С Т А Н О В И Щ Е

от доц. д-р **Иванка Цаковска**,
Институт по биофизика и биомедицинско инженерство - БАН

относно: дисертационен труд на **Веселина Христева Паскалева**

на тема: **„Компютърно генериране на тавтомери на органични съединения и приложения в химичната информатика"**

за присъждане на образователната и научна степен „**Доктор**"
Научна област „**Природни науки, математика и информатика**"
Професионално направление „**Химически науки**"
Докторска програма „**Теоретична химия**"

Научен ръководител: доц. д-р **Николай Кочев**

Представеният ми за становище дисертационен труд е в областта на **химичната информатика** – интердисциплинарно направление, прилагащо подходи на информатиката при решаване на различни химични проблеми. **Целта на изследването** е разработване на алгоритми за третиране на тавтомерията при обработката на структурна информация – **актуална задача**, върху която се работи все по-сериозно в полето на *in silico* моделирането.

Дисертационният труд има класическа структура, започвайки с въведение, последвано от литературен обзор. **Литературният обзор** представя различните аспекти на химичната информатика. По-нататък се фокусира върху тавтомерията и важността на нейното отчитане за коректността на *in silico* изследванията. Направен е преглед на голям набор софтуерни програми за генериране на тавтомери. Тук се наблюдава известна хаотичност в представянето. Например коментирани са накратко множество софтуерни системи, но липсват Mn.Tautomer и ChemAxon Marvin, които са използвани при външните тестове в Глава 4. Като цяло, докторантът показва добро познаване на състоянието на изследванията в научната област.

Методиката на дисертационното изследване и резултатите са представени в три глави, включващи съответно разработените алгоритми за генериране на тавтомери, тестването им, както и приложението им в различни аспекти на *in silico* моделирането. Цялостното впечатление е за разбиране и адекватно третиране на проблема. Тук, както и в литературния обзор се наблюдава известна хаотичност в представянето, както и доста технически

грешки при оформянето на работата. Препоръката ми е това да се има предвид и коригира особено ако (както е редно) дисертацията се направи публична за научната общност. Доброто оформяне на резултатите от едно научно изследване е важна и неделима част от него.

Приносите от дисертационния труд бих формулирала в две направления:

(i) **Научни приноси**, състоящи се в разработване на алгоритми за изчерпателно генериране на тавтомери, разработване на система за тяхното ранкиране, предлагане на теглова схема за модифициране на молекулните дескриптори в зависимост от тавтомерията, както и стратегия за включването на тази информация в QSAR моделирането.

(ii) **Научно-приложен принос**, състоящ се в имплементиране на разработените алгоритми в софтуерното приложение Ambit-Tautomer.

Към докторанта имам следните **въпроси**:

1. Под каноничен тавтомер, в разработената от авторите система, се разбира този с най-ниско енергийно тегло. В същото време, в литературния обзор при описание на един от използваните софтуери е казано, че „Изчисляването на каноничната форма се основава на емпирични правила и теглова схема (scoring scheme). Каноничният тавтомер не винаги е химически най-резонният или този с най-ниска енергия“. Тези формулировки предполагат различно разбиране за каноничен тавтомер в различните софтуерни системи. При това положение, обосновано ли е сравнението на каноничните тавтомери, генерирани в различни програми, направено в Секция 4.2.3.2?

2. Използваните при различните тестове извадки се състоят от структури, дефинирани по различен начин: Пример: извадки 7 и 8 са съответно: съединения с известна токсичност и съединения, използвани в селското стопанство, което обуславя висока вероятност от припокриване на химични пространства. Направена ли е проверка за дублиране на структури между извадките?

3. Секция 5.3 изследва влиянието на тавтомерията върху методите за търсене на структурно подобие. Коментират се различни методи за търсене на подобие – базирани на молекулни дескриптори, пръстови отпечатъци, най-голяма обща подструктура. Може ли да бъде откроен такъв, който би подпомогнал идентифицирането на наличните тавтомери на дадена структура в голяма база данни?

4. Предвид огромния размер на представените резултати логично е да има екипен подход в тяхното продуциране. Може ли докторантът да очертае своя личен принос в отделните етапи на изследването?

Наукометричните показатели на докторанта включват 3 публикации, една от които в авторитетното списание Molecular informatics (ИФ = 2.013 в годината на публикуване). Тази публикация е цитирана вече 10 пъти, което само по себе си говори за интереса към изследването. Има подобаващ брой участия в научни форуми (8), като в шест от тях Веселина е представящ автор.

Заключение: Дисертационният труд представлява впечатляващо по своя обем научно изследване в една модерна и перспективна област. Добро впечатление прави наличието на идеи за последващи разработки, които без съмнение са добър стимул за бъдещо развитие на един млад учен с доказан вече потенциал. Авторефератът отразява коректно съдържанието на дисертационния труд. Наукометричните параметри на кандидата покриват количествените критерии за развитие на академичния състав на Химическия факултет при Пловдивския университет ”Паисий Хилендарски”. Представените за защита документи са според изискванията на ЗРАСРБ и Правилника за неговото прилагане. На това основание давам **положителна оценка** и убедено подкрепям присъждането на образователната и научна степен «**Доктор**» на **Веселина Христева Паскалева**.

София,

15.08. 2016 г. доц. д-р Иванка Цаковска