf2]   | 3960        | 3714,46  | 3712,76  | 3874,05  |
| [EMIM]                                              | [NTf2]   | 3770        | 3321,01  | 2957,85  | 3220,94  |
| [EMIM]                                              | [BF4]    | 8110        | 5231,86  | 5610,14  | 5630,18  |
| [EMIM]                                              | [eFAP]   | 8210        | 6140,84  | 7079,94  | 7110,35  |
| [EMIM]                                              | [N(CN)2] | 3323        | 3812,16  | 3822,78  | 4118,29  |
| [HMIM]                                              | [pFAP]   | 2520        | 3582,35  | 3467,08  | 3805,64  |
| [HMIM]                                              | [eFAP]   | 3500        | 3149,65  | 2602,14  | 2908,29  |
| [HMIM]                                              | [NTf2]   | 11800       | 10883,78 | 10267,94 | 9878,12  |
| [HMIM]                                              | [Cl]     | 2160        | 3481,87  | 3275,33  | 3643,44  |
| [HMPy]                                              | [NTf2]   | 3280        | 3232,67  | 2775,11  | 3045,74  |
| [OMIM]                                              | [NTf2]   | 3000        | 3034,19  | 2332,90  | 2671,65  |
| [OMIM]                                              | [PF6]    | 5390        | 4590,72  | 4985,19  | 5080,89  |
| [OMIM]                                              | [BF4]    | 5140        | 4590,72  | 4985,19  | 5080,89  |
| [PMIM]                                              | [NTf2]   | 5200        | 5090,35  | 5470,73  | 5509,08  |
| [PMIM]                                              | [PF6]    | 3700        | 3256,39  | 2818,45  | 3099,84  |
| [PentMIM]                                           | [NTf2]   | 2020        | 3705,28  | 3610,25  | 3944,05  |
| [PentMIM]                                           | [bFAP]   | 2880        | 3378,62  | 3048,82  | 3303,43  |
| [HEEA]                                              | [Lac]    | 16214       | 17983,04 | 16846,12 | 15867,40 |
| [HEA]                                               | [HCOO]   | 9500        | 9896,03  | 10152,74 | 10093,29 |
| [HEA]                                               | [OAc]    | 13217       | 11967,05 | 11407,95 | 11155,06 |
| [HEA]                                               | [Lac]    | 14200       | 17273,59 | 15746,48 | 14824,96 |
| [HHMEEA]                                            | [Lac]    | 19500       | 19011,73 | 20468,41 | 19196,50 |
| [C3CNDMIM]                                          | [N(CN)2] | 13630       | 13565,94 | 12429,29 | 11923,87 |
| [PhDBIM]                                            | [NTf2]   | 6300        | 4036,82  | 4048,22  | 4185,10  |
| [BHEA]                                              | [OAc]    | 14100       | 14088,32 | 13446,56 | 13204,74 |
| [BHEA]                                              | [Lac]    | 17700       | 18609,55 | 19040,30 | 17936,41 |
| [ETT]                                               | [eFAP]   | 2940        | 4174,69  | 4374,85  | 4614,48  |
| [MHEA]                                              | [HCOO]   | 13950       | 12620,87 | 11781,12 | 11481,56 |
| [THEA]                                              | [OAc]    | 18238       | 18321,28 | 18573,71 | 17940,82 |
| [P66614]                                            | [Br]     | 2700        | 4437,03  | 5225,26  | 5492,79  |
| [P66614]                                            | [N(CN)2] | 5880        | 8728,80  | 8768,96  | 8573,22  |

Tabela A4. Wartości deskryptorów obliczonych na etapie modelowania stałej Henry'ego

| Kation                                              | Anion                 | GATS2s      | P_VSA_ppp_A | Chi_G/D     | B02[F-F] | Hy        | CATS3D_04_AA |
|-----------------------------------------------------|-----------------------|-------------|-------------|-------------|----------|-----------|--------------|
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | [NTf2]                | 0,885781    | 5,023407    | 0,994816    | 1        | -0,13023  | 2            |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]                | 0,932432    | 21,808937   | 0,996452    | 1        | -0,13023  | 2            |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [C(CN) <sub>3</sub> ] | 0,932432    | 21,808937   | 0,996452    | 0        | -0,534734 | 3            |
| [1,2BzMPY]                                          | [NTf2]                | 0,983914    | 1,899093    | 1,055888    | 1        | -0,13023  | 2            |
| [1,3BzMPY]                                          | [NTf2]                | 0,968932    | 1,899093    | 1,035659    | 1        | -0,13023  | 2            |
| [1,4BzMPY]                                          | [NTf2]                | 0,968932    | 1,899093    | 1,042548    | 1        | -0,13023  | 2            |
| [BzPY]                                              | [NTf2]                | 1,14374     | 1,899093    | 1,045291    | 1        | -0,13023  | 2            |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]                | 0,63664     | 3,798186    | 0,97067     | 1        | -0,13023  | 2            |
| [C <sub>8</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]                | 0,7309      | 3,798186    | 0,941765    | 1        | -0,13023  | 2            |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]                | 0,612256    | 0           | 1,002418    | 1        | -0,13023  | 2            |
| [HOEtMIM]                                           | [NTf2]                | 0,957244    | 46,481529   | 1,000807    | 1        | -0,13023  | 2            |
| [HOEtMIM]                                           | [PF6]                 | 0,957244    | 46,481529   | 1,000807    | 1        | -0,118299 | 0            |
| [HOEtMIM]                                           | [BF4]                 | 0,957244    | 46,481529   | 1,000807    | 1        | 0         | 0            |
| [HOEtMIM]                                           | [TfO]                 | 0,957244    | 46,481529   | 1,000807    | 1        | 0         | 0            |
| [C3CNMIM]                                           | [NTf2]                | 0,426344    | 34,85754    | 1,022942    | 1        | -0,13023  | 2            |
| [C3CNMIM]                                           | [N(CN) <sub>2</sub> ] | 0,426344    | 34,85754    | 1,022942    | 0        | -0,359298 | 1            |
| [DMPIM]                                             | [NTf2]                | 0,820214    | 3,798186    | 1,002274    | 1        | -0,13023  | 2            |
| [BzMPyrr]                                           | [NTf2]                | 0,800464    | 0           | 1,024656    | 1        | -0,13023  | 2            |
| [BzMIM]                                             | [NTf2]                | 0,935649    | 3,798186    | 1,005073    | 1        | -0,13023  | 2            |
| [BMPyrr]                                            | [TA]                  | 0,949405    | 0           | 1,017982    | 1        | -0,237111 | 0            |
| [BMPyrr]                                            | [eFAP]                | 0,949405    | 0           | 1,017982    | 1        | -0,13023  | 2            |
| [BMPyrr]                                            | [NTf2]                | 0,949405    | 0           | 1,017982    | 1        | -0,267367 | 0            |
| [BDMIM]                                             | [BF4]                 | 0,825418    | 3,798186    | 1,002311    | 1        | 0         | 0            |
| [BDMIM]                                             | [PF6]                 | 0,825418    | 3,798186    | 1,002311    | 1        | 0         | 0            |
| [BMIM]                                              | [NTf2]                | 0,919067    | 3,798186    | 0,980226    | 1        | -0,13023  | 2            |
| [BMIM]                                              | [TA]                  | 0,919067    | 3,798186    | 0,980226    | 1        | -0,267367 | 0            |
| [BMIM]                                              | [N(CN) <sub>2</sub> ] | 0,919067    | 3,798186    | 0,980226    | 1        | 0         | 0            |
| [BMIM]                                              | [PF6]                 | 0,919067    | 3,798186    | 0,980226    | 0        | -0,359298 | 1            |
| [BMIM]                                              | [BF4]                 | 0,919067    | 3,798186    | 0,980226    | 1        | 0         | 0            |
| [B3MPy]                                             | [NTf2]                | 0,862999    | 1,899093    | 1,028268    | 1        | -0,13023  | 2            |
| [B4MPy]                                             | [NTf2]                | 0,862999    | 1,899093    | 1,025569    | 1        | -0,13023  | 2            |
| [EDMIM]                                             | [NTf2]                | 0,790707    | 3,798186    | 1,003822    | 1        | -0,13023  | 2            |
| [EMIM]                                              | [NTf2]                | 0,906023    | 3,798186    | 0,984925    | 1        | -0,237111 | 0            |
| [EMIM]                                              | [BF4]                 | 0,906023    | 3,798186    | 0,984925    | 1        | -0,13023  | 2            |
| [EMIM]                                              | [eFAP]                | 0,906023    | 3,798186    | 0,984925    | 1        | 0         | 0            |
| [EMIM]                                              | [N(CN) <sub>2</sub> ] | 0,906023    | 3,798186    | 0,984925    | 0        | -0,359298 | 1            |
| [HMIM]                                              | [pFAP]                | 0,924233    | 3,798186    | 0,970218    | 1        | -0,262548 | 0            |
| [HMIM]                                              | [eFAP]                | 0,924233    | 3,798186    | 0,970218    | 1        | -0,237111 | 0            |
| [HMIM]                                              | [NTf2]                | 0,924233    | 3,798186    | 0,970218    | 1        | -0,13023  | 2            |
| [HMIM]                                              | [Cl]                  | 0,924233    | 3,798186    | 0,970218    | 0        | 0         | 0            |
| [HMPy]                                              | [NTf2]                | 0,872281    | 1,899093    | 0,978324    | 1        | -0,13023  | 2            |
| [OMIM]                                              | [NTf2]                | 0,938092    | 3,798186    | 0,959098    | 1        | -0,13023  | 2            |
| [OMIM]                                              | [PF6]                 | 0,938092    | 3,798186    | 0,959098    | 1        | 0         | 0            |
| [OMIM]                                              | [BF4]                 | 0,938092    | 3,798186    | 0,959098    | 1        | 0         | 0            |
| [PMIM]                                              | [NTf2]                | 0,923648    | 3,798186    | 0,980846    | 1        | -0,13023  | 2            |
| [PMIM]                                              | [PF6]                 | 0,923648    | 3,798186    | 0,980846    | 1        | 0         | 0            |
| [PentMIM]                                           | [NTf2]                | 0,919992    | 3,798186    | 0,991678    | 1        | -0,277374 | 0            |
| [PentMIM]                                           | [bFAP]                | 0,919992    | 3,798186    | 0,991678    | 1        | -0,13023  | 2            |
| [HEEA]                                              | [Lac]                 | 1,118682615 | 53,68323105 | 1,004252507 | 0        | 0,311391  | 0            |
| [HEA]                                               | [HCOO]                | 0,988203    | 42,683342   | 0,974413    | 0        | -0,26416  | 0            |
| [HEA]                                               | [OAc]                 | 0,988203    | 42,683342   | 0,974413    | 0        | -0,430677 | 0            |
| [HEA]                                               | [Lac]                 | 0,988203    | 42,683342   | 0,974413    | 0        | 0,311391  | 0            |
| [HHEMEA]                                            | [Lac]                 | 0,931437    | 85,366684   | 1,022308    | 0        | 0,311391  | 0            |
| [C3CNDMIM]                                          | [N(CN) <sub>2</sub> ] | 0,40437     | 34,85754    | 1,040675    | 0        | -0,359298 | 1            |
| [PhDBIM]                                            | [NTf2]                | 0,892604    | 3,798186    | 1,036818    | 1        | -0,13023  | 2            |
| [BHEA]                                              | [OAc]                 | 0,968922    | 85,366684   | 0,957539    | 0        | -0,430677 | 0            |
| [BHEA]                                              | [Lac]                 | 0,968922    | 85,366684   | 0,957539    | 0        | 0,311391  | 0            |
| [ETT]                                               | [eFAP]                | 0,875498    | 6,248628    | 0,997223    | 1        | -0,237111 | 0            |
| [MHEA]                                              | [HCOO]                | 0,958106    | 42,683342   | 0,988076    | 0        | -0,26416  | 0            |
| [THEA]                                              | [OAc]                 | 0,952991    | 128,050034  | 1,032079    | 0        | -0,430677 | 0            |
| [P66614]                                            | [Br]                  | 1,222253    | 0           | 0,95954     | 0        | -0,359298 | 1            |
| [P66614]                                            | [N(CN) <sub>2</sub> ] | 1,222253    | 0           | 0,95954     | 0        | 0         | 0            |

Tabela A5. Wartości deskryptorów oraz FFV obliczonych dla modelu FFV jako zmienna zależna

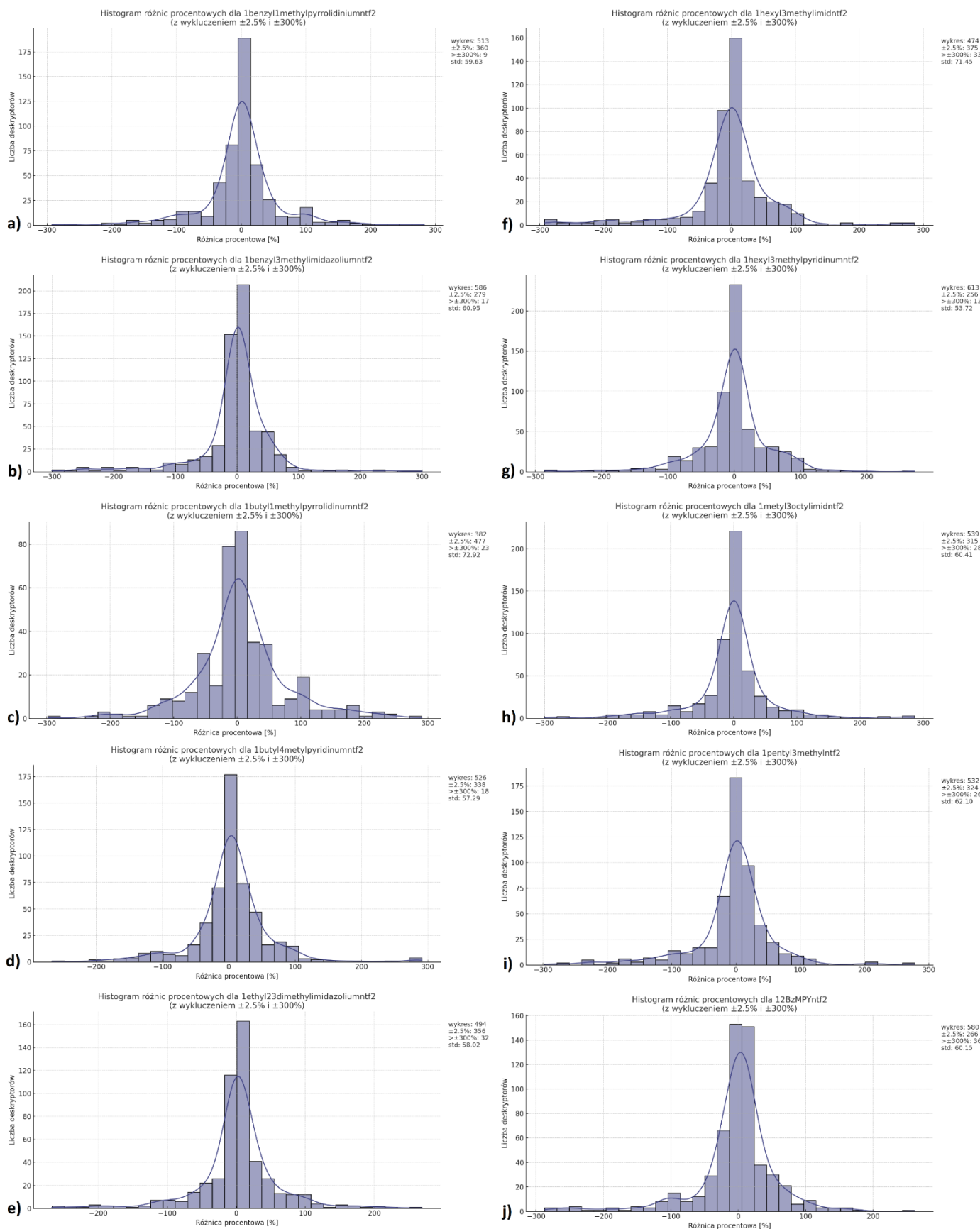
| Kation                                              | Anion                 | R4i   | R7s+  | nN+ | nHAcc | C-042 | DLS_04 | Rzeczywiste FFV | Obliczone FFV |
|-----------------------------------------------------|-----------------------|-------|-------|-----|-------|-------|--------|-----------------|---------------|
| [BMIM]                                              | [TfO]                 | 2.334 | 0.054 | 1   | 0     | 1     | 1      | 0.193           | 0.193034      |
| [PentMIM]                                           | [bFAP]                | 2.397 | 0.038 | 1   | 0     | 1     | 1      | 0.187           | 0.193752      |
| [HMIM]                                              | [pFAP]                | 2.45  | 0.04  | 1   | 0     | 1     | 1      | 0.182           | 0.194875      |
| [HMIM]                                              | [eFAP]                | 2.45  | 0.04  | 1   | 0     | 1     | 1      | 0.228           | 0.194875      |
| [P66614]                                            | [N(CN) <sub>2</sub> ] | 2.56  | 0.01  | 0   | 0     | 0     | 0.2    | 0.221           | 0.223398      |
| [C <sub>6</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]                | 3.052 | 1.336 | 1   | 13    | 1     | 1      | 0.204           | 0.196553      |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]                | 2.813 | 0.275 | 1   | 9     | 1     | 1      | 0.166           | 0.172675      |
| [BMPyrr]                                            | [eFAP]                | 3.477 | 0.041 | 1   | 0     | 0     | 0.4    | 0.190           | 0.176083      |
| [PentMIM]                                           | [NTf2]                | 2.397 | 0.038 | 1   | 0     | 1     | 1      | 0.180           | 0.193752      |
| [DMIM]                                              | [BF4]                 | 2.051 | 0     | 1   | 0     | 1     | 0.8    | 0.184           | 0.177314      |
| [ETT]                                               | [eFAP]                | 1.995 | 0.053 | 0   | 2     | 0     | 0.6    | 0.202           | 0.221805      |
| [OMIM]                                              | [NTf2]                | 2.556 | 0.027 | 1   | 0     | 1     | 0.8    | 0.186           | 0.18829       |
| [P66614]                                            | [Cl]                  | 2.56  | 0.01  | 0   | 0     | 0     | 0.2    | 0.230           | 0.223398      |
| [HMPy]                                              | [NTf2]                | 2.371 | 0.04  | 1   | 0     | 0     | 0.9    | 0.175           | 0.174653      |
| [BMIM]                                              | [NTf2]                | 2.334 | 0.054 | 1   | 0     | 1     | 1      | 0.184           | 0.193034      |
| [B3MPy]                                             | [NTf2]                | 2.289 | 0.07  | 1   | 0     | 0     | 0.9    | 0.170           | 0.174028      |
| [EMIM]                                              | [eFAP]                | 1.877 | 0.044 | 1   | 0     | 1     | 0.8    | 0.134           | 0.175326      |
| [P66614]                                            | [NTf2]                | 2.56  | 0.01  | 0   | 0     | 0     | 0.2    | 0.228           | 0.223398      |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]                | 2.485 | 0.062 | 1   | 2     | 0     | 1      | 0.170           | 0.173497      |
| [HMIM]                                              | [NTf2]                | 2.45  | 0.04  | 1   | 0     | 1     | 1      | 0.181           | 0.194875      |
| [PMIM]                                              | [NTf2]                | 2.234 | 0.04  | 1   | 0     | 1     | 1      | 0.208           | 0.190569      |
| [EMIM]                                              | [NTf2]                | 1.877 | 0.044 | 1   | 0     | 1     | 0.8    | 0.181           | 0.175326      |
| [B4MPy]                                             | [NTf2]                | 2.445 | 0.045 | 1   | 0     | 0     | 0.9    | 0.169           | 0.176296      |
| [DMPIM]                                             | [NTf2]                | 2.142 | 0.038 | 1   | 0     | 0     | 1      | 0.163           | 0.174151      |
| [BMIM]                                              | [TA]                  | 2.334 | 0.054 | 1   | 0     | 1     | 1      | 0.182           | 0.193034      |
| [BMPyrr]                                            | [NTf2]                | 3.477 | 0.041 | 1   | 0     | 0     | 0.4    | 0.159           | 0.176083      |
| [1,4BzMPY]                                          | [NTf2]                | 2.287 | 0.047 | 1   | 0     | 0     | 0.5    | 0.151           | 0.156692      |
| [EDMIM]                                             | [NTf2]                | 1.822 | 0.042 | 1   | 0     | 0     | 1      | 0.205           | 0.167906      |
| [1,3BzMPY]                                          | [NTf2]                | 2.241 | 0.066 | 1   | 0     | 0     | 0.5    | 0.149           | 0.156414      |
| [BzMIM]                                             | [NTf2]                | 1.749 | 0.048 | 1   | 0     | 1     | 0.6    | 0.185           | 0.164648      |
| [BzMPyrr]                                           | [NTf2]                | 2.782 | 0.037 | 1   | 0     | 0     | 0.6    | 0.149           | 0.170354      |
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | [NTf2]                | 2.759 | 0     | 1   | 1     | 0     | 0.8    | 0.173           | 0.172761      |
| [1,2BzMPY]                                          | [NTf2]                | 2.015 | 0.07  | 1   | 0     | 0     | 0.5    | 0.145           | 0.152044      |
| [C3CNMIM]                                           | [NTf2]                | 2.267 | 0.376 | 1   | 1     | 1     | 0.8    | 0.200           | 0.190126      |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]                | 4.13  | 0.037 | 1   | 0     | 0     | 0.2    | 0.169           | 0.180705      |
| [HOEtMIM]                                           | [NTf2]                | 1.839 | 0.188 | 1   | 1     | 1     | 1      | 0.190           | 0.183527      |
| [BzPY]                                              | [NTf2]                | 2     | 0.07  | 1   | 0     | 0     | 0.5    | 0.146           | 0.151744      |
| [C3CNDMIM]                                          | [NTf2]                | 2.167 | 0.475 | 1   | 1     | 0     | 0.8    | 0.156           | 0.176947      |
| [OMIM]                                              | [PF6]                 | 2.556 | 0.027 | 1   | 0     | 1     | 0.8    | 0.204           | 0.18829       |
| [PMIM]                                              | [PF6]                 | 2.234 | 0.04  | 1   | 0     | 1     | 1      | 0.192           | 0.190569      |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [C(CN) <sub>3</sub> ] | 2.485 | 0.062 | 1   | 2     | 0     | 1      | 0.189           | 0.173497      |
| [BMIM]                                              | [PF6]                 | 2.334 | 0.054 | 1   | 0     | 1     | 1      | 0.191           | 0.193034      |
| [OMIM]                                              | [BF4]                 | 2.556 | 0.027 | 1   | 0     | 1     | 0.8    | 0.204           | 0.18829       |
| [BMIM]                                              | [N(CN) <sub>2</sub> ] | 2.334 | 0.054 | 1   | 0     | 1     | 1      | 0.178           | 0.193034      |
| [BMIM]                                              | [BF4]                 | 2.334 | 0.054 | 1   | 0     | 1     | 1      | 0.191           | 0.193034      |
| [P66614]                                            | [Br]                  | 2.56  | 0.01  | 0   | 0     | 0     | 0.2    | 0.232           | 0.223398      |
| [BDMIM]                                             | [BF4]                 | 2.272 | 0.049 | 1   | 0     | 0     | 1      | 0.188           | 0.177112      |
| [HOEtMIM]                                           | [TfO]                 | 1.839 | 0.188 | 1   | 1     | 1     | 1      | 0.169           | 0.183527      |
| [BDMIM]                                             | [PF6]                 | 2.272 | 0.049 | 1   | 0     | 0     | 1      | 0.192           | 0.177112      |
| [EMIM]                                              | [TfO]                 | 1.877 | 0.044 | 1   | 0     | 1     | 0.8    | 0.188           | 0.175326      |
| [EMIM]                                              | [BF4]                 | 1.877 | 0.044 | 1   | 0     | 1     | 0.8    | 0.183           | 0.175326      |
| [EMIM]                                              | [N(CN) <sub>2</sub> ] | 1.877 | 0.044 | 1   | 0     | 1     | 0.8    | 0.172           | 0.175326      |
| [HOEtMIM]                                           | [PF6]                 | 1.839 | 0.188 | 1   | 1     | 1     | 1      | 0.171           | 0.183527      |
| [HEA]                                               | [OAc]                 | 1.502 | 0     | 1   | 1     | 0     | 0.6    | 0.152           | 0.139441      |
| [HOEtMIM]                                           | [BF4]                 | 1.839 | 0.188 | 1   | 1     | 1     | 1      | 0.167           | 0.183527      |
| [C3CNMIM]                                           | [N(CN) <sub>2</sub> ] | 2.267 | 0.376 | 1   | 1     | 1     | 0.8    | 0.163           | 0.190126      |
| [HMIM]                                              | [Cl]                  | 2.45  | 0.04  | 1   | 0     | 1     | 1      | 0.184           | 0.194875      |
| [HEA]                                               | [HCOO]                | 1.502 | 0     | 1   | 1     | 0     | 0.6    | 0.135           | 0.139441      |
| [MHEA]                                              | [HCOO]                | 1.865 | 0     | 1   | 1     | 0     | 0.6    | 0.152           | 0.146678      |
| [BHEA]                                              | [OAc]                 | 2.032 | 0.175 | 1   | 2     | 0     | 0.5    | 0.150           | 0.147617      |
| [HEA]                                               | [Lac]                 | 1.502 | 0     | 1   | 1     | 0     | 0.6    | 0.133           | 0.139441      |
| [HEEA]                                              | [Lac]                 | 1.427 | 0.232 | 1   | 2     | 0     | 0.5    | 0.153           | 0.137475      |

Tabela A6. Wartości deskryptorów oraz H<sub>CO2</sub> obliczonych dla modelu standardowego

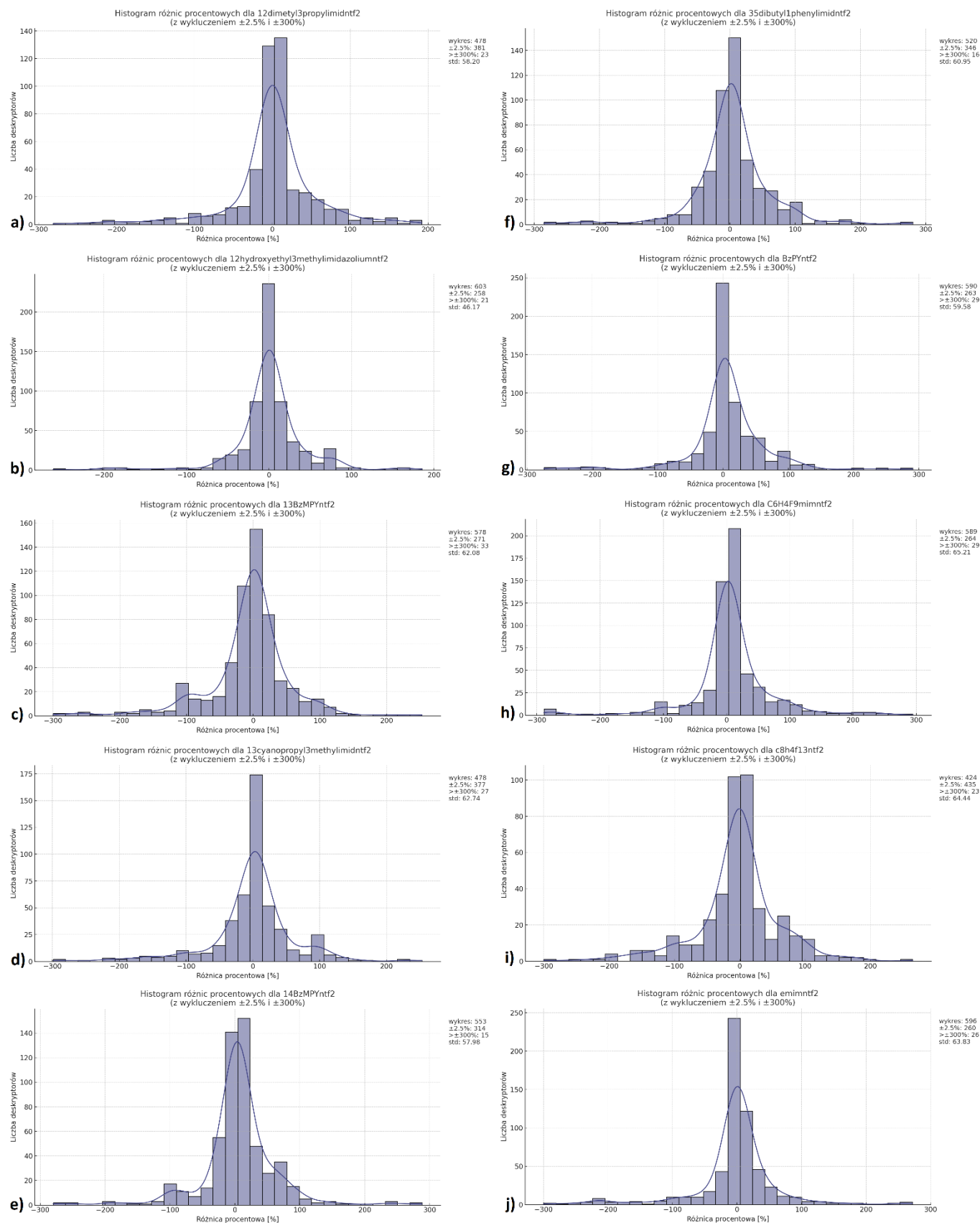
| Kation                                              | Anion                 | R3i   | R7i   | R7i+  | C-042 | CATS3D_06_AL | B06[F-F] | Rzeczywista H <sub>CO2</sub> | Obliczona H <sub>CO2</sub> |
|-----------------------------------------------------|-----------------------|-------|-------|-------|-------|--------------|----------|------------------------------|----------------------------|
| [BMIM]                                              | [TfO]                 | 2.325 | 0.72  | 0.056 | 1     | 1            | 0        | 1428                         | 6959.53                    |
| [PentMIM]                                           | [bFAP]                | 2.627 | 0.731 | 0.043 | 1     | 1            | 1        | 2020                         | 2751.96                    |
| [HMIM]                                              | [pFAP]                | 2.48  | 0.75  | 0.04  | 1     | 2            | 1        | 2160                         | 2565.71                    |
| [HMIM]                                              | [eFAP]                | 2.48  | 0.75  | 0.04  | 1     | 2            | 1        | 2520                         | 2565.71                    |
| [P66614]                                            | [N(CN) <sub>2</sub> ] | 3.27  | 1.42  | 0.01  | 0     | 0            | 0        | 2700                         | 4503.57                    |
| [C <sub>8</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]                | 3.439 | 1.032 | 0.05  | 1     | 0            | 1        | 2730                         | 2879.96                    |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]                | 2.995 | 0.879 | 0.045 | 1     | 0            | 1        | 2840                         | 2360.64                    |
| [BMPyrr]                                            | [eFAP]                | 3.223 | 1.132 | 0.049 | 0     | 0            | 1        | 2850                         | 4920.01                    |
| [PentMIM]                                           | [NTf2]                | 2.627 | 0.731 | 0.043 | 1     | 1            | 1        | 2880                         | 2751.96                    |
| [DMIM]                                              | [BF4]                 | 1.272 | 0     | 0     | 1     | 0            | 0        | 2900                         | 4485.62                    |
| [ETT]                                               | [eFAP]                | 1.399 | 1.34  | 0.078 | 0     | 0            | 1        | 2940                         | 725.75                     |
| [OMIM]                                              | [NTf2]                | 2.623 | 0.778 | 0.027 | 1     | 2            | 1        | 3000                         | 1863.74                    |
| [P66614]                                            | [Cl]                  | 3.27  | 1.42  | 0.01  | 0     | 0            | 0        | 3040                         | 4503.57                    |
| [HMPy]                                              | [NTf2]                | 2.449 | 0.762 | 0.032 | 0     | 1            | 1        | 3280                         | 4654.43                    |
| [BMIM]                                              | [NTf2]                | 2.325 | 0.72  | 0.056 | 1     | 1            | 1        | 3300                         | 2921.20                    |
| [B3MPy]                                             | [NTf2]                | 2.319 | 0.808 | 0.048 | 0     | 0            | 1        | 3300                         | 4578.23                    |
| [EMIM]                                              | [eFAP]                | 2.025 | 0.461 | 0.064 | 1     | 0            | 1        | 3323                         | 3840.55                    |
| [P66614]                                            | [NTf2]                | 3.27  | 1.42  | 0.01  | 0     | 0            | 1        | 3340                         | 465.24                     |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]                | 1.184 | 0.592 | 0.037 | 0     | 0            | 1        | 3370                         | 2226.36                    |
| [HMIM]                                              | [NTf2]                | 2.48  | 0.75  | 0.04  | 1     | 2            | 1        | 3500                         | 2565.71                    |
| [PMIM]                                              | [NTf2]                | 2.336 | 0.508 | 0.05  | 1     | 0            | 1        | 3700                         | 3390.25                    |
| [EMIM]                                              | [NTf2]                | 2.025 | 0.461 | 0.064 | 1     | 0            | 1        | 3770                         | 3840.55                    |
| [B4MPy]                                             | [NTf2]                | 2.449 | 0.899 | 0.053 | 0     | 0            | 1        | 3830                         | 4671.68                    |
| [DMPIM]                                             | [NTf2]                | 2.104 | 0.809 | 0.054 | 0     | 0            | 1        | 3850                         | 4416.15                    |
| [BMIM]                                              | [TA]                  | 2.325 | 0.72  | 0.056 | 1     | 1            | 0        | 3850                         | 6959.53                    |
| [BMPyrr]                                            | [NTf2]                | 3.223 | 1.132 | 0.049 | 0     | 0            | 1        | 3860                         | 4920.01                    |
| [1,4BzMPY]                                          | [NTf2]                | 2.197 | 0.857 | 0.048 | 0     | 0            | 1        | 3930                         | 3933.89                    |
| [EDMIM]                                             | [NTf2]                | 1.778 | 0.393 | 0.061 | 0     | 0            | 1        | 3960                         | 6761.09                    |
| [1,3BzMPY]                                          | [NTf2]                | 2.163 | 0.72  | 0.051 | 0     | 0            | 1        | 3980                         | 4948.51                    |
| [BzMIM]                                             | [NTf2]                | 1.812 | 0.701 | 0.058 | 1     | 1            | 1        | 3990                         | 1824.34                    |
| [BzMPyrr]                                           | [NTf2]                | 2.656 | 0.612 | 0.044 | 0     | 0            | 1        | 4000                         | 6479.28                    |
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | [NTf2]                | 1.118 | 0     | 0     | 0     | 0            | 1        | 4070                         | 3375.86                    |
| [1,2BzMPY]                                          | [NTf2]                | 2.122 | 1.039 | 0.054 | 0     | 0            | 1        | 4170                         | 2956.92                    |
| [C3CNMIM]                                           | [NTf2]                | 2.397 | 0.693 | 0.09  | 1     | 0            | 1        | 4210                         | 5101.21                    |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]                | 2.678 | 1.045 | 0.055 | 0     | 0            | 1        | 4350                         | 4460.01                    |
| [HOEtMIM]                                           | [NTf2]                | 2.076 | 0.542 | 0.081 | 1     | 0            | 1        | 4420                         | 4618.60                    |
| [BzPY]                                              | [NTf2]                | 2.345 | 0.775 | 0.055 | 0     | 0            | 1        | 4610                         | 5346.58                    |
| [C3CNDMIM]                                          | [NTf2]                | 2.164 | 1.012 | 0.06  | 0     | 0            | 1        | 4620                         | 3659.33                    |
| [OMIM]                                              | [PF6]                 | 2.623 | 0.778 | 0.027 | 1     | 2            | 0        | 5140                         | 5902.07                    |
| [PMIM]                                              | [PF6]                 | 2.336 | 0.508 | 0.05  | 1     | 0            | 0        | 5200                         | 7428.58                    |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [C(CN) <sub>3</sub> ] | 1.184 | 0.592 | 0.037 | 0     | 0            | 0        | 5250                         | 6264.69                    |
| [BMIM]                                              | [PF6]                 | 2.325 | 0.72  | 0.056 | 1     | 1            | 0        | 5340                         | 6959.53                    |
| [OMIM]                                              | [BF4]                 | 2.623 | 0.778 | 0.027 | 1     | 2            | 0        | 5390                         | 5902.07                    |
| [BMIM]                                              | [N(CN) <sub>2</sub> ] | 2.325 | 0.72  | 0.056 | 1     | 1            | 0        | 5590                         | 6959.53                    |
| [BMIM]                                              | [BF4]                 | 2.325 | 0.72  | 0.056 | 1     | 1            | 0        | 5600                         | 6959.53                    |
| [P66614]                                            | [Br]                  | 3.27  | 1.42  | 0.01  | 0     | 0            | 0        | 5880                         | 4503.57                    |
| [BDMIM]                                             | [BF4]                 | 2.14  | 0.888 | 0.048 | 0     | 1            | 0        | 6100                         | 8152.96                    |
| [HOEtMIM]                                           | [TfO]                 | 2.076 | 0.542 | 0.081 | 1     | 0            | 0        | 6140                         | 8656.93                    |
| [BDMIM]                                             | [PF6]                 | 2.14  | 0.888 | 0.048 | 0     | 1            | 0        | 6180                         | 8152.96                    |
| [EMIM]                                              | [TfO]                 | 2.025 | 0.461 | 0.064 | 1     | 0            | 0        | 7400                         | 7878.88                    |
| [EMIM]                                              | [BF4]                 | 2.025 | 0.461 | 0.064 | 1     | 0            | 0        | 8110                         | 7878.88                    |
| [EMIM]                                              | [N(CN) <sub>2</sub> ] | 2.025 | 0.461 | 0.064 | 1     | 0            | 0        | 8210                         | 7878.88                    |
| [HOEtMIM]                                           | [PF6]                 | 2.076 | 0.542 | 0.081 | 1     | 0            | 0        | 8940                         | 8656.93                    |
| [HEA]                                               | [OAc]                 | 2.934 | 0     | 0     | 0     | 0            | 0        | 9500                         | 12226.47                   |
| [HOEtMIM]                                           | [BF4]                 | 2.076 | 0.542 | 0.081 | 1     | 0            | 0        | 10800                        | 8656.93                    |
| [C3CNMIM]                                           | [N(CN) <sub>2</sub> ] | 2.397 | 0.693 | 0.09  | 1     | 0            | 0        | 11060                        | 9139.54                    |
| [HMIM]                                              | [Cl]                  | 2.48  | 0.75  | 0.04  | 1     | 2            | 0        | 11800                        | 6604.04                    |
| [HEA]                                               | [HCOO]                | 2.934 | 0     | 0     | 0     | 0            | 0        | 13217                        | 12226.47                   |
| [MHEA]                                              | [HCOO]                | 3.077 | 0     | 0     | 0     | 0            | 0        | 13950                        | 12605.42                   |
| [BHEA]                                              | [OAc]                 | 3.069 | 0.294 | 0.07  | 0     | 0            | 0        | 14100                        | 15490.44                   |
| [HEA]                                               | [Lac]                 | 2.934 | 0     | 0     | 0     | 0            | 0        | 14200                        | 12226.47                   |
| [HEEA]                                              | [Lac]                 | 2.634 | 0.635 | 0.076 | 0     | 0            | 0        | 16214                        | 12517.75                   |

Tabela A7. Wartości deskryptorów oraz H<sub>CO2</sub> obliczonych dla modelu FFV

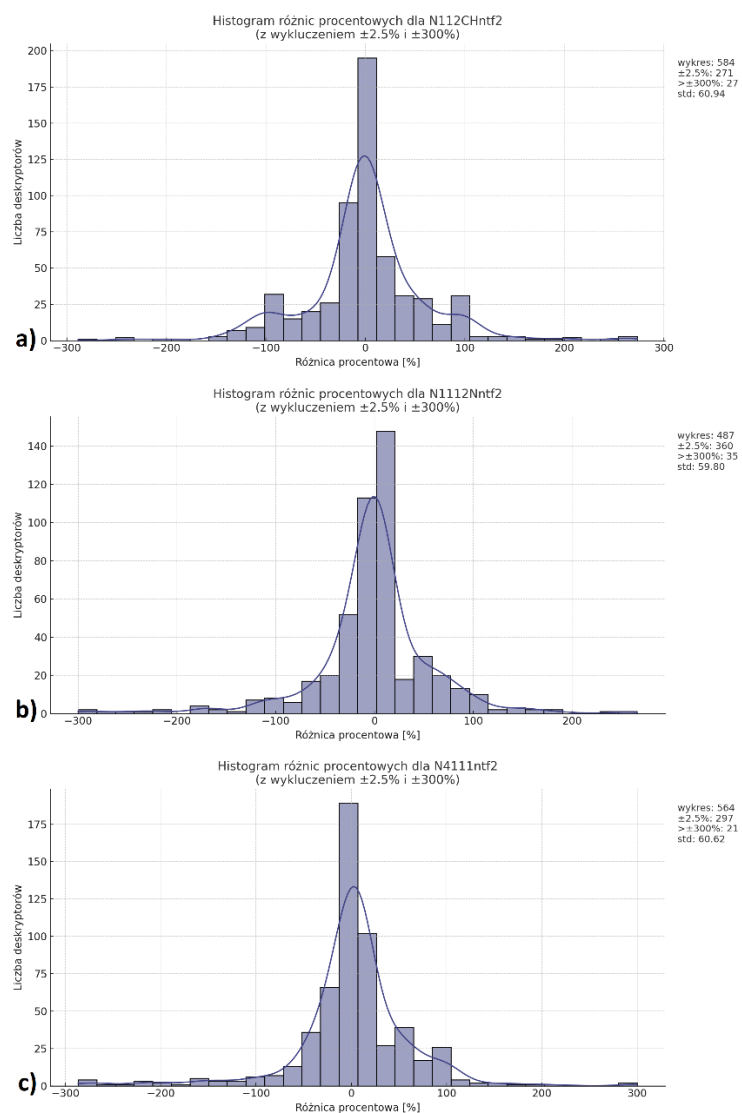
| Kation                                              | Anion                 | FFV   | R7i+  | B06[F-F] | F01[C-O] | F02[C-C] | DLS_04 | Rzeczywista H <sub>CO2</sub> | Obliczona H <sub>CO2</sub> |
|-----------------------------------------------------|-----------------------|-------|-------|----------|----------|----------|--------|------------------------------|----------------------------|
| [BMIM]                                              | [TfO]                 | 0.193 | 0.056 | 0        | 0        | 0        | 0.5    | 1428                         | 6358.45                    |
| [PentMIM]                                           | [bFAP]                | 0.187 | 0.043 | 1        | 0        | 9        | 0.1    | 2020                         | 1028.56                    |
| [HMIM]                                              | [pFAP]                | 0.182 | 0.04  | 1        | 0        | 6        | 0.1    | 2160                         | 2222.54                    |
| [HMIM]                                              | [eFAP]                | 0.228 | 0.04  | 1        | 0        | 3        | 0.2    | 2520                         | 2130.48                    |
| [P66614]                                            | [N(CN) <sub>2</sub> ] | 0.221 | 0.01  | 0        | 0        | 1        | 0.2    | 2700                         | 4416.76                    |
| [C <sub>6</sub> H <sub>4</sub> F <sub>13</sub> MIM] | [NTf2]                | 0.204 | 0.05  | 1        | 0        | 0        | 0.6    | 2730                         | 3099.88                    |
| [C <sub>6</sub> H <sub>4</sub> F <sub>9</sub> MIM]  | [NTf2]                | 0.166 | 0.045 | 1        | 0        | 0        | 0.6    | 2840                         | 3737.25                    |
| [BMPyrr]                                            | [eFAP]                | 0.190 | 0.049 | 1        | 0        | 3        | 0.2    | 2850                         | 3311.94                    |
| [PentMIM]                                           | [NTf2]                | 0.180 | 0.043 | 1        | 0        | 0        | 0.6    | 2880                         | 3353.11                    |
| [DMIM]                                              | [BF4]                 | 0.184 | 0     | 0        | 0        | 0        | 0.3    | 2900                         | 4951.89                    |
| [ETT]                                               | [eFAP]                | 0.202 | 0.078 | 1        | 0        | 3        | 0.2    | 2940                         | 4176.34                    |
| [OMIM]                                              | [NTf2]                | 0.186 | 0.027 | 1        | 0        | 0        | 0.6    | 3000                         | 2599.97                    |
| [P66614]                                            | [Cl]                  | 0.230 | 0.01  | 0        | 0        | 0        | 0      | 3040                         | 5192.98                    |
| [HMPy]                                              | [NTf2]                | 0.175 | 0.032 | 1        | 0        | 0        | 0.6    | 3280                         | 3035.04                    |
| [BMIM]                                              | [NTf2]                | 0.184 | 0.056 | 1        | 0        | 0        | 0.6    | 3300                         | 3770.79                    |
| [B3MPy]                                             | [NTf2]                | 0.170 | 0.048 | 1        | 0        | 0        | 0.6    | 3300                         | 3766.29                    |
| [EMIM]                                              | [eFAP]                | 0.134 | 0.064 | 1        | 0        | 3        | 0.2    | 3323                         | 5120.54                    |
| [P66614]                                            | [NTf2]                | 0.228 | 0.01  | 1        | 0        | 0        | 0.6    | 3340                         | 1020.06                    |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [NTf2]                | 0.170 | 0.037 | 1        | 0        | 0        | 0.6    | 3370                         | 3338.79                    |
| [HMIM]                                              | [NTf2]                | 0.181 | 0.04  | 1        | 0        | 0        | 0.6    | 3500                         | 3214.63                    |
| [PMIM]                                              | [NTf2]                | 0.208 | 0.05  | 1        | 0        | 0        | 0.6    | 3700                         | 3012.33                    |
| [EMIM]                                              | [NTf2]                | 0.181 | 0.064 | 1        | 0        | 0        | 0.6    | 3770                         | 4147.36                    |
| [B4MPy]                                             | [NTf2]                | 0.169 | 0.053 | 1        | 0        | 0        | 0.6    | 3830                         | 3982.49                    |
| [DMPIM]                                             | [NTf2]                | 0.163 | 0.054 | 1        | 0        | 0        | 0.6    | 3850                         | 4152.68                    |
| [BMIM]                                              | [TA]                  | 0.182 | 0.056 | 0        | 2        | 0        | 0.8    | 3850                         | 12556.52                   |
| [BMPyrr]                                            | [NTf2]                | 0.159 | 0.049 | 1        | 0        | 0        | 0.6    | 3860                         | 4045.91                    |
| [1,4BzMPY]                                          | [NTf2]                | 0.151 | 0.048 | 1        | 0        | 0        | 0.6    | 3930                         | 4182.13                    |
| [EDMIM]                                             | [NTf2]                | 0.205 | 0.061 | 1        | 0        | 0        | 0.6    | 3960                         | 3505.49                    |
| [1,3BzMPY]                                          | [NTf2]                | 0.149 | 0.051 | 1        | 0        | 0        | 0.6    | 3980                         | 4342.50                    |
| [BzMIM]                                             | [NTf2]                | 0.185 | 0.058 | 1        | 0        | 0        | 0.6    | 3990                         | 3826.63                    |
| [BzMPyrr]                                           | [NTf2]                | 0.149 | 0.044 | 1        | 0        | 0        | 0.6    | 4000                         | 4070.45                    |
| [(N <sub>11</sub> ) <sub>2</sub> CH]                | [NTf2]                | 0.173 | 0     | 1        | 0        | 0        | 0.6    | 4070                         | 1835.18                    |
| [1,2BzMPY]                                          | [NTf2]                | 0.145 | 0.054 | 1        | 0        | 0        | 0.6    | 4170                         | 4546.63                    |
| [C3CNMIM]                                           | [NTf2]                | 0.200 | 0.09  | 1        | 0        | 0        | 0.6    | 4210                         | 4741.96                    |
| [N <sub>4,1,1,1</sub> ]                             | [NTf2]                | 0.169 | 0.055 | 1        | 0        | 0        | 0.6    | 4350                         | 4060.22                    |
| [HOEtMIM]                                           | [NTf2]                | 0.190 | 0.081 | 1        | 0        | 0        | 0.6    | 4420                         | 4611.06                    |
| [BzPY]                                              | [NTf2]                | 0.146 | 0.055 | 1        | 0        | 0        | 0.6    | 4610                         | 4563.61                    |
| [C3CNDMIM]                                          | [NTf2]                | 0.156 | 0.06  | 1        | 0        | 0        | 0.6    | 4620                         | 4539.06                    |
| [OMIM]                                              | [PF6]                 | 0.204 | 0.027 | 0        | 0        | 0        | 0.3    | 5140                         | 5563.48                    |
| [PMIM]                                              | [PF6]                 | 0.192 | 0.05  | 0        | 0        | 0        | 0.3    | 5200                         | 6719.98                    |
| [(N <sub>111</sub> ) <sub>2</sub> N]                | [C(CN) <sub>3</sub> ] | 0.189 | 0.037 | 0        | 0        | 3        | 0.3    | 5250                         | 5079.28                    |
| [BMIM]                                              | [PF6]                 | 0.191 | 0.056 | 0        | 0        | 0        | 0.3    | 5340                         | 6975.04                    |
| [OMIM]                                              | [BF4]                 | 0.204 | 0.027 | 0        | 0        | 0        | 0.3    | 5390                         | 5563.48                    |
| [BMIM]                                              | [N(CN) <sub>2</sub> ] | 0.178 | 0.056 | 0        | 0        | 1        | 0.2    | 5590                         | 7145.60                    |
| [BMIM]                                              | [BF4]                 | 0.191 | 0.056 | 0        | 0        | 0        | 0.3    | 5600                         | 6975.04                    |
| [P66614]                                            | [Br]                  | 0.232 | 0.01  | 0        | 0        | 0        | 0      | 5880                         | 5149.21                    |
| [BDMIM]                                             | [BF4]                 | 0.188 | 0.048 | 0        | 0        | 0        | 0.3    | 6100                         | 6729.79                    |
| [HOEtMIM]                                           | [TfO]                 | 0.169 | 0.081 | 0        | 0        | 0        | 0.5    | 6140                         | 7855.31                    |
| [BDMIM]                                             | [PF6]                 | 0.192 | 0.048 | 0        | 0        | 0        | 0.3    | 6180                         | 6642.25                    |
| [EMIM]                                              | [TfO]                 | 0.188 | 0.064 | 0        | 0        | 0        | 0.5    | 7400                         | 6778.79                    |
| [EMIM]                                              | [BF4]                 | 0.183 | 0.064 | 0        | 0        | 0        | 0.3    | 8110                         | 7461.04                    |
| [EMIM]                                              | [N(CN) <sub>2</sub> ] | 0.172 | 0.064 | 0        | 0        | 1        | 0.2    | 8210                         | 7587.83                    |
| [HOEtMIM]                                           | [PF6]                 | 0.171 | 0.081 | 0        | 0        | 0        | 0.3    | 8940                         | 8384.36                    |
| [HEA]                                               | [OAc]                 | 0.152 | 0     | 0        | 2        | 0        | 0.9    | 9500                         | 10750.35                   |
| [HOEtMIM]                                           | [BF4]                 | 0.167 | 0.081 | 0        | 0        | 0        | 0.3    | 10800                        | 8471.91                    |
| [C3CNMIM]                                           | [N(CN) <sub>2</sub> ] | 0.163 | 0.09  | 0        | 0        | 1        | 0.2    | 11060                        | 8795.26                    |
| [HMIM]                                              | [Cl]                  | 0.184 | 0.04  | 0        | 0        | 0        | 0      | 11800                        | 7365.67                    |
| [HEA]                                               | [HCOO]                | 0.135 | 0     | 0        | 2        | 0        | 0.2    | 13217                        | 13127.30                   |
| [MHEA]                                              | [HCOO]                | 0.152 | 0     | 0        | 2        | 0        | 0.2    | 13950                        | 12755.23                   |
| [BHEA]                                              | [OAc]                 | 0.150 | 0.07  | 0        | 2        | 0        | 0.9    | 14100                        | 13514.57                   |
| [HEA]                                               | [Lac]                 | 0.133 | 0     | 0        | 3        | 1        | 0.9    | 14200                        | 14174.09                   |
| [HEEA]                                              | [Lac]                 | 0.153 | 0.076 | 0        | 3        | 1        | 0.9    | 16214                        | 16689.99                   |



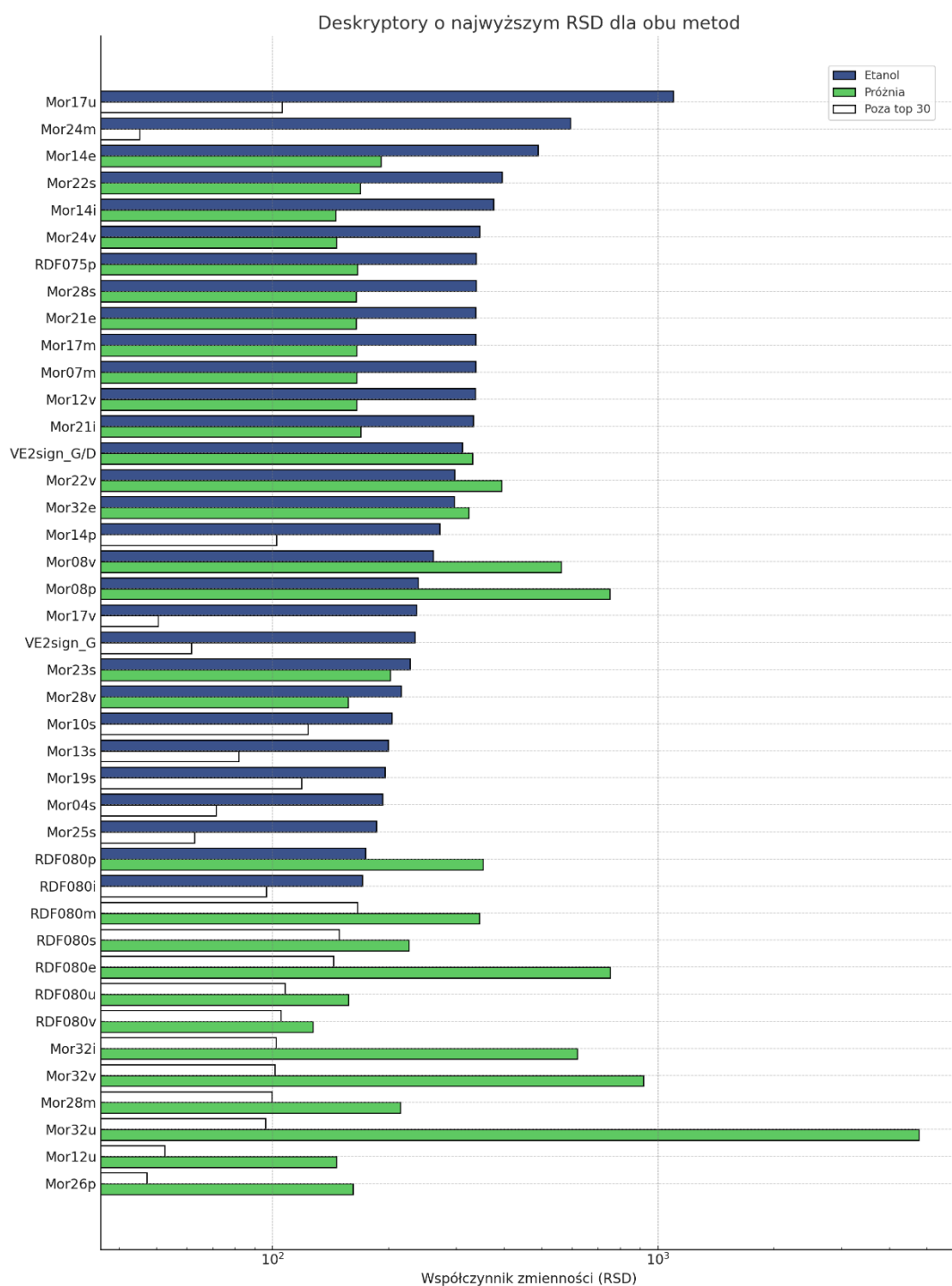
**Rysunek A1.** Histogramy różnic procentowych cieczy z anionem [NTf<sub>2</sub>]: a) [BzMPyrr], b) [BzMIM], c) [BMPyrr], d) [B4MPy], e) [EDMIM], f) [HMIM], g) [HMPy], h) [OMIM], i) [PentMIM], j) [1,2BzMPy]



**Rysunek A2.** Histogramy różnic procentowych cieczy z anionem [NTf<sub>2</sub>]: a) [DMPIM], b) [HOEtMIM], c) [1,3BzMPY], d) [C3CNMIM], e) [1,4BzMPY], f) [PhDBIM], g) [BzPY], h) [C<sub>6</sub>H<sub>4</sub>F<sub>9</sub>MIM], i) [C<sub>8</sub>H<sub>4</sub>F<sub>13</sub>MIM], j) [EMIM]



**Rysunek A3.** Histogramy różnic procentowych cieczy z anionem  $[\text{Ntf}_2]$ : a)  $[(\text{N}_{11})_2\text{CH}]$ , b)  $[(\text{N}_{111})_2\text{N}]$ ,  $[\text{N}_{4,1,1,1}]$



**Rysunek A4.** 30 deskryptorów o największej zmienności dla każdej z metod. Kolorem niebieskim oznaczono optymalizację w etanolu, kolorem zielonym optymalizację w próżni. Na biało zaznaczone są wartości współczynnika zmienności danego deskryptora dla przeciwnej metody optymalizacji, jeśli znajdują się one poza zbiorem 30 najwyższych wartości.