

## РЕЦЕНЗИЯ

от д-р Иван Петков Бангов, професор, Шуменски университет “Епископ Константин Преславски”

на дисертационен труд за присъждане на научна степен 'доктор на науките'

област на висше образование: 4. Природни науки, математика и информатика.

професионално направление: 4.2. Химични науки (Аналитична химия)

**Автор:** доцент д-р **Пламен Николов Пенчев** - Пловдивски университет “Паисий Хилендарски”.

**Тема:** КОМПЮТЪРНА ИНТЕРПРЕТАЦИЯ НА МОЛЕКУЛНИ СПЕКТРИ С ЦЕЛ РАЗКРИВАНЕ НА СТРУКТУРАТА НА ОРГАНИЧНИ СЪЕДИНЕНИЯ

### 1. Предмет на рецензиране:

Със заповед № Р33-2019 от 16.05 2016г. на Ректора на Пловдивския университет „Паисий Хилендарски“ (ПУ) съм определен(а) за член на научното жури за осигуряване на процедура за защита на дисертационен труд на тема

*КОМПЮТЪРНА ИНТЕРПРЕТАЦИЯ НА МОЛЕКУЛНИ СПЕКТРИ С ЦЕЛ РАЗКРИВАНЕ НА СТРУКТУРАТА НА ОРГАНИЧНИ СЪЕДИНЕНИЯ*

за придобиване на научната степен 'доктор на науките' на ПУ в област на висше образование 4. Природни науки, математика и информатика, професионално направление 4.2. Химични науки (Аналитична химия). Автор на дисертационния труд е доцент д-р **Пламен Николов Пенчев**.

Катедра “Аналитична химия и компютърна химия” към Факултет/Институт по Химически факултетна ВУЗ Пловдивски университет “Паисий Хилендарски”

Представеният от доцент д-р **Пламен Николов Пенчев** комплект материали на хартиен носител е в съответствие с Чл.45 (4) от Правилника за развитие на академичния състав на ПУ, включва следните документи:

- молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационен труд - да
- автобиография в европейски формат - да
- копие от диплома за образователната и научна степен „доктор” - да
- протоколи от катедрени съвети, свързани с откриване на процедурата и с предварителното обсъждане на дисертационния труд - да
- дисертационен труд - да
- автореферат - да
- списък на научните публикации по темата на дисертацията - да
- копия на научните публикации - да
- декларация за оригиналност и достоверност на приложените документи - да
- справка за спазване на специфичните изисквания на съответния факултет - да
- .....

Дисертантът има 24 публикации и Дисертационен труд

## **2. Кратки биографични данни**

Образователната и научно-преподавателска дейност на докторантът Пламен Пенчев е както следва:

**1980г.:** завършил средното си образование в Националната математическа гимназия “Акад. Л. Чакалов”-София, в паралелка химия (1977-1980г.);

**1986г.:** завършил висшето си образование във ВХТИ-София (1982-1986г.);

**1989г.:** завършва магистратура в ПУ “Паисий Хилендарски” (1986-1989г.);

**1989 -1996г.:** химик в катедра “Аналитична химия и компютърна химия”;

**1995-2001г.:** асистент, старши асистент и главен асистент в катедра “Аналитична химия и компютърна химия”;

**1998г.:** защитава ОНС Доктор, с диплома No 26061, от 23.04.1999г.;

**2001-2006г.:** Postdoctoral Research Associate, Department of Chemistry and Biochemistry, ASU, Arizona, USA;

**2006-2010г.:** главен асистент в катедра “Аналитична химия и компютърна химия”;

**2010г. - до сега:** доцент в катедра “Аналитична химия и компютърна химия”;

## **3. Актуалност на тематиката и целесъобразност на поставените цели и задачи**

Разработваният в дисертационния труд проблем се отнася към интердисциплинарната област “Хемоинформатика”. Това е едно направление, което обединява химичните науки с информационните технологии. Въпреки, че то е възникнало още през 60-те години с навлизането на компютрите в различните области на науката и практиката под името “Изкуствен интелект” (един от първите проекти беше “Евристичния Дендрал” на Стенфордския университет-САЩ, ръководен от професор Карл Джераси), тази област остава изключително актуална и днес с развитието на модерните компютърни технологии и тяхното внедряване в практиката. Проблемът, който се решава е как с помощта на компютъра да се определи структурата на дадено органично съединение, на основата на разнообразна спектрална информация. Решаването на този проблем има пряко практическо отношение към *Research and development* (R&D) изследванията в химическата и най-вече фармацевтичната промишленост. Трябва да се отбележи, че съвременните химични технологии не се основават на случайно открити вещества с търсените свойства. Последните се търсят систематично в огромни бази от данни (със стотици хиляди до милиони електронно кодирани съединения). На базата на подобни съединения, роботи синтезират едновременно стотици органични химични вещества. Компютърът трябва да даде бърз отговор, каква е структурата на тези вещества на основата на спектралната информация, получена от мас-газ спектрометрия, инфрачервена (ИЧ) и различните

видове ЯМР ( $H^1, C^{13}$ ) спектроскопии. Именно методите, разгледани и използвани в дисертацията, водят до тази цел. В тази дисертация са разгледани и използвани главно методите за търсене в бази от данни, съдържащи спектрална/структурна информация.

#### 4. Познаване на проблема

Дисертацията представлява обширен труд, както като изработка, така и като текст. Тя се състои от 371 страници. Общата част - *Литературен обзор* се състои от 107 страници. В него са разгледани всички налични подходи към решаването на проблемите, залегнали в дисертацията. Цитирани са 815 статии, 60 книги и справочници и 19 софтуерни продукта, което показва пълно запознаване с наличната информация по проблема и огромна ерудиция на дисертанта. Според мен литературният обзор покрива цялата научна информация по този проблем. Трябва също да се отбележи, че познаването на проблема, залегнал в дисертацията ясно личи от Глава 2. “Спектроскопски измервания и софтуер”. В тази глава са изнесени начините за добиване на различните видове информация от Raman, ATR, FT-IR, UV-VIS,  $^{13}C$ -NMR, различните 2D NMR спектри. Същото важи и за софтуера, който е разгледан в тази глава. Разгледани са основно редици софтуерни продукти, които третират спектралната информация, като **ToSim**, разработена от Warmuza и Scsibrani, програмата **Excerpt**, която съдържа редица хемоинформационни алгоритми, **MolGen**, програма за генериране на структури, разработена от групата на професор Kerber. **MSClass**, разработена от групата на проф. Warmuza. **CACTUS** на доц. Inenfeild за рисуване на химичните структури и трансформиране на различните структурни формати. Дискутирани са също други програми, които не са използвани в дисертацията, като **Euristic Dendral** на Университета в Станфорд-САЩ, **Darc** на Дюбоа и сътрудници, **C-SEARCH** на Робинен - Виена. И в двете последни разработки е използван **HOSE** кода на Бремза. Не е дискутирана подобаващо системата **CHEMICS** на групата на проф. Сасаки от Техническият университет в Тойохаши, Япония, въпреки че техните работи са цитирани, както и работите на групата на академик Зефиоров, също цитирани. Основно са разгледани методите на изкуствените невронни мрежи и програмите **Lora** и **SAS**. Разгледани са методите за подструктурно търсене. Може би за още по-голяма пълнота биха могли да бъдат споменати методите за структурна генерация на групата на проф. Квасничка от Чехословакия и разработките на Новосибирската група. Независимо от тези дребни забележки, дисертацията демонстрира голямата ерудиция на дисертанта в тази насока.

#### 5. Методика на изследването

Основна цел на представената ми за рецензия дисертация е разработката на една комплексна система, състояща се от бази от спектрални данни – УВ-ВИД,  $^{13}C$ -ЯМР, Раман, мас-газ, както и разработени от автора и използвани програми, включващи разнообразни алгоритми, като алгоритми за търсене в тези спектрални библиотеки, регресионен анализ, определяне на максимална подструктура, и класификация на молекулните спектри чрез използване на различни класификационни методи: линеен дескриминационен анализ, методът на най-близките съседи,

изкуствените невронни мрежи, максималната обща подструктура. Според автора само съвместното използване на тези методи може да даде надеждни резултати. Тази система се състои не само от автоматично извличане на спектралните данни от спектралните бази с оглед на наличната (query) информация, но и в даването на конкретна и надеждна информация относно химичната структура на изследваните химични съединения.

## **6. Характеристика и оценка на дисертационния труд**

Както започнах тази рецензия, дисертационният труд представлява една сериозна и обширна разработка на един сериозен и интересен проблем в модерната област на експертните системи за определяне на химичната структура на органични съединения на основата на спектрална информация от различни източници. Преди тази област се наричаше Изкуствен интелект в химията и използването на изкуствените невронни мрежи оправдава напълно и това наименование. Дисертационният труд се базира на най-модерните подходи на Хемоинформатиката. Освен, че дисертацията има методологично значение, разработения подход в нея има изключително практично значение.

## **7. Приноси и значимост на разработката за науката и практиката**

Разработката на нови кандидат-химични съединения с определени експлоатационни или биологично-активни свойства в съвременната химична и фармацевтична индустрия се основава на активно компютърното търсене в огромни химични бази, съставени от структурни и спектрални данни. Именно алгоритмите залегнали в дисертационния труд на доцент Пенчев се използват в практиката на търсене на такива съединения. При разработката на тази дисертация се използват както бази от данни и софтуер взети от други лаборатории, така и създадени или модифицирани от доц. Пенчев, като по-голямата част от софтуера е оригинални разработки и програмиран от доц. Пенчев на Делфи 2-5 за Windows . Тук трябва да се споменат следните оригинални разработки:

1. търсене в библиотеки с УВ и  $^{13}\text{C}$  ЯМР спектри;
2. автоматична интерпретация на  $^{13}\text{C}$  ЯМР спектри;
3. използване на структурната информация, извлечена от мас спектри и класификация на мас спектрите по структурни фрагменти;
4. подобряване на резултатите на метода на търсене на максимална подструктура;
5. пълно отнасяне на  $^1\text{H}$  и  $^{13}\text{C}$  ЯМР спектри на изолирани природни съединения.

Програмните продукти, които дисертантът е създал са:

1. програма за търсене в база от вибрационни спектри IRSS;
2. програма за търсене и интерпретация на  $^{13}\text{C}$  ЯМР спектри InferCNMR;
3. програма, използваща изкуствените невронни мрежи NueNet;
4. програм а за търсене на УВ-ВИД спектри UVLib;

В тези програми са заложили редица оригинални алгоритми и оригинални критерии за сравняване на спектри. Въпреки че те представляват отделни постижения, тук няма място да бъдат дискутирани. Направена е експериментална проверка на ефективността на тези алгоритми критерии.

### 8. Преценка на публикациите по дисертационния труд.

Авторът е дал следната таблица:

Категория	Брой
Научни статии по дисертацията	24
- статии в списания с импакт фактор	12 (4, 9)
- статии в български списания, които са цитирани	3 (0,3)
- международни специализирани научни списания	4 (0,2)
Общо за трите горни категории	19 (4,14)
Изисквания на ХФ, които включват трите горни категории	15
- в български научни списания	5 (0,1)
Изисквания на ХФ	5
Цитати на статиите по дисертацията	103
Изисквания на ХФ за цитати	40
Цитатите са последно обновени през април 2016.	109

Публикационната дейност на кандидата се състои само в статии. От таблицата се вижда, че като наукометрични показатели дисертацията напълно задоволява и надхвърля минималните изисквания на Факултета по химия към Пловдивския университет „Паисий Хилендарски“. От публикациите, текстовете, на които са предоставени, се вижда, че повечето са в специализирани реномирани списания, като Spectroscopy Letters, J. Mol. Structure, J. Chem. Inf. Comput. Sci., както и в изданието с новото му наименование Chemical Information and Modeling (едно от най-представителните списания на Американското химическо дружество в тази област), Vibrational Spectroscopy, Analytica Chimica Acta, Computers and Chemistry, Analytical Sciences, Natural Product communications (2 статии) и Progress in Nuclear magnetic Resonance Spectroscopy, 4 статии Asian Chemical Communications, 2 статии в Bulgarian Chemical Communications и останалите статии са в сборници на Пловдивския и Русенския университет (1 статия). Кандидатът е взел участие в 14 научни мероприятия, като в 3 е изнесъл лекция (едната е на мероприятие в щата Джорджия-САЩ). За високата научна стойност на научната продукция на доц. Пламен Пенчев говори факта за 109 цитирания.

### 9. Лично участие на автора

Разбира се повечето от статиите са в колектив, като тук искам да спомена плодотворното сътрудничество с професор Вармуца от Виенския технически университет, както и с проф. Андреев от Пловдивския университет. Обаче за мен не е трудно да разпознавам оригиналните разработки на доц. Пенчев в това сътрудничество. Професор Вармуца е предоставил някои програми, бази от данни. Доц. Пенчев обаче, ги модифицира, усъвършенства и така се получило едно добро сътрудничество. Както беше споменато по-горе важни програмни продукти за работата по тази

дисертация са създадени и кодирани от дисертанта. От цялата дисертация се вижда, че това е един самостоятелен труд на доц. Пенчев. И тук искам да отбележа, целенасочения характер на този труд. Нямаме сбор от статии с най-разнообразна тематика. За съжаление липсват публикации от престоя на доц. Пенчев в лабораторията на проф. Мънк в Arizona State University. Но това допълнително потвърждава характера на дисертацията, като един напълно оригинален труд на автора.

## **10. Автореферат**

Авторефератът е с размер на една нормална дисертация и дава точна картина на съдържанието на дисертацията.

## **11. Критични забележки и препоръки**

Обикновено рецензентите се чувстват задължени да намерят някои критични забележки. Аз ще се въздържа от това.

## **12. Лични впечатления**

Аз познавам доцент Пенчев отдавна. През целият период на своето научно развитие той работи упорито за решаването на проблемите, дискутирани в дисертацията. В своето развитие той е натрупал голяма ерудиция и се откроява, като един от водещите специалисти в тази област.

## **13. Препоръки за бъдещо използване на дисертационните приноси и резултати**

Аз съм говорил с дисертанта за едно бъдещо инкорпориране на тези инструменти за третиране на спектралната информация към нашата програма за генериране на структури и използването на дескрипторните фингърпринти за тази цел.

## **ЗАКЛЮЧЕНИЕ**

Дисертационният труд съдържа важни научни, научно-приложни и приложни резултати, които представляват оригинален принос в науката и отговарят на всички изисквания на Закона за развитие на академичния състав в Република България (ЗРАСРБ), на Правилника за прилагане на ЗРАСРБ и съответния Правилник на ПУ „Паисий Хилендарски“. Представените материали и дисертационни резултати напълно съответстват на специфичните изисквания на Факултета по Химия, приети във връзка с Правилника на ПУ за приложение на ЗРАСРБ.

Дисертационният труд показва, че дисертантът Пламен Николов Пенчев притежава задълбочени теоретични знания и професионални умения по научната специалност „компютърна химия“ и „хемоинформатика“, като демонстрира качества и умения за провеждане на изследвания с получаване на оригинални и значими научни приноси.

Поради гореизложеното, убедено давам своята положителна оценка за проведеното изследване, представено от рецензираните по-горе дисертационен труд, автореферат, постигнати резултати и приноси, и предлагам на почитаемото научно жури да присъди научната степен ‘доктор на науките’ на доцент Пламен Николов Пенчев в областта на висшето образование: 4. Природни науки, математика и информатика, професионално направление 4.2. Химични науки (Аналитична химия).

11.07.2017 г.

Рецензент: проф. дхн Иван Петков Бангов

.....